1:1 MOTIF FOR DNA RECOGNITION BY β-ALANINE-LINKED POLYAMIDES

Thesis by

Adam Robert Urbach

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

2002

(Submitted May 17, 2002)

To my Father

Acknowledgements

I would like to thank my Ph. D. advisor, Professor Peter Dervan, for giving me just the right mixture of freedom and guidance in order to pursue my interests and yet to turn my energy into useful science. The training I've received has far exceeded anything I could have imagined just a few years ago. I am particularly grateful to Professor Dervan for supporting my decision to leave Caltech during my second year and then taking me back into his group with enthusiasm and compassion. My thesis committee has taught me, among other things, the important lesson of how to better approach the writing and defending of research proposals. I am grateful to Professors Robert H. Grubbs, John D. Roberts, Richard W. Roberts, and Stephen L. Mayo for these lessons as well as helpful conversations over the years. Special thanks goes to Professor John D. Roberts for the privilege of his discerning attention during this past year.

The work of Professor Uli Laemmli provided much inspiration for the theses presented here. I look to Professor Laemmli with great admiration and respect as a scientist, and I am grateful to him for advocating my work during his visit to Caltech. Much appreciation goes to Professor John Love and Dr Scott Ross for teaching me macromolecular NMR and being supportive throughout our fruitful collaboration. Professors David Case and Tammy Dwyer provided many helpful suggestions that led to significant improvements in the structure calculations, and their generous input is most appreciated. During my last year at Caltech, I had the pleasure of collaborating with Michael "Meaty" Marques and Ray Doss on several interesting projects. I am very grateful to Michael and Ray for helping me finish research while I was writing props.

I wish to acknowledge an extraordinary group of colleagues in the Dervan lab who have managed somehow to put up with me over the years. David Herman is the embodiment of positive vibe, infecting all who know him. David's friendship has helped to keep me sane during many long nights of James Brown in the computer room–- adventures to the Goodland and afternoons with Abe will never be forgotten, not to mention learning to surf in overhead conditions. Thanks to John Trauger for showing me that ego is not a necessary component of ambition or success. I am grateful to John for his friendship and for generously sharing his apartment when I needed a place to live. Much appreciation goes to Ben Edelson for many stimulating discussions and for proofreading numerous works of literature. Ben's integrity and humility are an exceptional example to follow. Thanks to Will Greenberg and Dave Liberles for being excellent labmates and for many good times with Gamesa. I thank Ken Brameld for getting me interested in Himalayan trekking. Thanks to Clay Wang for great dinner parties and to Meredith Howard for lively conversations. Victor Rucker has made life in Church 308 memorable, and I thank him for being himself. I thank Shane Foister and Ramez Elgammal for much hospitality, sometimes southern. Roland Bürli and Ralf Jäger provided some much needed help with my candidacy props. I thank Aileen Chang, Jason Belitsky, Adam Kerstien, Amanda Cashin, Eric Fechter, Philipp Weyermann, Christoph Briehn, Leonard Prins, Anna Mapp, Paul Floreancig, Tom Minehan, Christian Melander, Bogdan Olenyuk, and (last but certainly not least) Bobby Arora for making the Dervan group a more pleasant place to live.

The staff at Caltech is very accommodating, and I would especially like to thank Dian Buchness, Chris Smith, Margot Hoyt, Lynne Martinez, Lindy Alo, Tom Dunn, and Steve Gould for their exceptional help over the years. I thank Darryl Willick for saving me from computer disaster on several occasions.

Along the path that led to Caltech, I thank David Arnold for taking in a problem child; Professor Jonathan Sessler for introducing me to organic chemistry and chemical research; and Faiz Kayyem and Cindy Bamdad for helping me find the path back to grad school.

Life is meaningless without friends and family. The enormous support provided by this special group has made my life complete. Thanks to Mimi, Papa, Mom, Jim, Rene, Joe, Lucy, Siggy, and Susan for their support and understanding. I thank Bryn, Allison, Dian, Mark, Kutty, Denise, Ashley, and the Herman family for making me part of their family. I am grateful to Dr Art Herman for solid advice when I needed it. I have had the privilege of getting to know Professor Michael Waring during my time at Caltech. Michael's support during my candidacy made a world of difference, and I will always be grateful. I thank Professor David Laude for his friendship and lessons outside the box. I am grateful to Sifu Ken Edwards for teaching me Tai Chi.

There are three people whom I would especially like to acknowledge for their endless support and friendship. Rudy Emmelot has been an outstanding friend to me these past few years. Rudy's generosity is sometimes overwhelming, and I cannot in these short words express to him the full extent of my gratitude for his companionship both at home and traveling across the globe. I could only be half the person I am without the love and support of my wonderful wife, Dana, whose support and devotion mean everything to me. I look forward most to our lives together. To my father, to whom this thesis is dedicated, goes my deepest respect and appreciation for his friendship, support, and countless self-sacrifices from the very beginning.

Abstract

Polyamides composed of N-methylpyrrole (Py), N-methylimidazole (Im), and 3-hydroxypyrrole (Hp) amino acids linked by beta-alanine (β) bind in the minor groove of DNA in 1:1 and 2:1 ligand:DNA complexes. Although the energetics and structure of the 2:1 motif have been explored extensively, there is remarkably less understood about 1:1 recognition beyond the initial studies on netropsin and distamycin. Laemmli and coworkers used β-linked polyamides, which bind in a 1:1 motif, to effect phenotypic changes in *Drosophila melanogaster*. The thesis work described here investigates Laemmli's 1:1 motif in order to further understand and exploit this novel mode of DNA recognition.

By selectively replacing Py residues with β it was found that the Im-β-Im subunit is important for high-affinity binding in 1:1 and 2:1 modes. This study also demonstrates that a single ligand can target very different DNA sequences based on 1:1 or 2:1 binding. This ambiguity of sequence targeting based on stoichiometry was addressed. It was discovered that hairpin and 1:1 binding modes, which are dependent on ligand conformation, are controlled by changing the linker between polyamide subunits.

The possibility of developing a 1:1 recognition code was explored by selectively mutating polyamide residues and DNA base pairs and comparing the association constants for the resulting complexes. It was found that Im residues tolerate all four Watson-Crick base pairs; Py and β residues are specific for A•T and T•A base pairs; and Hp specifies a single base pair, A•T, in the sequence context 5'-AAAGAGAAGAG-3'. Attempts to improve upon this recognition code using novel heterocyclic amino acids, such as furan, thiophene, thiazole,

and hydroxythiophene, are presented. The sequence-dependence of ligand orientation and the effect of ligand size on binding affinity were also explored.

The NMR structure of a 1:1 polyamide:DNA complex was determined. It reveals B-form DNA with a narrow minor groove and large negative propeller twist, which are shown to be stabilized by bifurcated hydrogen bonds between polyamide NH groups and purine N3 and pyrimidine O2 atoms. The first direct evidence is provided for hydrogen bond formation between Im-N3 and guanine NH2 in the 1:1 motif, thus confirming the original lexitropsin model.

Table of Contents

page

List of Figures and Tables

INTRODUCTION page

RESULTS AND DISCUSSION