

4-19-2018

# DNA Binding Kinetics of Large Antiviral Hairpin Polyamides

Jacquelyn Niederschulte  
jhhfc@mail.umsl.edu

Follow this and additional works at: <https://irl.umsl.edu/dissertation>

 Part of the [Biochemistry Commons](#), and the [Biophysics Commons](#)

---

## Recommended Citation

Niederschulte, Jacquelyn, "DNA Binding Kinetics of Large Antiviral Hairpin Polyamides" (2018). *Dissertations*. 739.  
<https://irl.umsl.edu/dissertation/739>

This Dissertation is brought to you for free and open access by the UMSL Graduate Works at IRL @ UMSL. It has been accepted for inclusion in Dissertations by an authorized administrator of IRL @ UMSL. For more information, please contact [marvinh@umsl.edu](mailto:marvinh@umsl.edu).

# DNA Binding Kinetics of Large Antiviral Hairpin Polyamides

Jacquelyn Niederschulte

M.S., Chemistry, University of Missouri – St. Louis, 2016

B.S. Chemistry, Lindenwood University, 2013

A Dissertation Submitted to The Graduate School at  
The University of Missouri – St. Louis  
In partial fulfillment of the requirements for the degree of  
Doctor of Philosophy in Chemistry with an emphasis in Biochemistry

May 2018

Advisory Committee

Cynthia Dupureur, PhD  
Chairperson

James Bashkin, PhD

Wesley Harris, PhD

Michael Nichols, PhD

Copyright, Jacquelyn Niederschulte, 2018

## Abstract

Although vaccines exist for the some of the most problematic strains of human papillomavirus (HPV), a double stranded DNA virus, there is currently no cure. HPV remains one of the most commonly sexually transmitted infections and is responsible for virtually all cervical cancers and genital warts. Natural products Distamycin A and netropsin have inspired the hairpin N-methylpyrrole (Py)/N-methylimidazole (Im) polyamides (PAs) studied here. The larger hairpin PAs, designed to bind to sites of 10 or more DNA bp, have been shown to be effective antivirals against oncogenic HPV strains 16, 18, and 31, while smaller hairpin PAs are not. Despite significant differences in potencies among the PAs tested, the PAs bind to DNA with similar binding affinity ( $K_d$ ). Drug discovery groups have historically used  $K_d$  as an indicator of drug efficacy. However, ample evidence has shown dissociation rate constants ( $k_{off}$ ) may be a better indicator of drug efficacy. While  $K_d$  is a function of association rate constant ( $k_{on}$ ) and  $k_{off}$ , respectively ( $K_d = k_{off}/k_{on}$ ), few studies have focused on the DNA binding kinetics of large hairpin PAs.

We are using fluorescence and CD spectroscopy to characterize binding affinities, obtain DNA binding kinetic rate constants, and determine binding stoichiometries as a function of PA size and N-terminal functional groups.  $K_d$  remains tight (low nM) for all PAs tested (6-20-rings) with our fluorescence assay, which is consistent with what is seen with other methods. The large PAs are characterized by slow DNA dissociation rates with half-lives ranging from 20-30 min, and dissociation slows as the size of PAs increases. The slow dissociation rates are likely the source of the difference in antiviral behavior. No correlation between antiviral efficacy and kinetic rate constants is observed when comparing PAs with different N-terminal groups.

Fluorescence and CD spectroscopic experiments indicate that saturation of the DNA does not occur until 2 or more equivalents of PA are added. Additionally,

association time courses for a 14-ring PA are multiphasic. Phases were determined to be concentration dependent indicating that they correlate to individual binding events. Thus, DNA association kinetics data and PA-DNA binding stoichiometry data are consistent with each other in that more than one PA binds to the DNA.

These results have profound implications on the antiviral mechanism of large hairpin PAs. The slow dissociation of multiple equivalents of PA bound to DNA may cause a large and prolonged disruption in DNA conformation, which in turn elicits or alters the DNA damage response, which is an integral part in the antiviral mechanism of large antiviral hairpin PAs and of the lifecycle of HPV.

## **DEDICATION**

To my husband, Dan, and daughter, Cordi.

## **ACKNOWLEDGEMENTS**

I am most grateful to my advisor, Dr. Cynthia Dupureur. Her support and friendship throughout my time here at UMSL have been invaluable. Her endless guidance, enthusiasm, and (most importantly) patience throughout my research and writing process were crucial to my success. I would also like to thank Dr. James Bashkin for his advice and scientific insights. I am also grateful to Dr. Michael Nichols for all his help with the Jasco. I would also like to thank the members of my committee for reading my thesis and for their continued support.

I am also extremely grateful to my current lab mate Yang Song who has been one of my closest friends throughout my time here at UMSL. His personal and professional insight and assistance has been greatly appreciated. I am also thankful to other current and former Dupureur group members: Elena Vasilieva, who helped to train me, Hyung Park and Kristin Bales, whose help was instrumental to the writing of this dissertation, and Helena Spikes. I am also appreciative to previous Bashkin group members whose research provided valuable insight into my own work: Gaofei He, Carlos Castaneda, Edith Csiki-Fejer, Jose Scuderi. And thanks to NanoVir for the materials provided for this study.

I am also grateful for my parents, Gina Ontai and Wayne Koizumi, my brother Ken, and my closest friends Emma Harris, Dani Chu, and Kori Kim. Thank you for your love and support, and for not understanding what I do but asking me anyway (and for pretending to find it interesting).

Last, but not least...words will never be able to properly describe how truly thankful I am to my husband. I could not have done this without his endless support. I am thankful for his ability to comfort a screaming baby while she watched me leave for work every day. I appreciate his never being mad at me for working late when I said I would be home early and his willingness to repeat something he told me during a moment where I was only half listening because I was preoccupied with science. He has been my rock despite the impossible situations he has found himself in over the past 5 years.

## TABLE OF CONTENTS

	Page
ABSTRACT	i
DEDICATION	iil
ACKNOWLEDGEMENTS	iv
TABLE OF CONTENTS	vi
LIST OF FIGURES	xi
LIST OF TABLES	xv
LIST OF SCHEMES	xvii
LIST OF ABBREVIATIONS	xviii
CHAPTER 1	INTRODUCTION
1.1. Structure of DNA	1
1.2. DNA binding ligands	4
1.2.1. Major groove binders	7
1.2.2. Minor groove binders	7
1.2.2.1. Non-therapeutic minor groove binders	7
1.2.2.2. Distamycin A and netropsin	8
1.3. N-methylpyrrole and N-methylimidazole polyamides	11
1.3.1. Hairpin polyamide structure and $\gamma$ -aminobutyric acid	11
1.3.2. $\beta$ -alanine as a flexible linker	14
1.3.3. N-methylpyrrole and N-methylimidazole	14
1.3.4. C-terminal functional groups (Tails)	19
1.4. Applications of polyamides	19
1.4.1. Polyamides and gene regulation	19
1.4.2. Polyamide antiviral applications	20
1.5. Human papillomavirus	20
1.51. Previous cell biology work with anti-HPV polyamides	23
1.5.1.1. Characterization of anti-HPV potency	23
1.5.1.2. HPV and the DNA Damage Response (DDR)	26
1.5.2. Previous biophysical work on anti-HPV polyamides	27



1.6. DNA-ligand binding kinetics	30
1.6.1. Basics of ligand binding kinetics	30
1.6.2. DNA binding kinetics of Distamycin A and Netropsin	32
1.6.3. Previous studies of hairpin PA-DNA binding kinetics	35
1.6.3.1. Kinetics studies of linked versus unlinked polyamides	35
1.6.3.2. Kinetics studies of PA-DNA sequence-specificity	38
1.6.3.3. Kinetics studies of $\beta$ -alanine and polyamide flexibility/rigidity	41
1.6.3.4. Kinetics studies of PA N- and C-terminal modifications	46
1.6.3.5. Studies of PA-DNA size-dependent binding kinetics	47
1.7. PA-DNA binding stoichiometry	48
1.7.1. Evidence of high linked PA-DNA binding stoichiometry in the literature	48
1.7.2. Multiple binding events and cooperativity	61
1.8. Overview of this dissertation	65

## CHAPTER 2            MATERIALS AND METHODS

2.1. Polyamide preparation and storage	66
2.2. DNA hairpins	66
2.2.1. DNA hairpin design	66
2.2.2. DNA preparation and storage	69
2.3. Buffer preparation	72
2.3.1. Recipes	72
2.3.2. Degassing	72
2.4. Fluorescence assays (equilibrium)	73
2.4.1. Direct titration to determine PA-DNA binding affinities ( $K_d$ )	73
2.4.2. Indirect determination of PA-DNA $K_d$ s via competition	77
2.4.3. PA-DNA binding stoichiometry	77
2.4.3.1. PA-DNA binding stoichiometry via fluorescence spectroscopy	77
2.4.3.2. PA-DNA binding stoichiometry via CD spectroscopy	78
2.5. Fluorescence assays (PA-DNA binding kinetics)	78
2.5.1. PA-DNA association kinetics via steady state instrument	78
2.5.2. PA-DNA association kinetics via stopped-flow fluorescence	83
2.5.3. PA-DNA dissociation kinetics	85

## CHAPTER 3 DNA BINDING KINETICS OF A LARGE ANTIVIRAL POLYAMIDE

3.1. Introduction and background	88
3.1.1. History and motivation	88
3.1.2. Design of DNA hairpins used for this study	90
3.2. Results	92
3.2.1. AT- versus TT-dependent NV1028-DNA binding affinities	92
3.2.2. DNA binding affinities determined indirectly via competition assay	94
3.2.3. Circular dichroism spectroscopy	94
3.2.4. Comparing binding stoichiometries obtained via fluorescence and CD spectroscopy	95
3.2.5. NV1028-DNA association kinetics	98
3.2.5.1. NV1028-DNA association rate constants	98
3.2.5.2. Multiple phases are observed when PA and DNA are 1:1	103
3.2.6. Macroscopic NV1028-DNA binding affinities ( $K_d$ )	103
3.2.7. NV1028-DNA dissociation kinetics	105
3.3. Discussion	110
3.3.1. Minor groove widths and NV1028-DNA binding kinetics	110
3.3.2. Reconciliation of kinetics and binding affinity data for AT and TT-dependence	110
3.3.3. Probing the slow phase for an additional binding event	114
3.3.3.1. Biphasic curves may fit well to monoexponential equations	114
3.3.3.2. Biexponential fits of slow phase curves	116
3.3.4. NV1028-DNA binding stoichiometry	118
3.3.4.1. High binding stoichiometries are consistent with previously reported data	118
3.3.4.2. NV1028-DNA binding modes	122

## CHAPTER 4 SIZE-DEPENDENCE OF HAIRPIN POLYAMIDES DNA AFFINITIES AND KINETICS

4.1. Introduction and background	129
4.1.1. Design of DNA hairpins used for this study	131
4.2. Results	132
4.2.1. PA size-dependence of DNA binding affinities ( $K_d$ )	132
4.2.2. PA-DNA binding affinities by competition	134
4.2.2.1. Competition DNA binding affinities for	134

	KA1039 and NV028	
	4.2.2.2. Difficulties with NV1042-DNA competition experiment	134
4.2.3.	Circular dichroism spectroscopy	135
	4.2.3.1. PA-DNA Size-dependent binding characterized via circular dichroism spectroscopy	135
	4.2.3.2. PA-DNA size dependent binding stoichiometries	138
	4.2.3.3. NV1042 binding stoichiometry determination via fluorescence spectroscopy	138
4.2.4.	Size-dependent PA-DNA association kinetics	140
	4.2.4.1. KA1039 and ODN-6-T3 CHAPS-dependent dye behavior	140
	4.2.4.2. Attempts at stopped-flow characterization of fast phase of KA1039	142
	4.2.4.3. Size-dependent PA-DNA association kinetic rate constants	143
	4.2.4.4. NV1042-DNA association rate constants (biexponential fitting)	148
4.2.5.	PA-DNA dissociation rate constants as a function of PA size	150
4.3.	Discussion	153
	4.3.1. Comparison of KA1039-DNA binding kinetics data collected by fluorescence spectroscopy and surface plasmon resonance (SPR)	153
	4.3.2. Internal consistency of $K_d$ and kinetic rate constants with respect to size	156
	4.3.3. PA-DNA association rate constants decrease with increasing PA size	158
	4.3.4. Macroscopic DNA $K_{off}$ correlates with PA size	161
	4.3.5. Implications for the large hairpin PA antiviral mechanism	161
CHAPTER 5	EFFECTS OF N-TERMINAL GROUPS ON THE DNA BINDING KINETICS OF LARGE ANTIVIRAL POLYAMIDES	
5.1.	Introduction and background	165
	5.1.1. Motivations	165
	5.1.2. DNA hairpins used to study DNA binding kinetics with respect to N-terminal functional groups	168
5.2.	Results	169
	5.2.1. DNA binding affinities for NV1028 and NV1042 series	169

5.2.2. DNA binding affinities determined via competition for NV1028 and NV1042 series	169
5.2.3. PA-DNA binding stoichiometry for NV1028 and NV1042 series	171
5.2.4 PA-DNA association kinetics for NV1028 and NV1042 series	173
5.2.4.1. NV1028 series-DNA association kinetics	173
5.2.4.2. NV1042 series-DNA association kinetics	177
5.2.5. PA-DNA dissociation rate constants for N-terminal series	181
5.3. Discussion	184
5.3.1. DNA binding affinities within an N-terminal series	182
5.3.2. Reconciliation of $K_d$ and kinetic rate constants with respect to N-terminal series	186
5.3.3. PA-DNA dissociation rate constants with respect to N-terminal function groups	187
5.3.4. PA-DNA association rate constants with respect to N-terminal functional groups	189
5.3.5. N-terminal functional groups and their implications on the viral mechanism	191
5.4. Concluding Remarks	195
REFERENCES	196

## LIST OF FIGURES

		Page
Figure 1.1	Structure of the DNA bases	2
Figure 1.2	B-DNA major and minor grooves	3
Figure 1.3	Types of DNA binding interactions	5
Figure 1.4	Structure of Distamycin A and Netropsin	9
Figure 1.5	Polyamide building blocks and structure	12
Figure 1.6	Cartoons of the different types of linked and unlinked PA structures	13
Figure 1.7	Perspective view of Py/Im polyamide bound to DNA in the minor groove	16
Figure 1.8	Novel five-membered heterocycles discussed in (Dervan & Edelson, 2003)	17
Figure 1.9	Curvature of polyamides with respect to DNA	18
Figure 1.10	HPV16 episome coding regions	22
Figure 1.11	Structures of NV1028 and NV1042 and the N-terminal groups of their guanidinylated analogs	24
Figure 1.12	Antiviral effects of NV1028 and Distamycin A determined in (Edwards et al., 2011)	25
Figure 1.13	NV1042 sequence map from (Vasilieva et al., 2016)	28
Figure 1.14	Kinetic traces of Dst-DNA binding where Dst is in excess	34
Figure 1.15	PA-DNA binding stoichiometry: CD spectra of 6-ring H-pin from (O'Hare et al., 2007)	52
Figure 1.16	PA-DNA binding stoichiometry: CD spectra of C7 cross-linked H-Pin PA from (Burckhardt et al., 2000)	53
Figure 1.17	PA-DNA binding stoichiometry: 6-ring hairpin in single mismatch site from (Pilch et al., 1996)	54
Figure 1.18	PA-DNA binding stoichiometry: 6-ring hairpin ZT65B with 3 sequences (Buchmueller et al., 2005)	56
Figure 1.19	PA-DNA binding stoichiometry: 8-ring hairpin PAs from (Wang, S. et al., 2012)	57

Figure 1.20	PA-DNA binding stoichiometry: 8-ring hairpin PAs from (Wang, S. et al., 2014)	58
Figure 1.21	PA-DNA binding stoichiometry: 8-ring hairpin PAs from (Liu, B. et al., 2017)	59
Figure 1.22	CD titration experiments of NV1028 and NV1028 with chiral $\gamma$ -(R)NH <sub>2</sub> by (Castaneda, 2017)	60
Figure 1.23	Independent versus dependent binding models	62
Figure 2.1	DNA binding orientations of hairpin polyamides	68
Figure 2.2	DNA hairpin overall design	70
Figure 2.3	Fluorescence spectra of free DNA versus DNA bound	75
Figure 2.4	Example of fitted equilibrium data	76
Figure 2.5	Example of binding stoichiometry data collected via fluorescence spectroscopy	79
Figure 2.6	Example of stoichiometry data collected via CD spectroscopy	81
Figure 2.7	Example of fitted association kinetics experimental data	82
Figure 2.8	Comparison of Jasco cutoff filters versus TAMRA bandpass filter	84
Figure 2.9	Example of dissociation kinetics experimental data	87
Figure 3.1	CD spectra collected by Kristin Balesfor NV1028 with different DNA oligonucleotides	96
Figure 3.2	Overlay of stoichiometry data collected by CD and fluorescence spectroscopy	97
Figure 3.3	Representative examples of association kinetics curves for NV1028 with different DNA oligonucleotides	99
Figure 3.4	Secondary association kinetics plots for NV1028 with different DNA oligonucleotides	101
Figure 3.5	Association kinetics curve of NV1028 with ODN-14-LFTT T6 at 1:1	104
Figure 3.6	K <sub>d</sub> collected while monitoring real-time changes in fluorescence of ODN-14-SFAT T3	106
Figure 3.7	Representative examples of K <sub>off</sub> data collected via fluorescence spectroscopy	108
Figure 3.8	DNA minor groove width prediction for ODN-	111

	14-SFTT and ODN-14-SFAT	
Figure 3.9	Simulated association curves generated using a biexponential equation	115
Figure 3.10	Secondary plots of NV1028 slow phase association curves when fit to a biexponential equation	117
Figure 3.11	CD spectrum and titration plot for NV1028 with GH6084C	121
Figure 3.12	Binding modes for multiple equivalents of NV1028	124
Figure 3.13	Sequence maps of NV1028 data collected via affinity cleavage and DNase I footprinting by Elena Vasilieva	126
Figure 4.1	NV1042 competition experiment with ODN-20-T3 collected by Yang Song	136
Figure 4.2	CD spectra collected by Kristin Bales for PA of KA1039, NV1028, and NV1042	137
Figure 4.3	Overlay of binding stoichiometry data collected by CD and fluorescence spectroscopy	139
Figure 4.4	Comparison of KA1039 and ODN-6-T3 CHAPS-dependent behavior	141
Figure 4.5	Example of stopped-flow data for KA1039 with ODN-6-T3	144
Figure 4.6	Representative examples of association kinetics curves collected for KA1039, NV1028, and NV1042	145
Figure 4.7	Secondary association kinetics plots for KA1039, NV1028, and NV1042	146
Figure 4.8	Representative secondary plots of NV1028 and NV1042 association curves when fit to a biexponential equation.	149
Figure 4.9	Representative dissociation rate constant plots for KA1039, NV1028, NV1042, and KA1002 replicated from (Dupureur et al., 2012)	151
Figure 4.10	SPR sensorgram of KA1039 replicated from (Wang, S. et al., 2012)	154
Figure 4.11	Comparison of all-at-once binding versus zipper-like binding	159
Figure 4.12	PA size-dependence of DNA dissociation rate constants	162
Figure 4.13	Possible <i>in vivo</i> binding model for at least 2:1 NV1028-DNA binding	164
Figure 5.1	N-terminal functional groups studied in Chapter 5	166

Figure 5.2	Stoichiometry data for NV1028 and NV1042 series	172
Figure 5.3	Association kinetics traces of NV1028 and NV1042 series	174
Figure 5.4	NV1028 series secondary association plots	175
Figure 5.5	NV1042 series secondary association plots	178
Figure 5.6	Representative secondary plots of NV1042 series when association curves were fit to a biexponential equation	180
Figure 5.7	Dissociation curves for NV1028 and NV1042 series	182
Figure 5.8	Residence time and potency with respect to N-terminal functional groups	185
Figure 5.9	Proposed structures for repulsion-induced altered conformations of polyamides with charged N-terminal groups	192
Figure 5.10	Association rate constants and potency with respect to N-terminal functional groups	193



## LIST OF TABLES

		Page
Table 1.1	DNA binding ligands	6
Table 1.2A	Summary of data from previous hairpin PA-DNA binding kinetics studies (linked vs unlinked PA)	37
Table 1.2B	Summary of data from previous hairpin PA-DNA binding kinetics studies (DNA sequence dependence)	39
Table 1.2C	Summary of data from previous hairpin PA-DNA binding kinetics studies ( $\beta$ -alanine, C-terminal groups, size-dependence)	42
Table 1.3	Index of published PA-DNA binding stoichiometry	50
Table 2.1	Polyamides characterized in this dissertation	67
Table 2.2	DNA hairpins discussed in this dissertation	71
Table 2.3	Settings used for CD spectroscopy experiments	80
Table 3.1	NV1028 and the DNA hairpins relevant to Chapter 3	91
Table 3.2	DNA binding affinities for NV1028	93
Table 3.3	Association rate constants for NV1028	102
Table 3.4	DNA dissociation rate constants ( $K_{off}$ ) for NV1028	109
Table 3.5	Comparison of experimental and computed NV1028-DNA binding constants	113
Table 3.6	Concentration dependent rate constants for NV1028 slow phase when fit to a mono or biexponential equation	119
Table 4.1	Polyamides and DNA hairpins relevant to Chapter 4	130
Table 4.2	Comparison of binding affinities determined via fluorescence assay versus other methods	133
Table 4.3	DNA association rate constants ( $k_{on}$ ) for PAs by size	147
Table 4.4	Comparison of dissociation rate constants determined via fluorescence assay versus other methods for a 1:1 PA:DNA complex	152
Table 4.5	Comparison of experimental and computed PA-DNA binding constants for KA1039, NV1028, and NV1042	157

Table 5.1	Polyamides and DNA hairpins relevant to Chapter 5	167
Table 5.2	Binding affinities for NV1028 and NV1042 series	170
Table 5.3	Association rate constants for NV1028 and NV1042 series	176
Table 5.4	Dissociation rate constants and percent fluorescence recovery for NV1028 and NV1042 series	183
Table 5.5	Comparison of experimental and computed PA-DNA binding constants for NV1028 and NV1042 series	188

## LIST OF SCHEMES

	Page
Scheme 1.1    Ligand binding reversibly to DNA	30

## LIST OF ABBREVIATIONS

PA, polyamide; ODN, oligonucleotide; HPV, human papillomavirus; LCR, long control region;  $\beta$ ,  $\beta$ -alanine; Im, N-methylimidazole; Py, N-methylpyrrole; Dp, N,N-dimethylaminopropylamine; Ta, 3,3-diamino-N-methyldipropylamine; TAMRA, 6-carboxytetramethylrhodamine; CD, circular dichroism spectroscopy, UV-VIS, ultraviolet-visible spectrophotometer; SPR, surface plasmon resonance; AC, affinity cleavage

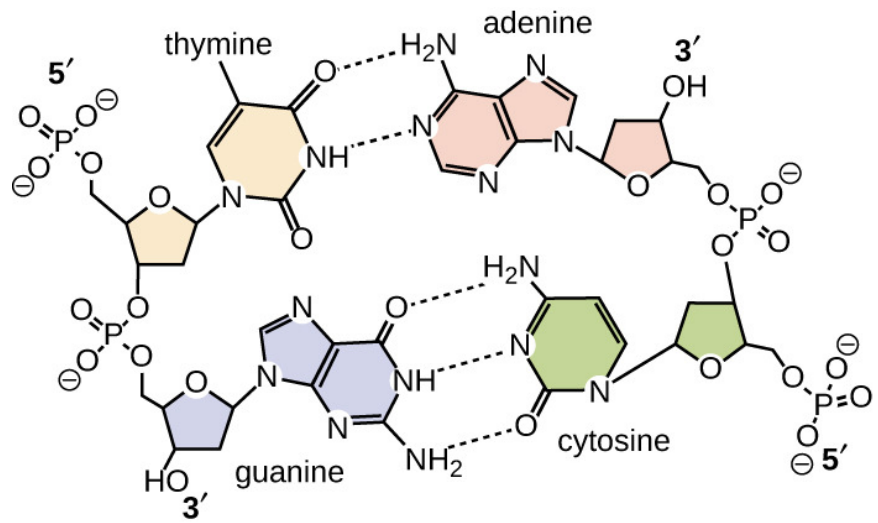
# CHAPTER I

## INTRODUCTION

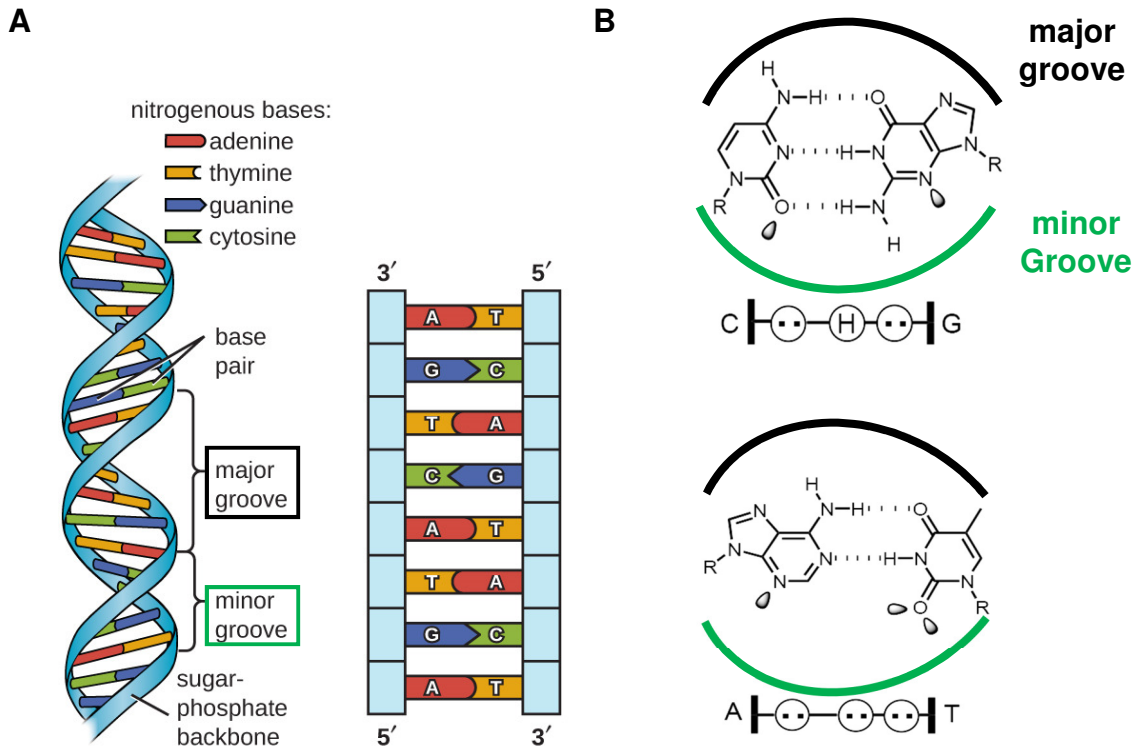
### 1.1. Structure of DNA

DNA is a biopolymer that stores the genetic information in living organisms, as well as some viruses, and contains instructions used for function, development, growth, and replication. Double-stranded DNA is made up of two complementary strands containing four nucleotides which are composed of a sugar, a phosphate, and nitrogen-containing base: adenine (A), which pairs with thymine (T); and guanine (G), which pairs with cytosine (C) (**Fig. 1.1 and 1.2**) in what is referred to as Watson-Crick base-pairing. The strands together form the double-helix shape first elucidated in 1953 (Watson & Crick, 1953) (**Fig. 1.2A**). The sugars of the bases are linked together to form the polymer strand via a phosphodiester bond giving DNA a negatively charged backbone. Each strand is then bound and stabilized to the other via ring stacking and hydrogen bonding interactions between the complementary DNA bases.

The double-helix structure of DNA has both a major and minor groove which vary in size depending on the conformation of DNA: A-, B-, or Z-DNA (Voet et al., 2013). The major and the minor grooves are characterized by their hydrogen bond topology (**Fig. 1.2B**): different hydrogen bonding acceptors and donors are present in the major versus the minor groove. B-DNA is the most common and biologically abundant conformation of DNA and is characterized by



**Figure 1.1** Structure of the DNA bases. Reproduced with permission from ("Structure and Function of DNA," 2018) under Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License.



**Figure 1.2** B-DNA major and minor grooves. A, DNA structure reproduced from ("Structure and Function of DNA," 2018) under Creative Commons Attribution-Noncommercial-Share Alike 3.0 United States License. B, corresponding hydrogen bond topology adapted with permission from (Dervan & Edelson, 2003) Copyright 2003 Elsevier. For each pair, the face of the major groove is shown in **black**, and the the face of the minor groove is shown in **green**.

a right-handed helix, and with a large major groove and smaller minor groove (**Fig. 1.2A**).

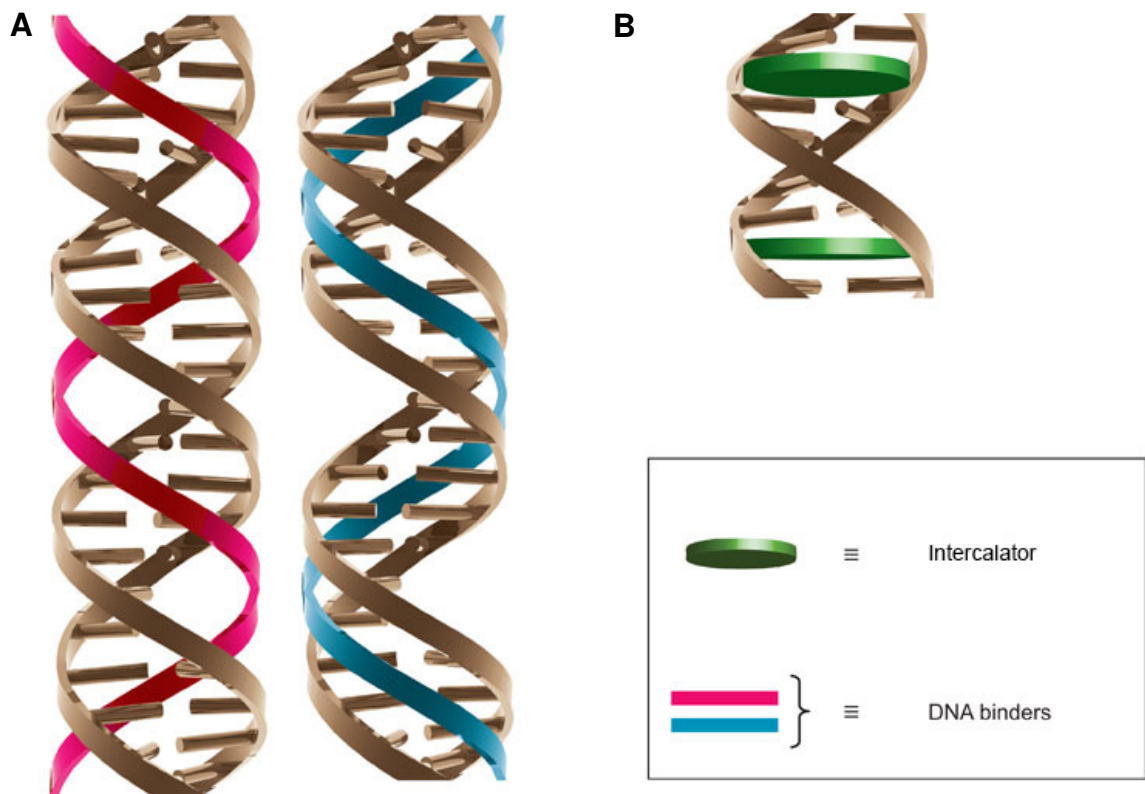
## 1.2. DNA binding ligands

There are a number of different natural and synthetic DNA binding ligands described throughout the literature. Given the unique shape, structure, and properties of DNA, there are a variety of different modes by which a ligand is able to bind DNA. Some common types of DNA binding are covalent, electrostatic with the phosphate backbone, intercalation, and minor and major groove binding, and any combination thereof (Hurley, 2002). A covalent interaction is typically considered irreversible, whereas a non-covalent interaction can likely be reversed.

The negatively charged sugar-phosphate backbone of DNA allows for electrostatic interactions to occur between ligand and DNA. An example of an electrostatic interaction occurs between the enzyme *EcoRI*, a restriction enzyme, and the recognition site GAATTC (Lesser et al., 1990). The protein-phosphate contacts help to anchor and position recognition elements of the protein so it can make specific binding interactions in the major groove.

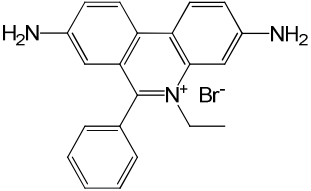
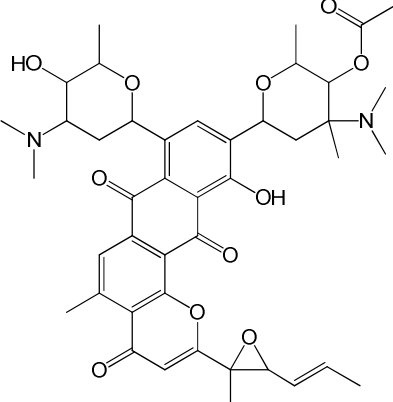
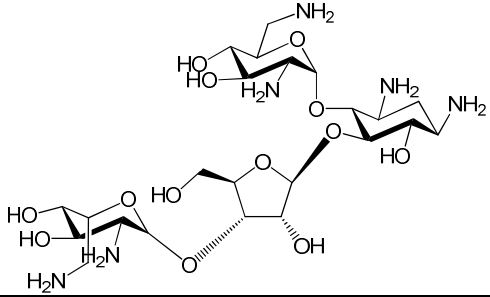
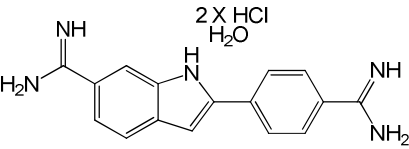
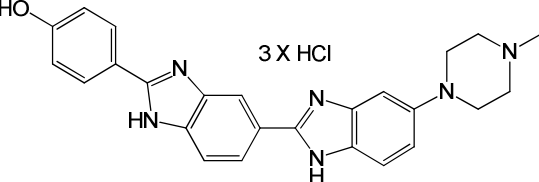
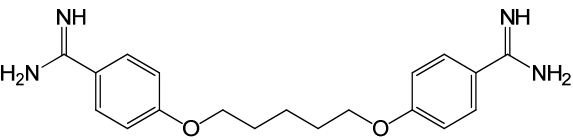
Intercalators associate with DNA via ring-stacking and  $\pi$ -orbital interactions (**Fig. 1.3b**). The insertion of these planar molecules into the DNA typically results in deformation of DNA and can result in mutagenesis. A well-known example of an intercalator is ethidium bromide, a fluorescent dye used to stain DNA. Structures of representative DNA ligands can be found in **Table 1.1**.





**Figure 1.3** Types of DNA binding interactions. A, DNA groove binding: pink, major groove; blue, minor groove. B, intercalation. Adapted with permission from (Urbach, 2011) Copyright 2011 Springer Nature, license number 4316320252982.

**Table 1.1** Examples of DNA binding ligands

DNA Ligand	Structure	Type
Ethidium bromide		Intercalation <sup>a</sup>
Pluramycin A		Major groove <sup>b</sup>
Neomycin		Major groove <sup>c</sup>
DAPI		Intercalation Minor groove <sup>cd</sup>
Hoechst 33258		Minor groove <sup>ac</sup>
Pentamidine		Minor groove <sup>d</sup>

<sup>a</sup>(Nelson et al., 2007) <sup>b</sup>(Hamilton & Arya, 2012) <sup>c</sup>(Reddy, B. S. P. et al., 2000)  
<sup>d</sup>(Barrett et al., 2013)

### **1.2.1. Major groove binders**

Major groove binding interactions are typically favored by proteins, such as the aforementioned *EcoRI* (Kim et al., 1990) or proteins that utilize the leucine zipper motif (Ellenberger et al., 1992), while minor groove binders tend to be smaller peptides or small molecules, such as drug compounds. Several of the major groove binders, such as pluramycin A and its analogs, are capable of binding specifically, but typically rely on an intercalating or alkylating moiety (Hamilton & Arya, 2012). Recently, there has been interest in aminoglycosides, such as neomycin, which have a carbohydrate scaffold capable of binding reversibly to DNA. Although there have been a number of major groove binders investigated, research has focused mainly on developing minor groove chemotherapeutics.

### **1.2.2. Minor groove binders**

#### **1.2.2.1. Non-therapeutic minor groove binders**

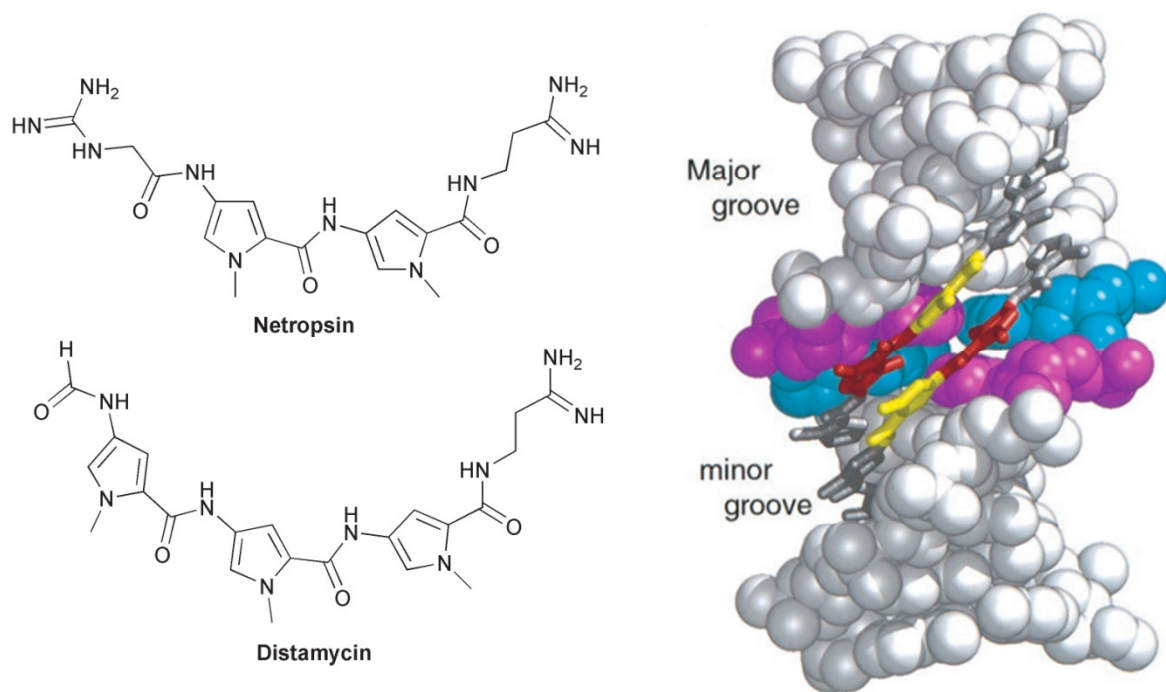
Minor groove binders have many therapeutic and non-therapeutic uses. Well-known examples of minor groove binders with non-therapeutic uses are the Hoechst stains and 4',6-diaminidino-2-phenylidole (DAPI). Both Hoechst stains and DAPI bind in the minor groove of AT-rich regions of DNA and are used as fluorescent DNA stains in both living and fixed cells. Their use as cell stains come from their ability to traverse cell membranes. However, since DAPI is less efficient at crossing cell membranes in living cells, it is typically used with fixed cells. Hoechst stains are fluorescent dyes that are characterized as bis-

benzimidazoles that can easily enter living cell membranes and are popular for live-cell staining (Reddy, B. S. P. et al., 2000). Given the desirable properties of Hoechst dyes, there has also been interest in developing anti-cancer Hoescht analogs (Nelson et al., 2007). DAPI has shown antitrypanosomal activity, but further development has not been the focus of recent studies; DAPI remains a powerful tool in cellular microscopy (Barrett et al., 2013). Minor groove binders have a number of therapeutic applications. For example, pentamidine and its furan analogs have been shown to be powerful antiparasitics with  $IC_{50}$ s in the low nM range (Barrett et al., 2013).

#### **1.2.2.2. Distamycin A and netropsin**

Distamycin A (Dst) and netropsin (Net) are natural products isolated from the *Streptomyces* species (**Fig. 1.4**). These small linear polyamides (PA) consist of 3 or 2 adjacent pyrrole units, respectively, that allow them to bind in the minor groove of AT-regions of DNA: the amide protons form hydrogen bonds with the O2 of pyrimidines and N3 of purines (Baliga & Crothers, 2000b), while the N2 amino group of guanine causes steric hindrance making binding unfavorable in GC-rich regions of the DNA (Marverti et al., 2012). When the double bonds of Dst are saturated, its binding affinity to DNA is weakened, showing the importance of the heterocyclic pyrrole residues for DNA binding (Woods et al., 2002).

Additionally, van der Waals contacts and electrostatic interactions of the terminal amidine group with the phosphate backbone also contribute to the preferential binding of Dst in the minor groove.



**Figure 1.4** Structures of Distamycin A (Dst) and netropsin (Net) (left). Figure adapted with permission from (Khalaf et al., 2004) Copyright 2004 American Chemical Society. A small linear polyamide bound in DNA minor groove in 2:1 mode (right), adapted with permission from (Kielkopf et al., 1998b) Copyright 1998, The American Association for the Advancement of Science, license number 4316320563478.

Dst is capable of binding in either a 1:1 or a 2:1 binding mode, where the linear molecules bind in an antiparallel conformation (Chen & Sha, 1998; Fagan & Wemmer, 1992). These different modes have been shown to be sequence dependent where the 2:1 complex generally tends to prefer mixed AT sites (Chen & Sha, 1998). This preference has been attributed to the difference in minor groove widths between regions of 4 or more consecutive A (A-tracts) and mixed AT regions ((AT)<sub>n</sub>) sequences. Narrowing of the minor groove, as well as an intrinsic curvature of the DNA towards the minor groove, is observed for A-tracts (Burkhoff & Tullius, 1988), while (AT)<sub>n</sub> sequences tend to be less narrow due to the flexibility of the TpA steps (Crothers & Shakked, 1999; Wu & Crothers, 1984). To accommodate the second equivalent of Dst binding, a widening of the minor groove of at least 3.5 Å is required (Fagan & Wemmer, 1992). Only 1:1 binding has been observed for Net which is likely due to the electrostatic repulsion of its terminal groups: in place of the uncharged amide group of Dst, the slightly smaller Net has a guanidinium group which leaves both of its termini Net charged (**Fig. 1.4**).

The interest in Dst and Net comes from their antimicrobial and antitumor properties (Neidle, 2001). However, despite Dst and Net's sequence specificity and desired chemotherapeutic properties, these molecules are cytotoxic, limiting their clinical usage (Edwards et al., 2013b; Finlay et al., 1951). Additionally, studies have shown Dst to be a potent antiviral (Baraldi et al., 2004). The ability of Dst and Net to bind specifically in the minor groove of AT-rich regions of DNA has made them important lead compounds in anti-cancer and antiviral research.

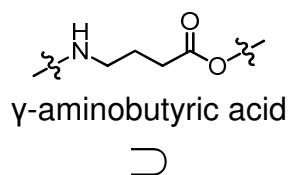
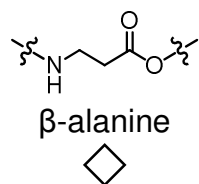
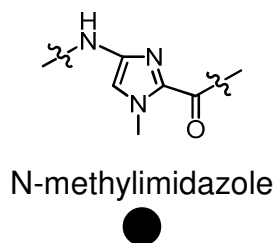
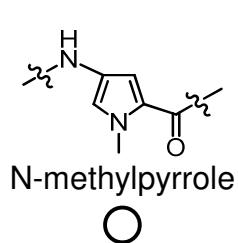
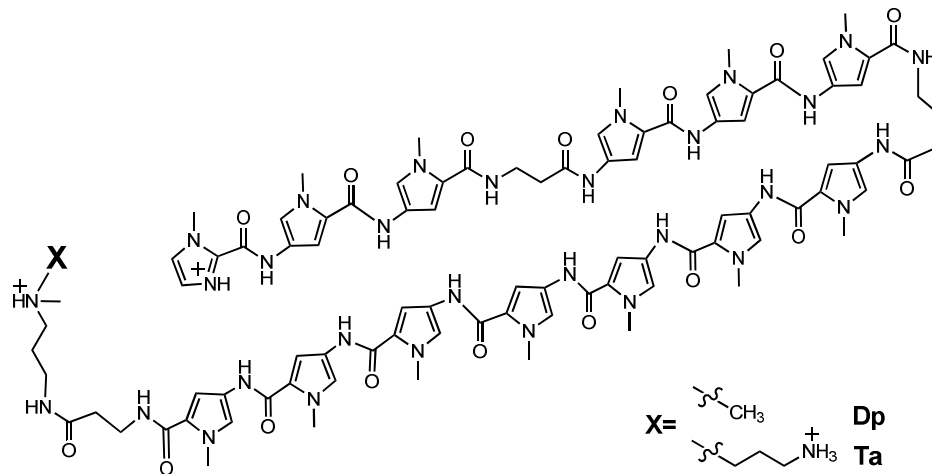
These natural products have inspired the development of other minor groove binding compounds with lower toxicity, such as N-methylimidazole/N-methylpyrrole polyamides, referred to here simply as polyamides (PA).

### **1.3. N-methylpyrrole and N-methylimidazole polyamides**

There are currently four basic building blocks of PAs which allow them to bind selectively to DNA: N-methylpyrrole (Py), N-methylimidazole (Im),  $\beta$ -alanine ( $\beta$ ), and  $\gamma$ -aminobutyric acid ( $\gamma$ ) (**Fig 1.5**). While there are variations and additional building blocks used by different investigators, these four are the most common.

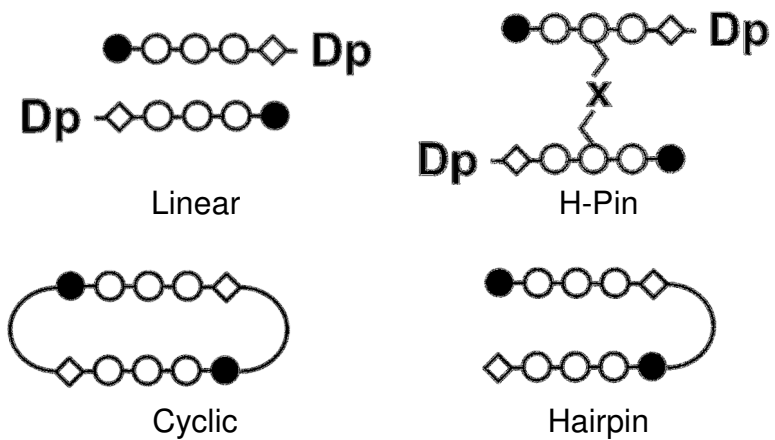
#### **1.3.1. Hairpin polyamide structure and $\gamma$ -aminobutyric acid**

The initial PAs, also referred to as lexitropsins in earlier publications, mimicked the linear structure of Dst and Net (Lown, 1988, 1992). However, the observed 2:1 DNA binding of Dst has inspired the design of linked PAs where linear monomers are connected in an antiparallel configuration: the hairpin where monomers are linked N-termini to C-termini with  $\gamma$ -aminobutyric acid ( $\gamma$ ) at one side of the molecules, circular where both the monomers are linked by  $\gamma$  at both N-termini and C-termini, and H-pins where the monomers are crosslinked at an internal position by an alkyl spacer (**Fig. 1.6**). It has been shown that linked molecules bind with greater affinity and specificity than their unlinked monomers (Cho et al., 1995; Mrksich & Dervan, 1994). The class of PA molecules discussed in this dissertation are hairpin PAs and, thus, the remaining discussion

**A****B**

**Figure 1.5** Polyamide building blocks and structure. A, polyamide building blocks; B, structure of a 14-ring hairpin PA in DNA binding site with C-terminal groups shown as X. Since  $\beta$  serves as a Py substitute, they are counted as “rings”. Protonated Im’s must deprotonate in order to properly H-bond with G.





**Figure 1.6** Cartoons of the different types of linked and unlinked PA structures. The Im/Py/ $\beta$  sequence is arbitrary; the loop represents  $\gamma$ ; the "X" of the H-pin represents an alkyl spacer of varying length and structure.

will focus on hairpin PAs.

### **1.3.2. $\beta$ -alanine as a flexible linker**

Selectivity was thought to increase with size: a bigger, more unique binding site should result in fewer non-specific interactions. By increasing the number of Py/Im rings, torsional strain is increased which causes the hairpin to curve and misalign with the DNA bp (Han et al., 2012; Turner et al., 1998; Wang, C. C. et al., 2001). Although PA residues do allow for sequence specificity, the shape of the groove can be equally important. It is for this reason  $\beta$ -alanine ( $\beta$ ) is used in larger PAs (PAs that have 5 or more consecutive Py or Im) to provide flexibility by replacing a Py. The inclusion of  $\beta$  into hairpin PAs of 8 and 12 rings was shown to improve specificity of the hairpin PAs as determined by quantitative DNase I footprinting analysis (Turner et al., 1998). More recently, SPR has been used to study the PA-DNA binding affinities and binding kinetics, and the benefits of including  $\beta$  seem to be dependent on a number of factors, such as how many  $\beta$  are included and where they are included with respect to overall PA sequence and structure (Han et al., 2013; Wang, S. et al., 2014). Over the past 2 decades  $\beta$  has continued to be included as a flexible linker but the exact effects on binding affinities and binding kinetics are still being studied. See **section 1.6.3.3** for discussion of  $\beta$ -related PA-DNA binding kinetics studies.

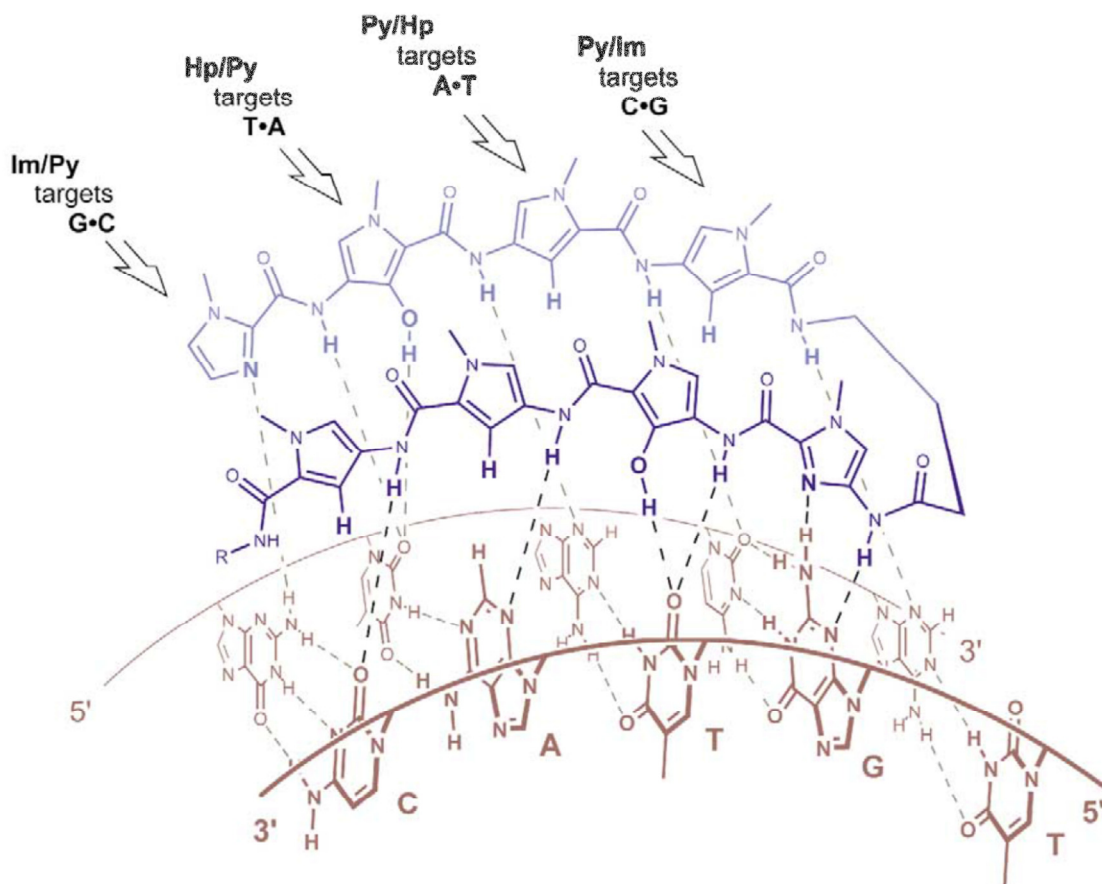
### **1.3.3. N-methylpyrrole (Py) and N-methylimidazole (Im)**

Py binds selectively to W (where W stands for A or T) and C. Im binds

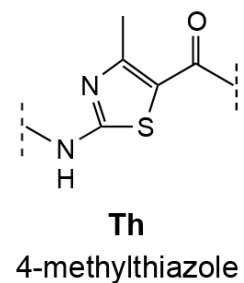
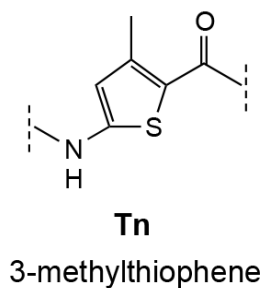
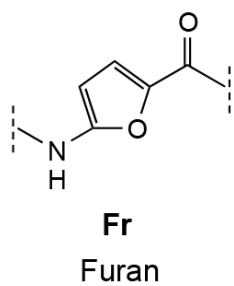
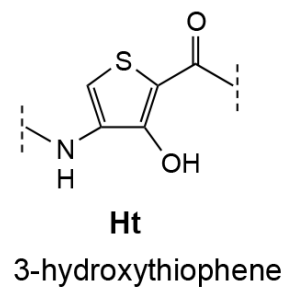
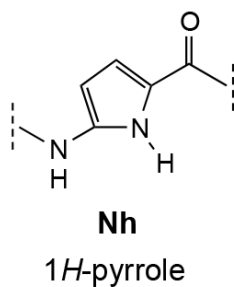
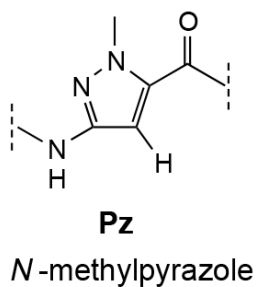
selectively to G, where the Im nitrogen forms a hydrogen bond with the exocyclic amine of G (Kielkopf et al., 1998a) (**Fig. 1.7**). These two units are amino acid building blocks which are condensed by solid-phase Merrifield chemistry to form amides, hence their name Py/Im polyamides. Py and Im serve as the main components of PA sequence recognition.

Sugiyama and Dervan have developed thymine-selective building blocks which can discriminate TA from AT bp via modification of existing residues (Farkas et al., 2009; Zhang et al., 2012). A series of five-membered heterocycles were investigated by Dervan within an 8-ring hairpin structure (**Fig 1.8**). Heterocycles, such as N-methylpyrazole (Pz) and 3-methylthiophene (Tn), had enhanced AT selectivity over GC but only modest T selectivity over A. Other heterocycles, such as furan (Fn) and 3-hydroxythiophene (Ht) were also found to have a dramatic effect on the overall curvature of the 8-ring hairpin PA resulting in a complete loss of binding (**Fig 1.9**).

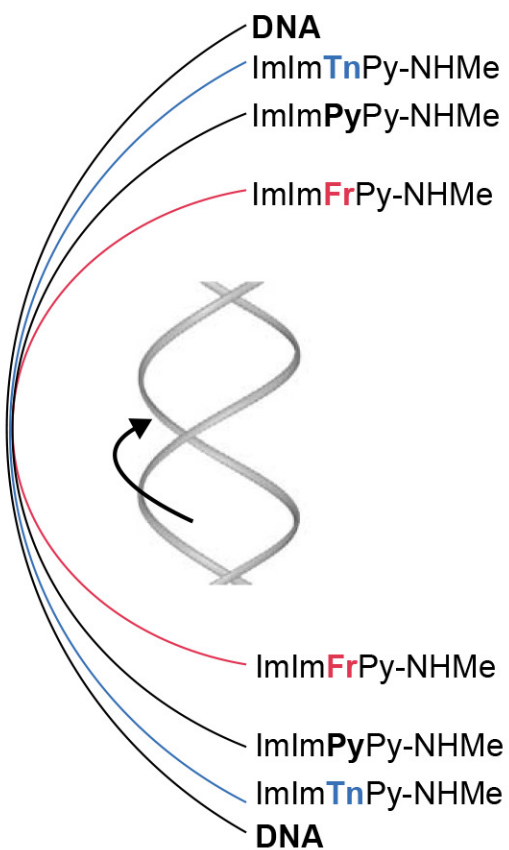
N-methyl-3-hydroxypyrrole, Hp, was one of the novel heterocycles that showed considerable thymine-selectivity. Unfortunately, Hp was found to degrade in acidic environments or in the presence of free radicals making it unsuitable for biological uses (Marques et al., 2002). Work with benzimidazoles, such as hydroxybenzimidazole (Hz) and imidazopyridine (Ip), has also shown potential as more-selective residues. Not only do these heterocycles show selectivity for T and G, respectively, they are also chemically stable, and do not affect the PA curvature unfavorably (Dervan & Edelson, 2003).



**Figure 1.7** Perspective view of Py/Im polyamide bound to DNA in minor groove. Solid line represents phosphate backbone; dashed lines represent hydrogen bonds. Reproduced with permission from (Dervan & Edelson, 2003) Copyright 2003 Elsevier.



**Figure 1.8** Novel five membered heterocycles. Adapted with permission from (Dervan & Edelson, 2003) Copyright 2003 Elsevier.



**Figure 1.9** Curvature of PAs with respect to DNA demonstrating how a small difference in a single heterocycle can result in a large change in curvature. Fr, furan; Tn, 3-methylthiophene; Py, N-methylpyrrole; Im, N-methylimidazole. Structures of Fr and Tn can be found in **Fig. 1.8**. Reproduced with permission from (Dervan & Edelson, 2003) Copyright 2003 Elsevier.

### **1.3.4. C-terminal functional groups (Tails)**

Dst has an amidine, and Net has a guanidinium and amidine flanking their 3 or 2 Py residues, respectively. Early smaller PAs featured a Dp (3-(dimethylamino)propylamine) tail that affords additional AT/TA binding specificity (Swalley et al., 1999) (**Fig. 1.5B**). The Ta (3,3'-diamino-N-methyldipropylamine) tail, which has two positive charges versus Dp's single positive charge has also been used as a tether for functional groups, such as EDTA or dyes (White, S. et al., 1997a). Issues with aggregation of larger PAs led to the inclusion of Ta tail in the large antiviral PAs studied in this thesis (Wang, S. et al., 2014). The additional charge of the Ta is assumed to have a twofold effect: first, the positive charges can interact electrostatically with the negative backbone of DNA; second, the additional positive charge may aid in repulsion of other PAs resulting in decreased aggregation.

## **1.4. Applications of polyamides**

### **1.4.1. Polyamides and gene regulation**

PAs bind to DNA with affinities comparable to DNA binding proteins giving them the potential to inhibit gene expression through a variety of different mechanisms. For example, minor groove binders have been shown to be able to inhibit binding of transcription factors (TF) to DNA by preventing TF from making necessary contacts for recognition or by causing structural or conformational changes to the DNA that disrupt protein-DNA dynamics important for TF recognition and binding (He et al., 2015). A PA can be designed to bind

competitively to the recognition site of a transcription factor thereby displacing and/or preventing it from binding (Dervan & Edelson, 2003; Reddy, B. S. P. et al., 2000). An example of this is micronogotropen (MGT), a minor groove binder inspired by Hoechst 33342 (White, C. M. et al., 2001). MGT binds in the minor groove of AT rich regions of DNA through its benzimidazole moiety while its polyamine tail makes electrostatic contacts with the backbone of DNA. It was also shown that the inclusion of a tripyrrole/pyrrole functional group was crucial for inhibition of gene expression in whole-cell assays.

Conversely, PAs can also be used to upregulate certain genes by inhibiting the binding of repressor proteins to DNA (Dervan & Edelson, 2003). An example of this was the use of PAs to inhibit UL122 immediate early protein 2 (IE86) from binding to human cytomegalovirus (HCMV) DNA by preventing IE86 from binding (Dickinson et al., 1999). Being able to control the recruitment of transcriptional machinery is crucial to gene expression. PAs and other minor groove binders are not only selective, but are amenable to design.

#### **1.4.2. Polyamide antiviral applications**

In addition to the studies of PAs in the treatment of HCMV, PAs have been the focus of other antiviral studies. Dimeric Dst and Net analogs linked by a glycine *cis*-diamino platinum group have shown enhanced binding to the herpes simplex virus (HSV) origin of replication with low levels of cytotoxicity (Surovaya et al., 2016). A chlorambucil-hairpin PA conjugate (HP-CHL) (Wurtz & Dervan,

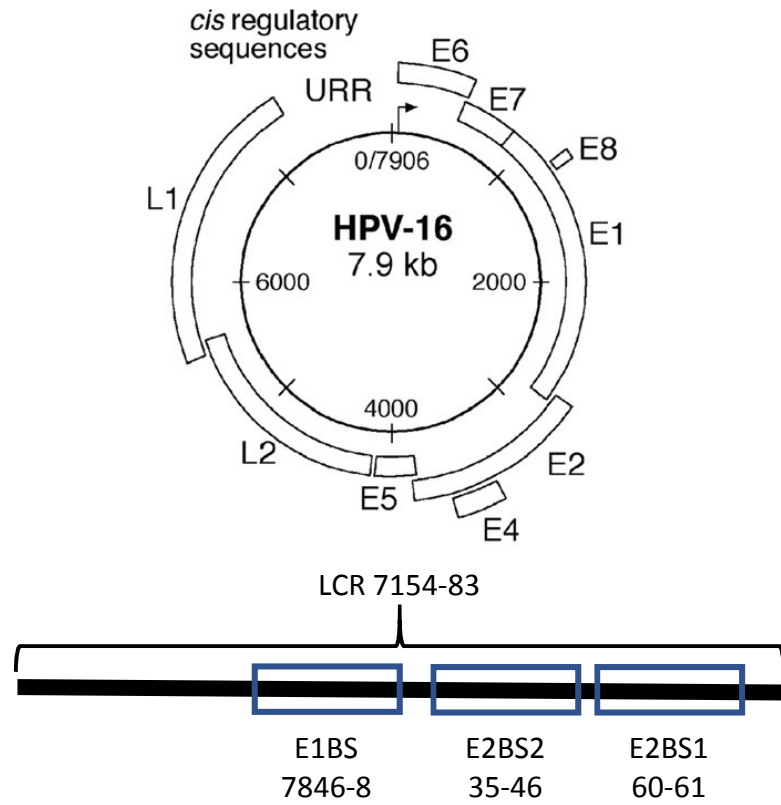


2000) was shown to be capable of alkylating specific sites of HIV-1 promoter effectively blocking HIV-1 (human immunodeficiency virus) replication.

### **1.5. Human papillomavirus (HPV)**

Human papillomavirus (HPV) is the most common sexually transmitted infection and the cause of virtually all cervical cancers (CDC, 1999; NCI, 2018). There are over 200 related viruses that make up the group, but the strains HPV16 and HPV18 are responsible for more than 70% of cervical cancers (Crow, 2012; Lacey et al., 2011; NCI, 2018). Although there are vaccines for the most common HPV strains, the vaccines serve only as a prophylactic; there is currently no cure or palliative antiviral treatment for HPV. Typically, HPV infection is cleared naturally by the host, but on rare occasions, HPV infections with oncogenic strains can result in cancer. The frequency of cervical cancer has decreased significantly due to routine cervical cytology screenings (i.e. Pap smears) (NCI, 2018). However, cervical cancer still remains a problem worldwide (Cutts et al., 2007). Additionally, it is still unclear why HPV infection is more persistent in some populations but not others, and studies of risk factors are still underway.

The HPV genome exists of an 8 kbp double-stranded episome and has 3 distinct regions: early (E), late (L), and long control region (LCR) (**Fig. 1.10**). These regions are further subdivided into a total of 8 open reading frames: 6 in the early region (E1, E2, E4, E5, E6, E7), 2 in the late region (L1, L2) (Williams et al., 2011). The LCR does not contain any protein coding regions but contains



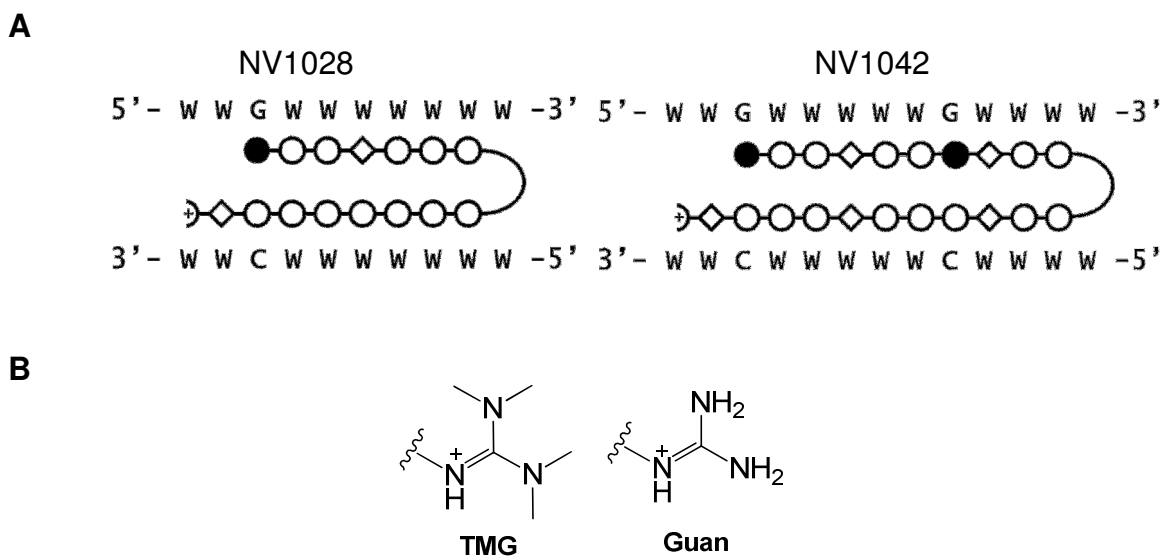
**Figure 1.10** HPV16 episome coding regions. Top, The HPV-16 episome. Bottom, E1 and E2 binding sites within the LCR. Adapted with permission from (Lace et al., 2011) Copyright 2011 American Society for Microbiology.

several transcription factor binding sites, such as the E1 and E2 binding sites, important for viral replication. Aside from E1 and E2 proteins, viral replication relies predominantly on host cell machinery. Thus, replication of HPV can be inhibited by preventing E1 and E2 from binding to the LCR.

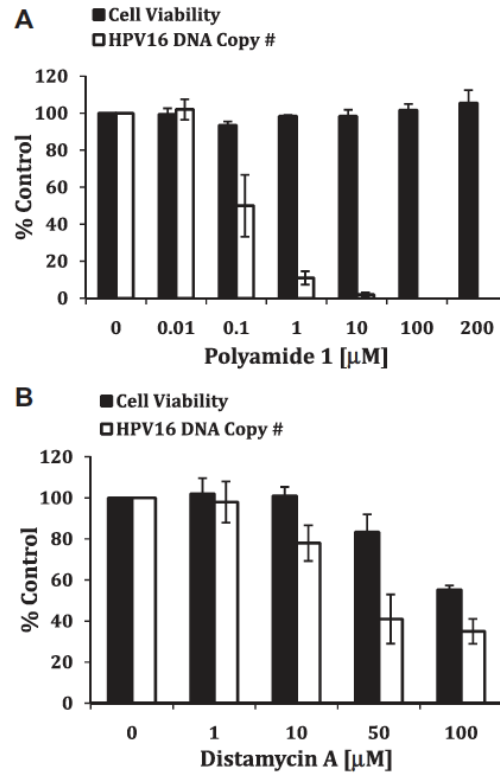
### **1.5.1. Previous cell biology work with anti-HPV polyamides**

#### **1.5.1.1. Characterization of anti-HPV potency**

The AT-rich nature of the LCR of the HPV genome (greater than 60%) makes it an excellent target for Dst-inspired PAs. Large hairpin PAs, NV1028 and NV1042, were designed to bind to the LCR in order to prevent replication of the HPV viral genome (**Fig. 1.11**). NV1028 was designed to bind to the sequence  $W_2GW_7$  and NV1042 was designed to bind to the sequence  $W_2GW_5GW_4$ , where W represents A or T. Previous characterization of the antiviral behavior of NV1028 and NV1042 was performed using a range of cells for HPV16, HPV18, and HPV31, including W12E cells, HPV16 infected human cervical epithelial cells. Cell viability was determined using an MTT assay described in (Edwards et al., 2011), and  $IC_{50}$ 's and  $IC_{90}$ 's (the concentration at which copies of viral DNA was reduced by 50% and 90%, respectively) were determined by quantifying the copies of HPV16 per cell using the previously described q-PCR technique (Garner-Hamrick & Fisher, 2002). NV1028 and NV1042 were shown to be not only more effective against HPV16 than Dst with  $IC_{50}$ 's of 100 nM and 36 nM, respectively, but also much less cytotoxic (Edwards et al., 2011) (**Fig. 1.12**, NV1042 data not shown).



**Figure 1.11** Structures of NV1028 and NV1042 and N-terminal groups. A, cartoon structures of NV1028 and NV1042 in their binding sites based on Dervan's recognition rules. B, structures of tetramethylguanidinium (TMG) and guanidinium (Guan).



**Figure 1.12** Effects of NV1028 (PA1), top, and Dst, bottom, on W12E HPV 16 infected cells. Reproduced with permission from (Edwards et al., 2011) Copyright 2011 Elsevier, license number 4316321296988.

Recent studies have focused on NV1028 analogs where the N-terminal des-amino of NV1028 was substituted with tetramethylguanadinium (TMG) or guanidinium (Guan) to produce NV1028-TMG and NV1028-Guan, respectively (**Fig. 1.11**) (Castaneda et al., 2017). It was shown that while the N-terminal substitutions had no effect on the  $IC_{50}$ 's, a 3-fold improvement of  $IC_{90}$  was observed for NV1028-Guan versus NV1028, and a 10-fold worsening in antiviral activity was observed for NV1028-TMG versus NV1028 (i.e. Guan>H>TMG).

#### **1.5.1.2. HPV and the DNA Damage Response (DDR)**

Using a pathway consisting of a network of proteins that detect, signal, and repair damaged DNA, collectively referred to as the DNA damage response (DDR), eukaryotic cells are able to maintain genomic integrity (Harper & Elledge, 2007). Two important sensors of DNA damage are the ataxia-telangiectasia-mutated (ATM) and ataxia telangiectasia and Rad3-related (ATR) serine/threonine protein kinases. ATM primarily responds to double-strand breaks, while ATR primarily responds to single-strand breaks. Activation of these sensors results in a signal transduction cascade that results in recruitment of DNA repair factors or cell-cycle checkpoints. Activation of the DDR can be unfavorable for viral replication. Thus, some viruses, such as HIV, Epstein-Barr virus, and HPV, have evolved to exploit DDR responses (Edwards et al., 2013a; Moody & Laimins, 2009; Sakakibara et al., 2011).

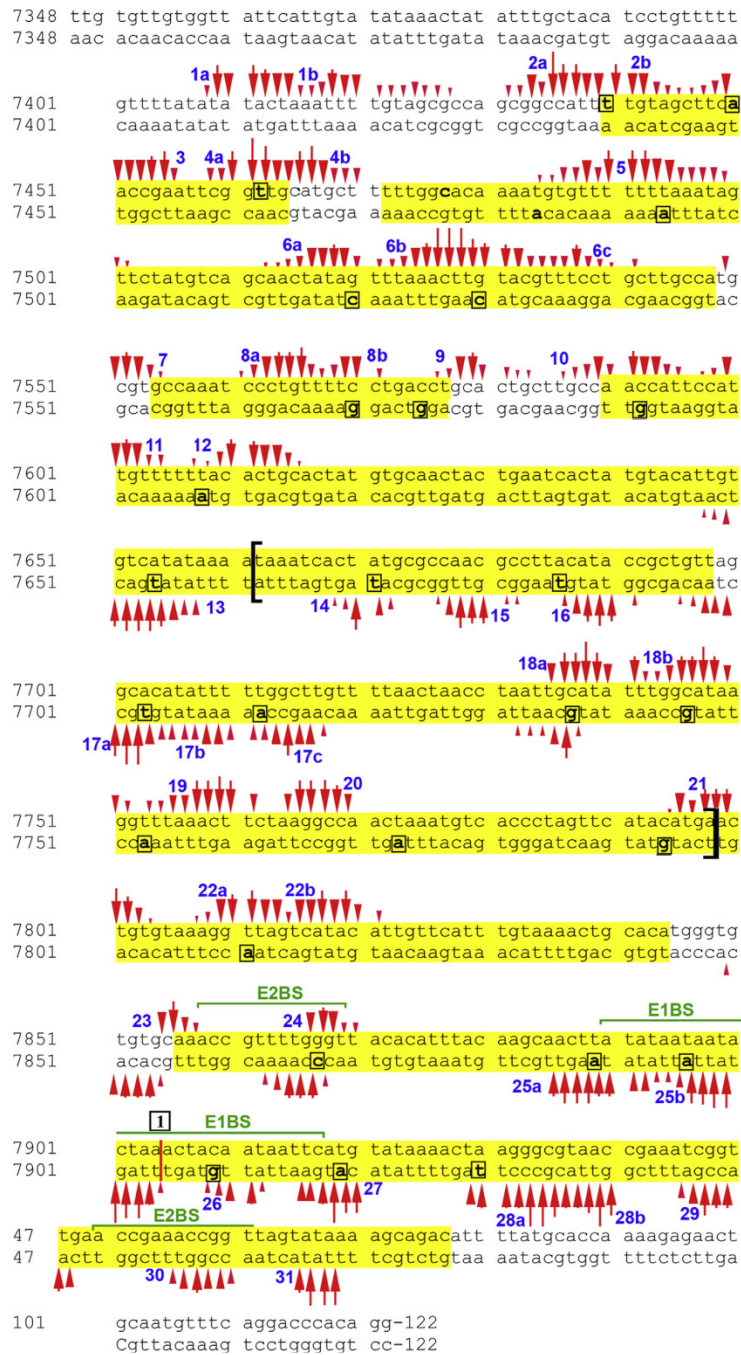
Edwards and coworkers (Edwards et al., 2013b) determined that NV1042 (referred to as PA25) was able to specifically target HPV episomes. Notably, it

was also shown via Southern blots and other methods that the elimination of HPV episomes was mediated by the DDR. An inactive 8-ring hairpin PA was also examined via similar methods and determined to not produce a DDR. Thus, the DDR is important for both the lifecycle of HPV and the antiviral mechanism of large hairpin PAs. Implications of the results on the antiviral mechanism (i.e. the DDR) are discussed in further detail in **Chapters 3** and **4**.

### **1.5.2. Previous biophysical work on anti-HPV polyamides**

Studies of antiviral PAs NV1028 and NV1042 have been performed by this and the Bashkin group using DNase I footprinting via capillary electrophoresis (CE) and affinity cleavage to determine binding affinities and DNA binding patterns (Castaneda et al., 2017; He et al., 2014; Vasilieva et al., 2016). In general, despite both the large antiviral PAs being designed to bind to specific sequences using Dervan's selectivity rules, they have been shown to tolerate several mismatches and can bind sequences with as many as 4 mismatches at low nM affinities. Not only do replicable footprinting and affinity cleavage patterns show significant coverage of the viral DNA fragments used; conversely there are also specific portions of the DNA that have 4 or fewer mismatches that the PAs do not bind even at high concentrations of excess PA (**Fig. 1.13**, NV1042 shown). Additionally, isotherms collected via DNase I footprinting experiments were sigmoidal in shape and Hill coefficients indicate that PA may be binding to the DNA in a cooperative manner (Vasilieva, 2014).

Biophysical characterizations of NV1028 analogs have also been



**Figure 1.13** NV1042 sequencing map from (Vasilieva et al., 2016) in the 7348-122 LCR region of HPV16. Red arrows represent affinity cleavage sites observed at 50 nM NV1042-FeEDTA conjugate; size of the arrows indicate intensities relative to other arrows in a site, not to other sites. Yellow represents footprinting observed at 10 nM NV1042. Permission to reproduce is currently pending.



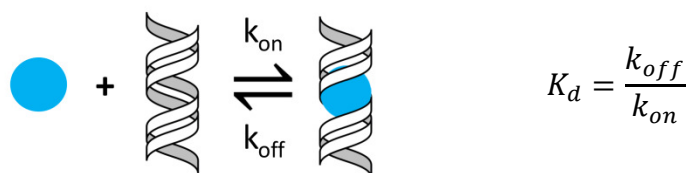
performed (Castaneda et al., 2017; Vasilieva et al., 2016). Castaneda and coworkers showed that, despite having significantly different antiviral activity, the DNA binding affinities of NV1028, NV1028-TMG, and NV1028-Guan are similar (Castaneda et al., 2017). However, despite the binding affinities being similar for all 3 PAs, the guanidinylated analogs prefer to bind to different sequences than NV1028: NV1028-Guan binds to the most sites on the DNA fragment tested when compared to NV1028 and NV1028-TMG. In other work on NV1028, Vasilieva and coworkers showed that the addition of a chiral  $\gamma$  turn does not strongly influence DNA binding orientation of NV1028 (Vasilieva et al., 2016) despite previous results of 8- and 10-ring hairpins with a chiral  $\gamma$  showing otherwise (Herman et al., 1998; Reddy, P. M. et al., 2005; Rucker et al., 2003).

While NV1028 and its analogs have been well-studied, NV1042 characterization has been limited due to challenges in synthesis and purification (Castaneda et al., 2017). In the singular biophysical study of NV1042, it was revealed through analysis of NV1042 DNA binding patterns that even with two Im present, G was not required for NV1042 to bind to DNA with high affinity (Vasilieva et al., 2016). Similar to NV1028 results, NV1042 is capable of binding to sites that contain multiple mismatches with low nM affinity. Additionally, replicable footprinting patterns show areas where broad coverage is observed and, like NV1028, there are areas where no coverage is observed even at high concentrations of excess PA (**Fig. 1.13**). Notably, while some overlap is observed, the areas of coverage are different for NV1028 and NV1042.

## 1.6. DNA-Ligand binding kinetics

### 1.6.1. Basics of ligand binding kinetics

During initial structure activity relationship (SAR) studies, focus is typically placed on target selectivity and specificity—increased selectivity of the drug and/or a more selective delivery method can reduce non-target binding and possibly limit undesired adverse effects (Cusack et al., 2015; Swinney, 2004). Selectivity and specificity are generally determined by and discussed in terms of binding affinities ( $K_d$ ). Even though  $K_d$  is a function of the association and dissociation rate constants ( $k_{on}$  and  $k_{off}$ , respectively) (**Scheme 1**), the individual kinetic rate constants and their contributions to binding affinity are largely ignored during initial structure-activity relationship studies. However, it has been shown that in vitro  $k_{on}$  and  $k_{off}$ , rather than in vitro  $K_d$ , can be better predictors of in vivo drug action (Copeland, 2016; Copeland et al., 2006; Cusack et al., 2015; Vilums et al., 2013)



**Scheme 1** Ligand binding reversibly to DNA

The rate of diffusion places an upper limit on  $k_{on}$ .  $k_{on}$  is concentration dependent: an increase in dosage would be an effective way to compensate for a slower  $k_{on}$  (Copeland et al., 2006). Additionally,  $k_{on}$  values do not always

translate to better drug efficacy as faster  $k_{on}$  values can result in more off-target toxicity (de Witte et al., 2016). In addition to the rate of diffusion, the rate of desolvation of the drug molecule, molecular orientation, and what is required in vivo for the drug to reach its target, are also obstacles medicinal chemists face when considering  $k_{on}$  (Copeland et al., 2006).

In the past decade, more focus has been placed on  $k_{off}$  as a metric for lead optimization because it is independent of the restraints placed on  $k_{on}$  by diffusion. Specifically, more emphasis has been placed on residence times or drug lifetimes (RT), the time a ligand remains bound to its target receptor, and its relation to drug efficacy and safety. RT can be defined as either the ligand-receptor complex relaxation time ( $1/k_{off}$ ) or as the ligand-receptor dissociative half-time ( $\ln[2]/k_{off}$ ) (Copeland et al., 2006; Cusack et al., 2015). Ideally, longer residence times would result in a longer therapeutic response duration and lower concentrations of the drug required to elicit a therapeutic effect.

An example of kinetics-driven drug discovery is the drug Finasteride (Proscar and Propecia) (Copeland et al., 2006). This drug blocks the production of dihydrotestosterone (DHT) by inhibiting  $5\alpha$ -reductase. It binds tightly with a  $K_i$  of 0.3 pM which is largely driven by an extremely slow dissociation rate constant. While this is technically considered a reversible inhibitor, Finasteride has a half-life greater than 30 days and there is some debate over whether the return of  $5\alpha$ -reductase function is the result of the drug disassociating or the production of new enzyme.

More recently, researchers studying Bruton's tyrosine kinase (BTK) identified inhibitors with tunable residence times (Bradshaw et al., 2015). To do this, Bradshaw et al. targeted the noncatalytic cysteine (Cys 481) with an inverted cysteine-reactive cyanoacrylamide electrophile. By varying the linker and the structure of the functional group, residence times of up to 7 days were observed *in vitro* and up to 18 hours *in vivo*.

### **1.6.2. DNA binding kinetics of Distamycin A and Netropsin**

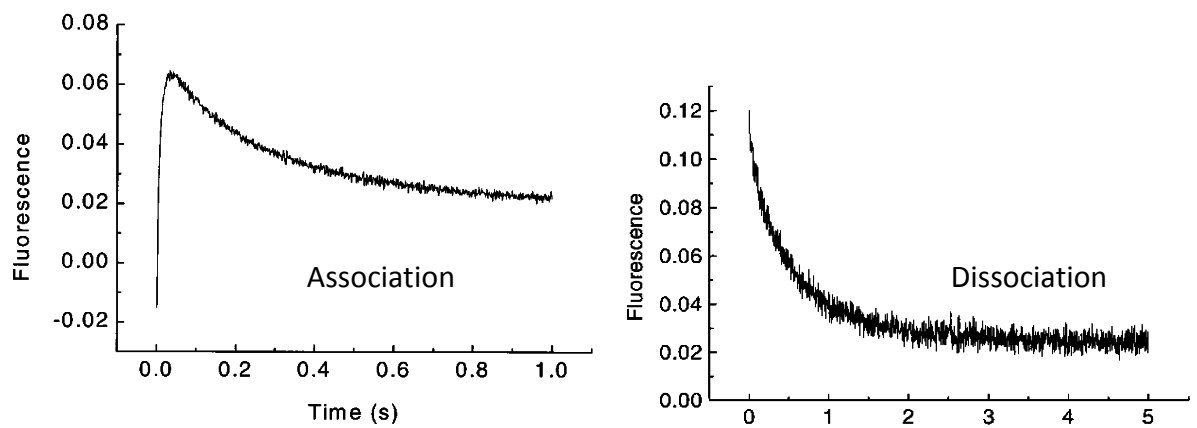
Thermodynamic characterization of DNA binding by Dst and Net is extensive. It has been shown that DNA binding by both molecules is easily detected through a number of different methods, such as CD, surface plasmon resonance (SPR), and NMR (Baliga & Crothers, 2000b; Chen & Sha, 1998; Fagan & Wemmer, 1992; Gambari et al., 2000). However, characterization of the DNA binding kinetics by both molecules is limited. Most of the Dst and Net DNA binding kinetics literature available tends to focus on the DNA binding kinetics of other analytes in their presence.

Two studies were performed by Baliga and Crothers of Dst binding kinetics with fluorescence spectroscopy by observing the changes in the intrinsic fluorescence of Dst (Baliga & Crothers, 2000a, 2000b). The first study characterized the sequence-dependent DNA binding kinetics of Dst for 1:1 binding. They determined that macroscopic association and dissociation rate constants were influenced by both the number of mismatches (where an A/T was

replaced by a G/C) and by the location of the mismatch (Baliga & Crothers, 2000a).

The second study by Baliga and Crothers focused on Dst DNA binding kinetics for Dst using sequences where monomeric or dimeric states were preferred (Baliga & Crothers, 2000b). It was shown that 2:1 binding had a unique effect on fluorescence signal (**Fig. 1.14**): a uniform change in fluorescence was observed for 1:1 complexes, but a change in direction was observed for 2:1 complexes which produced a biphasic curve. The second slower phase observed was determined to be inversely concentration dependent, which the authors speculated may be caused by non-specific transient binding where the molecule binds at a non-competitive binding site then quickly dissociates. DNA association rate constants for the 1:1 and 2:1 sites were both close to diffusion limited,  $\sim 70 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ . Biphasic DNA dissociation kinetics were also observed for 2:1 sites but were attributed to dissociation events of non-specific binding interactions. DNA dissociation kinetics traces did not change directions.

SPR is the current favored method for studying PA-DNA binding, but, to my knowledge, there are no published studies of Dst-DNA binding kinetics by SPR, only studies of Dst-DNA binding equilibrium (Gambari et al., 2000; Wang, S. et al., 2011). The lack of SPR data might be due to the weak binding affinity of Dst ( $\mu\text{M}$ ) or to binding kinetics that fall outside of the limits of detection. Additionally, there has been no report of biphasic binding kinetics for any small hairpin PAs studied via SPR.



**Figure 1.14** Kinetic trace of Dst-DNA binding where Dst is in excess. Left, the association curve of the 2:1 binding site. Right, the dissociation curve of the 2:1 binding site. Reproduced with permission from (Baliga & Crothers, 2000b) Copyright (2000) National Academy of Sciences, U.S.A.

### 1.6.3. Previous studies of hairpin PA-DNA binding kinetics

DNA binding kinetics have been characterized for several small hairpin PA, typically 10 rings or fewer throughout the past two decades. While there is at least one study that used fluorescence spectroscopy to characterize hairpin PA-DNA binding kinetics (Baliga et al., 2001), the favored method for studying PA-DNA binding kinetics is surface plasmon resonance (SPR); however, concerns about aggregation have limited studies of larger molecules by this method. The following is a discussion of the previous hairpin PA-DNA binding kinetics studies which is summarized in **Table 1.2** which is separated into 3 parts (A, B, and C). The **line numbers** refer to lines within **Table 1.2A-C**.

#### 1.6.3.1. Kinetics studies of linked versus unlinked polyamides

As discussed previously, hairpin PAs are 2 linear PAs linked by a  $\gamma$  to produce a hairpin structure capable of binding with greater affinity and specificity (Cho et al., 1995; Mrksich & Dervan, 1994). The earliest hairpin PA-DNA binding kinetics study examined the differences in association and dissociation rate constants for linear, hairpin, and cyclic PAs (Baliga et al., 2001). The linear monomer contained 4 rings and was designed to bind in a 2:1 mode, while the hairpin and cyclic PAs contained 8 rings (**lines 1-6, Table 1.2A**). The DNA association and dissociation rate constants were determined for these PAs with a matched and mismatched binding sequence. Association rate constants ( $k_{on}$ ) could not be determined directly and were calculated using directly determined equilibrium constants and dissociation rate constants ( $k_{off}$ ). This study shows that

linking the PAs does result in more favorable binding kinetics, which is likely the reason for the tighter binding affinities observed. However, association rate constants are similar for the hairpin PA and cyclic PA. Additionally, better specificity was observed for the hairpin PA compared to the cyclic given that the hairpin dissociates 75-fold more slowly from the matched sequence than the mismatched sequence; only a 30-fold difference in  $k_{\text{off}}$  is observed for the cyclic PA. Thus, this study found that kinetically there is no significant benefit of cyclic PAs over hairpin PAs.

Morinaga and coworkers used a cysteine to cyclize 6, 8, and 10-ring hairpin PAs resulting in molecules with both a  $\gamma$  and a cysteine linker (**lines 7-16**) (Morinaga et al., 2011). The cyclization reaction can produce both R and S isomers, but for ease of discussion only the values for the R-isomer are shown since R and S values and/or trends are similar. Binding affinities improved for the hairpin PAs as the number of rings increased. The weaker affinity of the cyclic 10-ring was attributed to a lack of flexibility resulting in a misaligned molecule. This hypothesis was supported given that the tightest binding affinities were observed for the 10-ring  $\beta$ -analogs with matched sites (**lines 13, 15**). Compared with previous studies, cysteine cyclic PAs bound weaker than hairpin PAs. Additionally, DNA binding rate constants were much slower for the 8-ring hairpin and cyclic PAs (**lines 9-10**) compared to the calculated DNA association rate constants for cyclic PAs (**lines 3, 5**). It is unclear whether these differences are the result of structural differences between the previously studied cyclic (2  $\gamma$  linkers) and the cysteine cyclic PAs, or are due to differences in the methods



**Table 1.2A** Summary of previous hairpin PA-DNA binding kinetics studies (linked versus unlinked polyamides)

Line	Size <sup>a</sup>	Structure <sup>b</sup>	Rec <sup>c</sup>	K <sub>d</sub> <sup>d</sup>	k <sub>on</sub> <sup>e</sup>	k <sub>off</sub> <sup>f</sup>	Reference
1	4		P	50 <sup>g</sup>	0.5	25±3	(Baliga et al., 2001)
2			M	800 <sup>g</sup>	0.5	396±21	
3	8		P	0.03 <sup>g</sup>	70	2±0.1	
4			M	2.5 <sup>g</sup>	60	151±5	
5	8		P	0.01 <sup>g</sup>	150	2±0.2	
6			M	0.8 <sup>g</sup>	80	65±3	
7	6		P	64	0.26	16	(Morinaga et al., 2011)
8	6		P	230	0.0077	1.8	
9	8		P	5.8	0.98	5.7	
10	8		P	62	0.03	1.9	
11	10		P	290	0.011	3.3	
12	10		P	5700	0.0056	32	
13	10		P	3.1	0.35	1.1	
14			M	3.7	0.3	1.1	
15	10		P	2	0.35	0.69	
16			M	8.2	0.35	2.9	
17	14		M	5.9	0.24	1.4	(Kashiwazaki et al., 2016)
18	14		M	7	0.22	1.6	

<sup>a</sup> number of rings; <sup>b</sup> Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop, cys cyclic closures (>), 3-(dimethylamino)propylamine (Dp), 3,3'-diamino-N-methyldipropylamine (Ta); <sup>c</sup> recognition sequence: P, perfect match; M, mismatch <sup>d</sup> K<sub>d</sub> (nM); <sup>e</sup> k<sub>on</sub> (10<sup>6</sup> M<sup>-1</sup> s<sup>-1</sup>), <sup>f</sup> k<sub>off</sub> (10<sup>-3</sup> s<sup>-1</sup>) <sup>g</sup> K<sub>d</sub> computed from k<sub>off</sub>/calculated k<sub>on</sub>

used to determine the values.

In the most recent work comparing the DNA binding kinetics of hairpin and cyclic PAs, DNA binding kinetic rate constants and DNA binding affinities were shown to be similar for a hairpin PA and a cysteine cyclic PA containing 14-rings (**lines 17-18**) (Kashiwazaki et al., 2016). Despite the larger size, there was no benefit to DNA binding affinity or kinetic rate constants when compared to the the 10-ring hairpin and cysteine cyclic PAs studied by the same group (**lines 13, 15**). The authors note that while cyclizing the PAs may not improve the biophysical parameters, it still helps to eliminate binding mode ambiguity, and can sometimes improve biological benefits (Li et al., 2013). To my knowledge, this is the largest hairpin PA for which the DNA binding kinetics have been published.

#### **1.6.3.2. Kinetics studies of PA-DNA sequence specificity**

An early paper by Henry and coworkers studied the sequence-dependent DNA binding kinetics of a 6-ring hairpin designed to bind to the sequence WWGGW (JH-37) (**lines 19-20, Table 1.2B**) (Henry et al., 2004). It was shown that JH-37 bound to its cognate sequence with nearly 5-fold stronger binding affinity than to a single mismatch sequence. The reduction in binding strength was the result of a much faster DNA dissociation rate constant, given that DNA association rate constants were similar for both the matched and mismatched sequences.

As discussed previously, while hairpin PAs have been shown to exhibit

**Table 1.2B** Summary of previous hairpin PA-DNA binding kinetics studies (DNA specificity)

Line	Size <sup>a</sup>	Structure <sup>b</sup>	Rec <sup>c</sup>	K <sub>d</sub> <sup>d</sup>	k <sub>on</sub> <sup>e</sup>	k <sub>off</sub> <sup>f</sup>	Reference
19	6		P	36	3.3	90	(Henry et al., 2004)
20			M	154	5.9	780	
21	6		P	588	0.27	160	(Flores et al., 2006)
22	6		P	40	0.05	2	
23	8		P	110	0.36	40	(Zhang et al., 2006)
24			M	2800	0.13	360	
25	8		P	390	0.07	27	(Zhang et al., 2006)
26			M	3400	0.04	140	
27	10		P	6.8	1.9	13	(Zhang et al., 2012)
28			M	260	0.65	170	
29	10		P	20	1.8	36	(Zhang et al., 2012)
30			M	360	0.58	210	
31	12		P	220	0.0079	1.8	(Taylor Rhys et al., 2013)
32			M	350	0.0063	2.2	
33			M	380	0.0038	1.4	

<sup>a</sup> number of rings; <sup>b</sup> Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop, cysteine cyclic closures (>), mitonafide (Mit), acetyl (Ac), indole (Ind), 3-(dimethylamino)propylamine (Dp); <sup>c</sup> recognition sequence: P, perfect match; M, mismatch <sup>d</sup> K<sub>d</sub> (nM); <sup>e</sup> k<sub>on</sub> (10<sup>6</sup> M<sup>-1</sup> s<sup>-1</sup>), <sup>f</sup> k<sub>off</sub> (10<sup>-3</sup> s<sup>-1</sup>)

sequence specificity, the canonical residues do not discriminate between A/T and T/A bp. Developing more sequence-specific PAs remains a focus of both the Dervan and Sugiyama groups. However, of the two groups only Sugiyama's group has focused on characterizing the sequence-dependent PA-DNA binding kinetics. Zhang and coworkers examined the effects of adding a chiral  $-OH$  or  $NH_2$  to the  $C\alpha$  of  $\gamma$  of an 8-ring hairpin PA (**lines 23-26**) (Zhang et al., 2006). It was determined that the (S)- $\gamma^{\alpha-OH}$  (**lines 23-24**) and (S)- $\gamma^{\alpha-NH_2}$  (**lines 25-26**) provided more specificity for TA bp than the (R)- $\gamma^{\alpha-OH}$  and (R)- $\gamma^{\alpha-NH_2}$ , respectively (not shown). The weakened  $K_d$  was the result of much faster  $k_{off}$  for AT (mismatched) versus TA (matched).

A second study by Zhang and coworkers examined sequence specificity of  $\gamma^{\beta-OH}/\beta$  pairs within 10-ring hairpin PAs (**lines 27-30**) (Zhang et al., 2012). Similar to the previous study, the S isomers showed more sequence specificity than the R-isomers (not shown). (S)- $\gamma^{\beta-OH}$  showed preference for TA regardless of position: there was no difference in specificity for TA whether (S)- $\gamma^{\beta-OH}/\beta$  was adjacent to Py or Im, or if the pair was within a certain proximity of the N- or C-termini. Consistent with the results for the 8-ring (S)- $\gamma^{\alpha-OH}$ , the weaker  $K_d$ s observed for the mismatched sites were primarily due to faster DNA dissociation rate constants

In a recent study, DNA binding kinetics were examined for a 12-ring hairpin PA with indole at its C-terminus (Taylor Rhys et al., 2013). The indole was used as a linker for an alkylating moiety, but the binding kinetics were determined only for a minimized structure (without the alkylating moiety) with a perfect, single

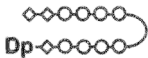
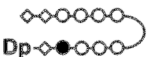

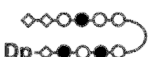
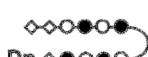


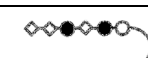
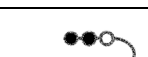
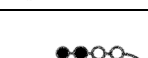
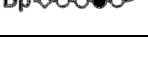
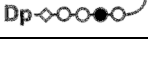
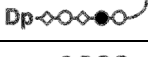
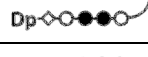
mismatch, and double mismatch binding sequences (**lines 31-33**). DNA binding affinity only differed slightly for the mismatched sites (**lines 32-33**) and was due to a small reduction (less than 2-fold) in association rates: perfect>single>double.

### **1.6.3.3. Kinetics studies of $\beta$ -alanine and hairpin flexibility/rigidity**

Torsional strain on PA structures is typically observed for PA with greater than 4 consecutive rings. Im/Py and Py/Im bind selectively to G/C and C/G, respectively. Given these facts, one may infer that an 8-ring hairpin PA can contain up to 4 Im/Py or Py/Im pairs to bind to GC-rich sequences. However, Han and coworkers showed that for an 8-ring hairpin PA with 4 Im (maximum for an 8-ring hairpin), increased planarity of the molecule was observed resulting in slower DNA association rates and, thus, weaker binding affinities (**lines 34-38, Table 1.2C**) (Han et al., 2012). It is believed that sequentially adding Im groups to a PA eventually induces a planarity in the PA conformation that is less favorable for DNA binding than that of an all-Py structure. Dissociation rate constants are similar regardless of the number of Im in the hairpin PA.

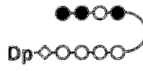
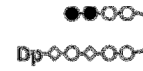



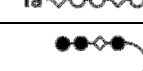
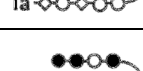
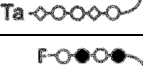
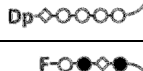
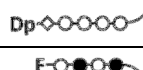
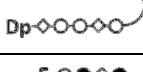
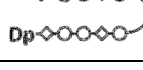
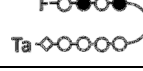
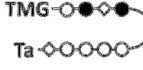
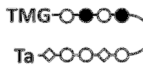
To relieve the torsional strain on larger PA, a Py can be substituted by  $\beta$ -alanine ( $\beta$ ), which acts as a flexible linker; however,  $\beta$  can influence hairpin PA-DNA binding affinity and kinetics depending on its position within the hairpin. In a follow-up study, Han and coworkers probed the effects of  $\beta$  position in an 8-ring hairpin PA designed to bind to the sequence WGCGCW (Han et al., 2013). Due to the symmetry of the binding sequence a single mismatch oligonucleotide was used as the control to characterize the influence of  $\beta$  in a hairpin PA binding to

**Table 1.2C** Summary of previous hairpin PA-DNA binding kinetics studies  
( $\beta$ -alanine, C-terminal functional groups, size dependence), part 1

Line	Size <sup>a</sup>	Structure <sup>b</sup>	Rec <sup>c</sup>	$K_d$ <sup>d</sup>	$k_{on}$ <sup>e</sup>	$k_{off}$ <sup>f</sup>	Reference
34	8		P	2.5±0.5	1.5±0.29	3.9±1.2	(Han et al., 2012)
35	8		P	3.7±2.3	1.2±0.17	4.3±2.5	
36	8		P	36±10	0.28±0.24	9.5±7.0	
37	8		P	48±4.7	0.30±0.045	14±0.8	
38	8		P	54±9.9	0.13±0.029	6.4±1.3	
39			M	270±52	0.082±0.042	22±6.3	(Han et al., 2013)
40	8		M	0.63±0.43	0.094±0.032	0.59±0.20	
41	8		M	0.70±0.54	13±10	9.1±4.0	
42	8		M	0.44±0.22	15±0.20	6.5±3.2	
43	6		P	7.1±0.7	3±1.5	21±5	
44	8		P	0.3±0.1	44±15	12±2	(Dupureur et al., 2012)
45 <sup>h</sup>			P	2.3±0.3	ND	7	
46	8		P	35±3	5±2	170±10	(Wang, S. et al., 2012)
47	8		P	0.9±0.2	5±2	6±1	(Bashkin, J. K. et al., 2013)
48	8		P	0.71±0.14	12	8.5	
49	8		P	106±11	tf <sup>i</sup>	tf <sup>i</sup>	


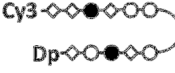
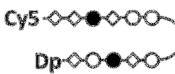
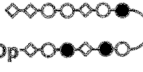
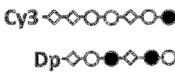
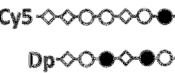
<sup>a</sup> number of rings; <sup>b</sup> Pyrrole (○), Imidazole (●);  $\beta$ -alanine (◇),  $\gamma$ -aminobutyric acid shown as a loop, 3-(dimethylamino)propylamine (Dp); <sup>c</sup> recognition sequence: P, perfect match; M, mismatch  
<sup>d</sup>  $K_d$  (nM); <sup>e</sup>  $k_{on}$  ( $10^6 M^{-1} s^{-1}$ ), <sup>f</sup>  $k_{off}$  ( $10^{-3} s^{-1}$ ) <sup>h</sup> collected via fluorescence spectroscopy, ND not determined; <sup>i</sup> tf—too fast to be determined

**Table 1.2C** Summary of previous hairpin PA-DNA binding kinetics studies ( $\beta$ -alanine, C-terminal functional groups, size dependence), part 2

Line	Size <sup>a</sup>	Structure <sup>b</sup>	Rec <sup>c</sup>	$K_d$ <sup>d</sup>	$k_{on}$ <sup>e</sup>	$k_{off}$ <sup>f</sup>	Reference	
50	8		P	2.5±0.4	4.7±0.7	12±1	(Wang, S. et al., 2014)	
51	8		P	36.4±0.4	96±6	2.6±0.3		
52	8		P	12.1±0.4	3.7±0.4	45.7±7		
53	8		P	2.0±0.2	1.3±0.2	2.6±0.2		
54	8		P	18.6±0.8	0.7±0.2	13±2		
55	8		P	26.3±2.2	1.1±0.2	29±5		
56	8		P	2.6±0.2	1.8±0.3	4.7±0.3		
57	8		P	1.3±0.01	15.8±0.3	7.6±0.1		(Liu, B. et al., 2017)
58	8		P	0.67±0.01	10±1	6.9±0.1		
59	8		P	1.62±0.04	21±10	34.1±0.5		
60	8		P	0.54±0.01	15±1	8.4±0.1		
61	8		P	tf	tf	79±4		
62	8		P	0.16±0.02	50±0.1	7.9±0.1		
63	8		P	tf <sup>i</sup>	Tt <sup>i</sup>	35±2		
64	8		P	1.65±0.03	32±0.1	52.7±0.6		

<sup>a</sup> number of rings; <sup>b</sup> Pyrrole (○), Imidazole (●);  $\beta$ -alanine (◇),  $\gamma$ -aminobutyric acid shown as a loop, tetramethylguanidium (TMG), formamido (F), 3-(dimethylamino)propylamine (Dp), 3,3'-diamino-N-methyldipropylamine (Ta); <sup>c</sup> recognition sequence: P, perfect match; M, mismatch <sup>d</sup>  $K_d$  (nM); <sup>e</sup>  $k_{on}$  ( $10^6 M^{-1} s^{-1}$ ), <sup>f</sup>  $k_{off}$  ( $10^{-3} s^{-1}$ ); <sup>i</sup> tf—too fast to be determined

**Table 1.2C** Summary of previous hairpin PA-DNA binding kinetics studies( $\beta$ -alanine, C-terminal functional groups, size dependence), part 3

Line	Size <sup>a</sup>	Structure <sup>b</sup>	Rec <sup>c</sup>	$K_d$ <sup>d</sup>	$k_{on}$ <sup>e</sup>	$k_{off}$ <sup>f</sup>	Reference
65	8		P	1.0±0.12	8.3±4.1	8.7±3.9	(Han et al., 2014)
66	8		P	350±88	0.11±0.04	38±6.2	
67	8		P	64±50	0.59±0.44	38±19	
68	10		P	0.6±0.4	2.9±1.3	1.7±1.4	
69	10		P	310±71	0.064±0.0013	20±5.0	
70	10		P	510±32	0.055±0.0094	28±3.9	

<sup>a</sup> number of rings; <sup>b</sup> Pyrrole (○), Imidazole (●);  $\beta$ -alanine (◇),  $\gamma$ -aminobutyric acid shown as a loop, Cy3 fluorophore (Cy3), Cy5 fluorophore (Cy5), 3-(dimethylamino)propylamine (Dp); <sup>c</sup> recognition sequence: P, perfect match; M, mismatch <sup>d</sup> $K_d$  (nM); <sup>e</sup>  $k_{on}$  ( $10^6 M^{-1} s^{-1}$ ), <sup>f</sup>  $k_{off}$  ( $10^{-3} s^{-1}$ )



mismatched sequences (**lines 39-42**). Interestingly, while  $K_d$ s were similar, the binding kinetics were very different for the different  $\beta$ -analogs. A single  $\beta$  replacement resulted in at least 150-fold increase in  $k_{on}$  values (**lines 40-41**), while a double  $\beta$  replacement resulted in only a 10-fold increase in  $k_{on}$  (**line 42**). Conversely, a single  $\beta$  replacement resulted in only a 2-3-fold reduction of  $k_{off}$ , while a double  $\beta$  replacement reduced  $k_{off}$  values by over 30-fold. This trend was consistent for all the sequences studied (not shown).

The Wilson group has the largest breadth of work characterizing the influence of  $\beta$  on the hairpin PA-DNA binding kinetics using 8-ring hairpin PAs (**lines 44-64**). In contrast to what was reported by the Han and coworkers, the addition of  $\beta$  at either the N- or C-terminus resulted both an increase and in a reduction in binding affinity of 3-100-fold (Bashkin, J. K. et al., 2013; Liu, B. et al., 2017; Wang, S. et al., 2014; Wang, S. et al., 2012). The loss of DNA binding affinity could not be attributed to only one cause as the addition of  $\beta$  could result in slower association rate constants (**line 54**) or faster dissociation rate constants (**line 47, 53**), and resulted in a combination of both (**lines 46, 51, 55**). Improved binding affinities were the result of improved dissociation rate constants (**line 58**), but also the result of unfavorable effects on both binding rate constants that mathematically yielded a more favorable binding affinity (**line 59**). Additionally, there does not seem to be a strong correlation in DNA binding kinetics when  $\beta$  is paired with Py, Im, or another  $\beta$ . The effects of  $\beta$  on DNA binding kinetics is—at best—complicated.

#### 1.6.3.4. Kinetics studies of PA N- and C-terminal modifications

Another method groups have used to improve PA design is to add functional groups to the N- or C-termini of the PAs. In terms of improving DNA binding kinetics, this method has been met with varying success. The previously discussed JH-37 was modified with an intercalating molecule, mitonafide (Mit) conjugated to the N-terminal Py (**line 22**) (Flores et al., 2006). A tighter  $K_d$  was observed for Mit-JH-37 than JH-37 which was attributed to Mit-JH-37 dissociating from DNA more slowly (Flores et al., 2006) (**lines 21-22**). Given that the same biotinylated oligonucleotide sequence are used in both studies and that the method are the same, it is unclear why there is such a large discrepancy between biophysical parameters determined by Henry (**line 19**) and then later by Flores (**line 21**).

The addition of C-terminal fluorophores, Cy3 and Cy5, was examined by Han and coworkers using 8- and 10-ring hairpin polyamides (Han et al., 2014) (**lines 65-70**). The addition of a C-terminal fluorophore weakened binding affinity due to a significant reduction in DNA association rate constants. Dissociation rate constants also increased, but were small (10-fold) in comparison to the differences in association rate constants (50-100-fold).

Wang and coworkers studied hairpin PA modified with an N-terminal tetramethylguanidinium (TMG) or formamido (F) (**lines 57-64**) (Liu, B. et al., 2017). It is unclear whether the differences in DNA binding affinities and kinetics are due to the different N-terminal groups or because the F-analogs have Dp at their C-termini and the TMG-analogs have Ta. In a previous study by Wang the

Ta (**line 53**) and Dp (**line 50**) analogs bound with similar binding affinities, but with slightly different binding kinetics. Despite there being published data for a number of Dp and Ta 8-ring hairpin PAs these are the only 2 PA with identical Py/Im/ $\beta$  patterns, thus it is difficult to draw a conclusion about the effects on binding kinetics of Ta versus Dp C-terminal groups.

#### **1.6.3.5. Studies of PA-DNA size-dependent binding kinetics**

To my knowledge there have been no studies of PA-DNA binding kinetics with respect to size: studies tend to focus on comparing PAs of the same size with variations to their structure. In the 2 studies that compare the DNA binding kinetics of PAs of different sizes only 2 sizes are compared. The first study, Wang and coworkers, compared a 6- and 8-ring hairpin PA (**lines 43-44**) (Wang, S. et al., 2012). Weakened binding affinity is observed for the 6-ring and is due to a much slower  $k_{on}$  (14-fold) and only a slightly faster  $k_{off}$  (less than 2-fold). In the second study, Han and coworkers compare 3 sets of 8- and 10-ring hairpin PAs (**lines 65-70**) (Han et al., 2014). Tighter binding affinities are observed for 2 of the 10-ring hairpins, while weaker binding is observed for the third. The differences in binding affinities are not due to a change in a single parameter, nor are the changes in the parameters consistently favorable or unfavorable.

Despite there being a number of published PA-DNA binding kinetics studies, they tend to focus on small hairpin PAs; data is limited for PAs with greater than 10-rings. Additionally, the diversity of PA structures and sequences makes it difficult to make a comparison between the different studies because

many of the PA have multiple differences in their structures (chiral residues, C-termini, N-termini, etc). Varying Im/Py/ $\beta$  content makes it especially hard to compare data since we do not know the precise kinetic contributions of the individual residues. Given these issues, it is difficult to draw a conclusion about size dependence from current published data.

## **1.7. PA-DNA binding stoichiometry**

### **1.7.1. Evidence of high linked PA-DNA binding stoichiometry in the literature**

As discussed previously, sequence dependent PA-DNA binding stoichiometry is observed for some linear polyamides (Chen & Sha, 1998; Pelton & Wemmer, 1990; Wemmer et al., 1994), such as Dst, but not for others, such as Net. And Dst's ability to form 2:1 complexes with DNA is part of what inspired the circular, hairpin, and H-pin PA design of modern PAs. While their structural linkages may be different, they are believed to bind to DNA similarly.

A common method used to confirm PA-DNA binding is circular dichroism spectroscopy (CD), which is used for both linear and hairpin PAs. PA-DNA binding is commonly observed in the 300-370 nm range: no CD signals in that range are observed for DNA when PA is not present, nor by PA alone (Lyng et al., 1992; Pilch et al., 1996). The published CD spectra typically show incremental titrations ranging from 0-2 equivalents of PA with a DNA duplex or hairpin which contains a single isolated PA binding site. Given our understanding of how these linked PAs should bind to DNA, saturation should be observed at 1:1. However, for some hairpin and H-Pins, this is not always the case (**Figs**

**1.15-22**). Notably, while there is much discussion about the PA-DNA binding stoichiometry of linear polyamides and/or Dst-inspired small molecules, further scrutiny of DNA binding stoichiometry and DNA binding modes has not occurred despite the published data indicating non-canonical binding may be happening. The following is a summary of the literature where CD is used to investigate DNA binding of H-pin and hairpin PAs where PA-DNA binding stoichiometry ratios can be gleaned from the published data. For convenience, **Table 1.3** contains an index of the papers and figures discussed in this section.

A study by O'Hare and coworkers of 6-ring H-pin PA designed to bind to the site ACGCGT showed greater than 2:1 PA-DNA binding via CD (**Fig. 1.15**) (O'Hare et al., 2007). The authors note that this higher PA-DNA binding stoichiometry is likely due to non-specific binding due to the high concentrations required for CD and the PAs tendency to aggregate.

Another study of 6-ring H-Pins by Burckhardt and coworkers also used CD to examine PA-DNA binding (**Fig. 1.16**) (Burckhardt et al., 2000). In this study not only did the CD spectra show unlinked monomers saturate the DNA at greater than 2:1, but also the CD spectra of the H-Pins shows saturation above 2:1. The authors do not comment on this high PA-DNA binding stoichiometry.

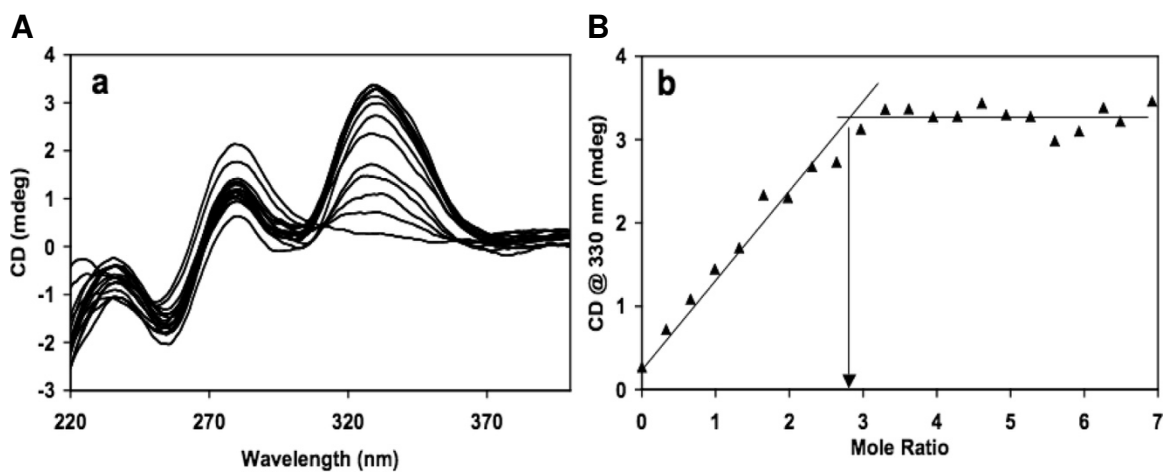
Pilch and coworkers studied the thermodynamics of a 6-ring hairpin PA designed to bind to the site TGTTA when bound to its perfect site (duplex 1), and 2 single mismatch sites: TGGTA (duplex 2), TATTA (duplex 3) (**Fig. 1.17**) (Pilch et al., 1996). While tighter binding was observed for duplex 1, binding was still observed for duplex 2 and duplex 3. Interestingly, while CD spectra for duplex 1

**Table 1.3** Index of published high PA-DNA binding stoichiometry

<b>Fig</b>	<b>Size</b>	<b>Type</b>	<b>Note</b>	<b>Reference</b>
<b>1.15</b>	6	H-Pin	H-Pin bound to 3 duplex of different sequence. 2 of 3 CD titrations published show greater than 1:1 stoichiometry. 1 of 3 shows at least 2:1. Shown CD titration of 6-ring H-Pin bound at least 2:1.	(O'Hare et al., 2007)
<b>1.16</b>	6	H-Pin	Multiple CD titrations show greater than 2:1 binding for H-Pin and linear monomers. Experiments were performed at 200 mM and 2M salt with 4 12-mer DNA duplexes. Shown: CD spectra of H-pin at 200 mM NaCl along with titration plots of H-Pin with other sequences.	(Burckhardt et al., 2000)
<b>1.17</b>	6	Hairpin	3 CD titrations of 6-ring hairpin bound to 3 duplexes of different sequence. Shown: 1 of 3 CD spectra shows no saturation at 1.4 equivalents.	(Pilch et al., 1996)
<b>1.18</b>	6	Hairpin	3 CD titrations of 6-ring hairpin bound to DNA hairpins of different sequence. Shown: all 3 published spectra show saturation at 2:1 or greater.	(Buchmueller et al., 2005)
<b>1.19</b>	8	Hairpin	4 CD titrations published of 1 6-ring and 3 8-ring hairpin PAs with the same DNA hairpin. Shown: All 3 8-ring PAs do not show saturation at 2 equivalents of PA.	(Wang, S. et al., 2012)
<b>1.20</b>	8	Hairpin	7 CD titrations published of 8-ring hairpin PAs with identical binding sites. At least 5 of 7 do not show saturation at 2 equivalents of PA. All spectra are shown.	(Wang, S. et al., 2014)
<b>1.21</b>	8	Hairpin	10 CD titrations published of 8-ring hairpin PAs with identical binding sites. At least 4 of the 10 CD titrations do not show saturation at 2 equivalents of PA. All spectra are shown.	(Liu, B. et al., 2017)

**Table 1.3, cont.** Index of published high PA-DNA binding stoichiometry

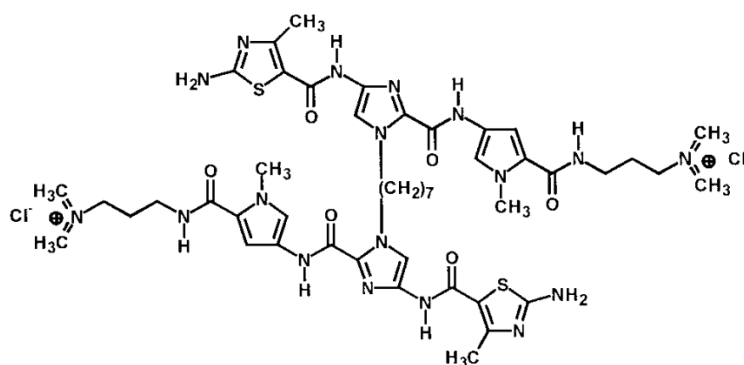
<b>Fig</b>	<b>Size</b>	<b>Type</b>	<b>Note</b>	<b>Reference</b>
1.22	14	Hairpin	6 CD titrations published of NV1028 (14-ring hairpin) and NV1028 with a chiral $\gamma$ . Saturation was observed above 1:1 in experiments performed without CHAPS. All CD titrations are shown.	(Castaneda, 2017)



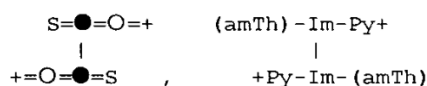
**Figure 1.15** A, CD spectra of f-ImPylm-EG-8 binding to 5'-ACGCGT-3'. B, CD response of the CD signal at  $\lambda_{\text{max}}$  (328 nm) versus the mole ratio of ligand to ACGCGT. 9  $\mu\text{M}$  DNA, 10 mM phosphate, 200 mM  $\text{Na}^+$ , 1 mM EDTA, pH 6.2. Reproduced with permission from (O'Hare et al., 2007). Copyright 2007 American Chemical Society.



A



3. Structure of C7 cross-linked polyamide dimer



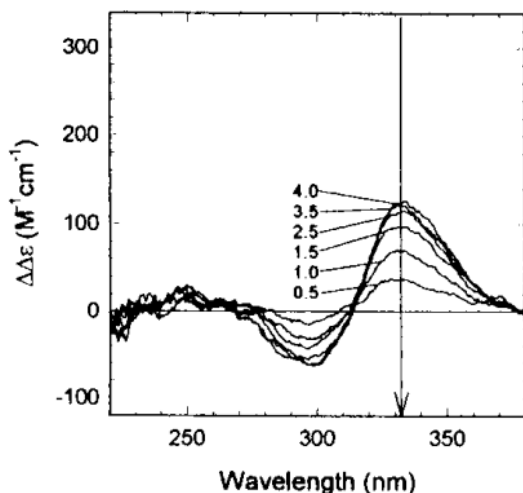
1, 5'-CGCAACGTTGCG-3'

3, 5'-CGCAGCGCTGCG-3'

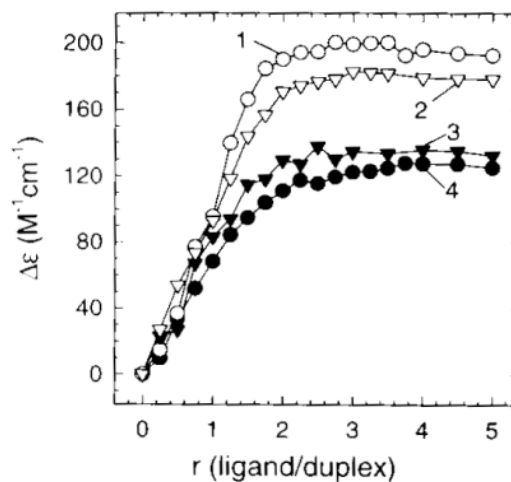
2, 5'-CGCAAAGTTGCG-3'

4, 5'-CGCAAATTTGCG-3'

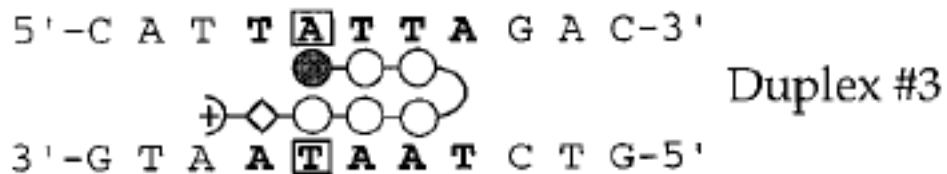
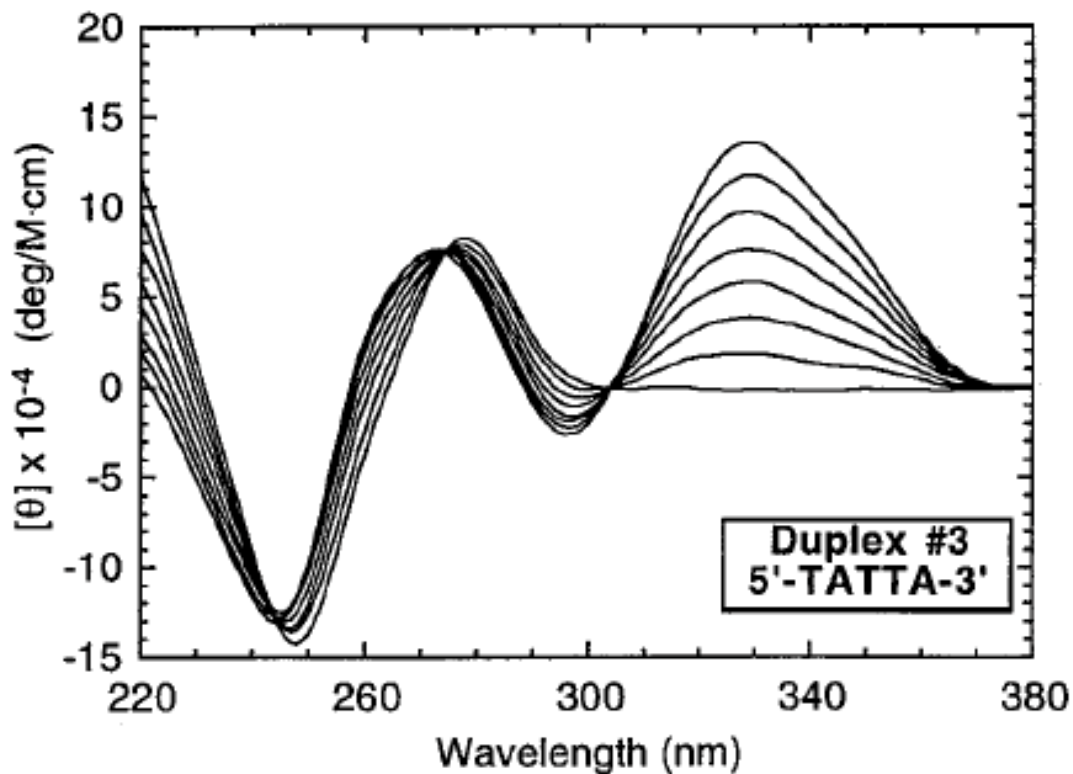
B



C



**Figure 1.16** CD titrations of a C7 cross-linked H-Pin polyamide. A, structure of the 7-ring H-Pin and the dodecamers used in CD titration experiments. B, CD titration of 6-ring H-Pin in complex with dodecamer 1. C, CD titration plots of 6-ring H-Pin with dodecamers 1-4. Adapted with permission from (Burckhardt et al., 2000) Copyright 2000 Taylor & Francis.



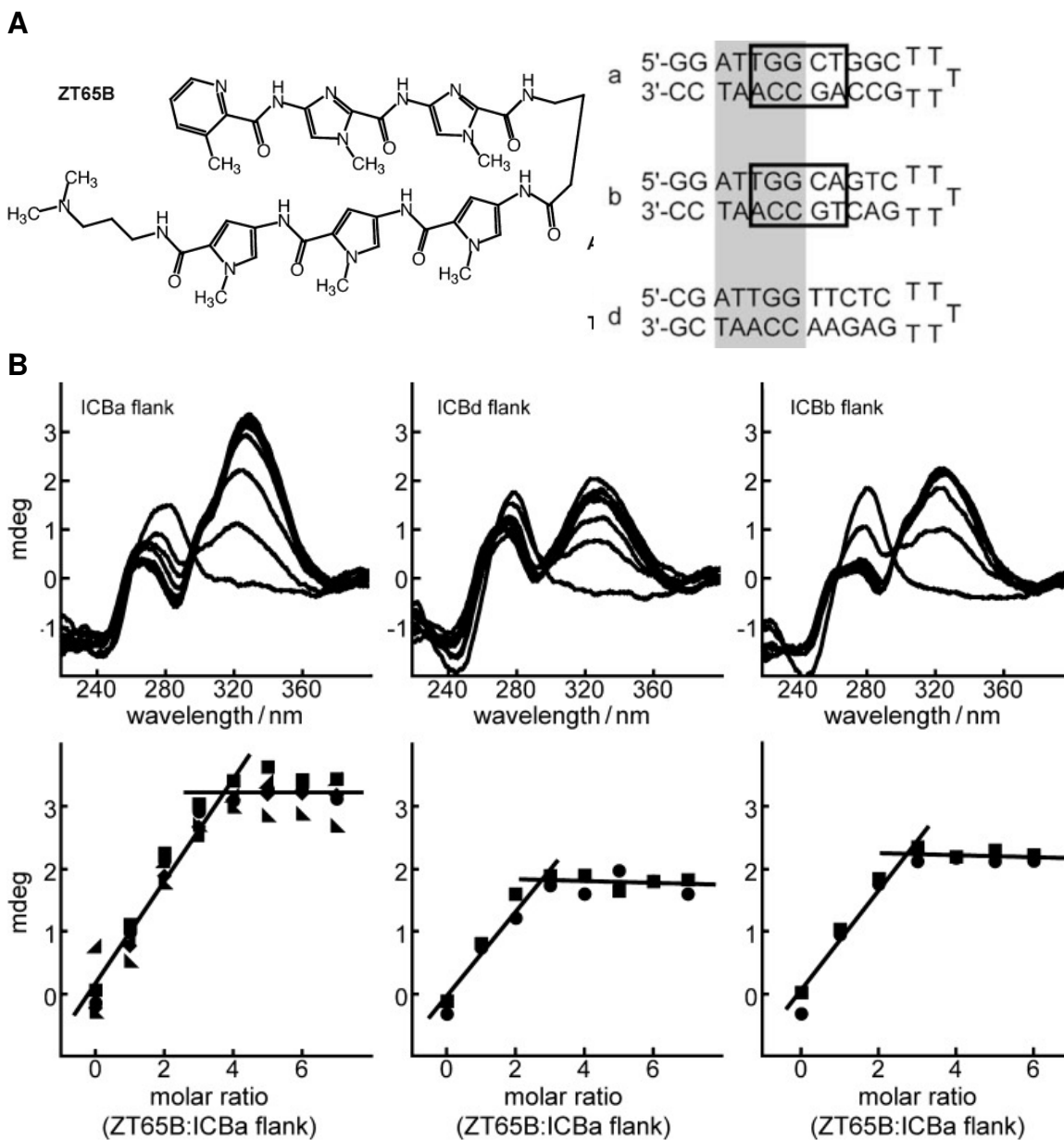
**Figure 1.17** CD spectra for 6-ring hairpin ImPyPyyPyPyPYβDp in single mismatch site. CD titration ends at 1.4 equivalents. Reproduced with permission from (Pilch et al., 1996) Copyright (1996) National Academy of Sciences, U.S.A..

and duplex 2 are consistent with 1:1 binding, the spectra for duplex 3 does not show saturation at 1.4 equivalents. The authors note that the PA likely binds to each duplex with modified structural and electronic properties due to sterics and/or different hydrogen bonding available by each sequence of DNA.

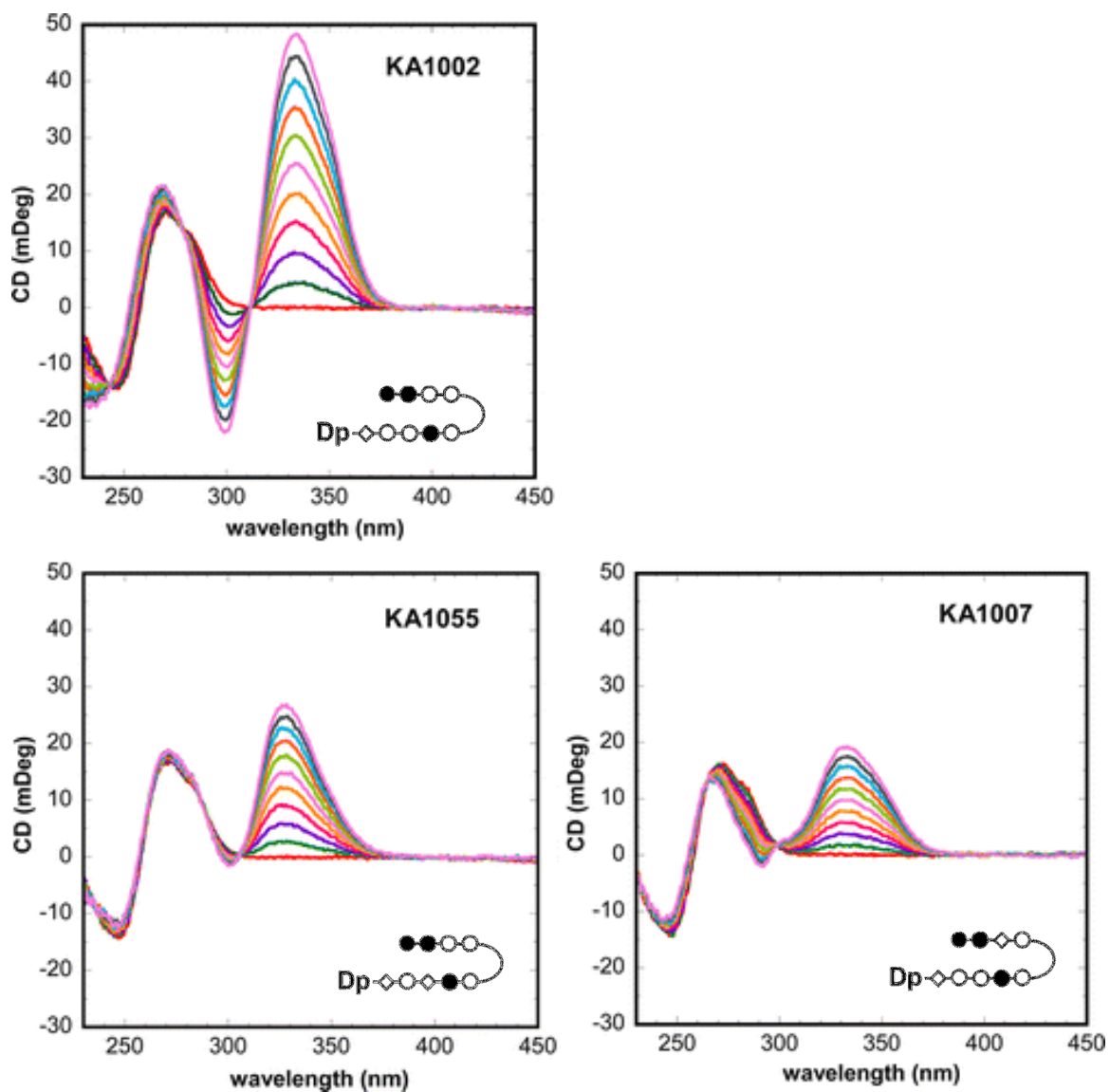
In a study of a 6-ring hairpin PA ( $MpImIm\gamma PyPyPy\beta Dp$ , where  $Mp$  is methylpicolinate), Buchmueller and coworkers observed high molar PA-DNA binding ratios when testing their hairpin with different sequences of DNA (**Fig. 1.18**) (Buchmueller et al., 2005). The authors dismiss these values as stoichiometric values since they believed greater than 2:1 binding was unlikely. Instead they attribute the high binding stoichiometric ratios to equilibration of free and bound states observed for molecules with weaker binding affinities.

A number of published CD spectra are of 6- and 8-ring hairpin PA examined by the Wilson group (**Fig. 1.19-21**) (Liu, B. et al., 2017; Wang, S. et al., 2014; Wang, S. et al., 2012). Greater than 1:1 binding is observable in over half of the published spectra as saturation is not observed at 2.0 eq of PA. The authors make no comment about the high stoichiometry.

CD spectra were collected by Carlos Castaneda of NV1028 and a chiral analog with a (R)- $NH_2$  at the  $\gamma$  (Castaneda, 2017) (**Fig. 1.22**). Using these two PAs, Castaneda examined DNA binding stoichiometries with respect to the PA binding orientations (forward or reversed, discussed in **Chapter 2**). **Fig. 1.22** includes the assumed binding orientation for each CD titration performed. For NV1028, greater than 1:1 binding was observed with saturation around 2:1. For NV1028 with the chiral  $\gamma$ , saturation was not observed until around 2.5:1.

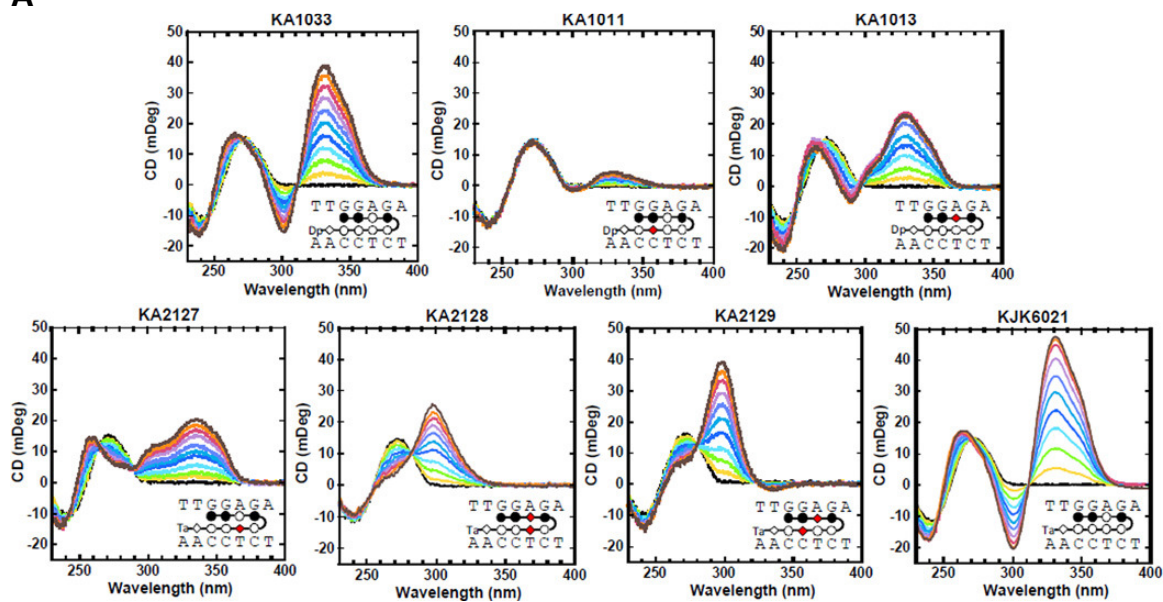


**Figure. 1.18** 6-hairpin with 3 DNA hairpins of different sequence. A, the hairpin PA and the DNA hairpins used in B. B, CD titrations of the 6-ring hairpin in complex with DNA. Adapted with permission from (Buchmueller et al., 2005) Copyright 2005 John Wiley & Sons, license number 4315710816292.

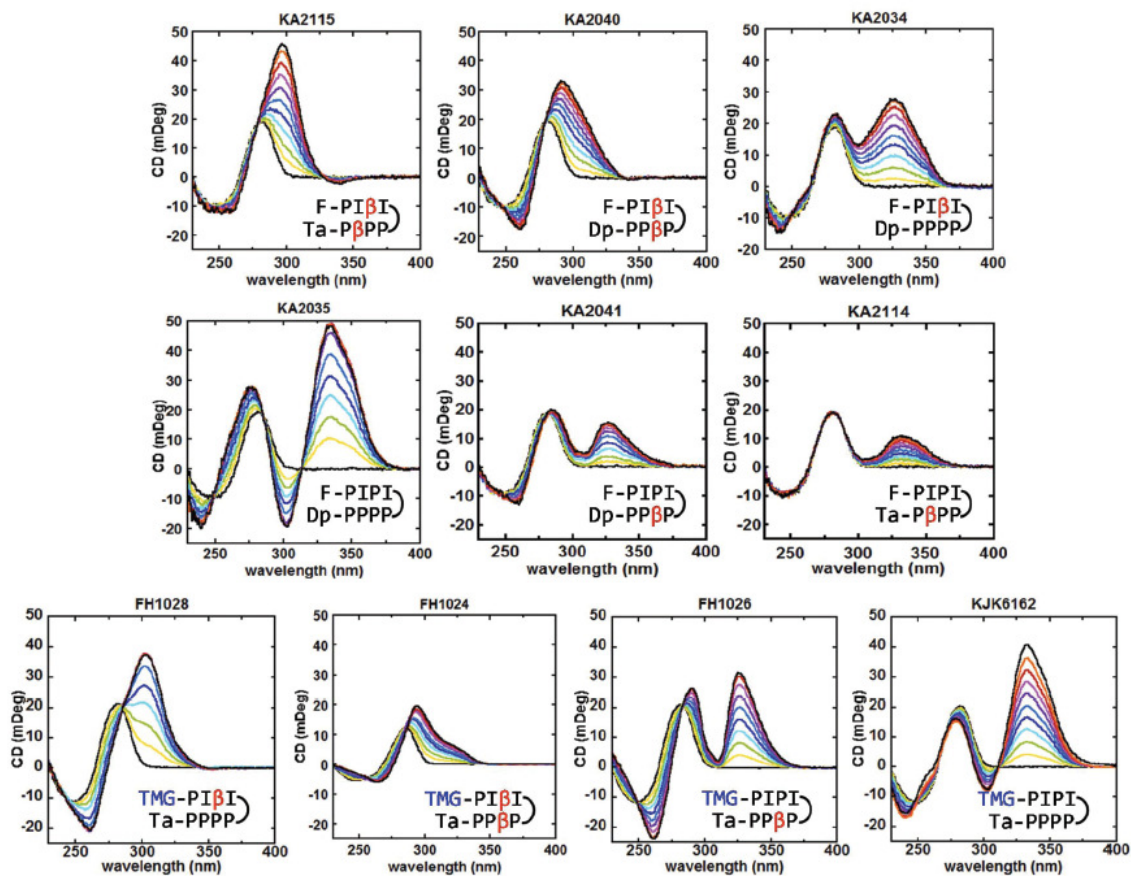


**Figure 1.19** CD spectra of 8-ring hairpin PAs titrated with up to 2.0 equivalents of PA. Reproduced with permission from (Wang, S. et al., 2012) Copyright 2012 American Chemical Society.

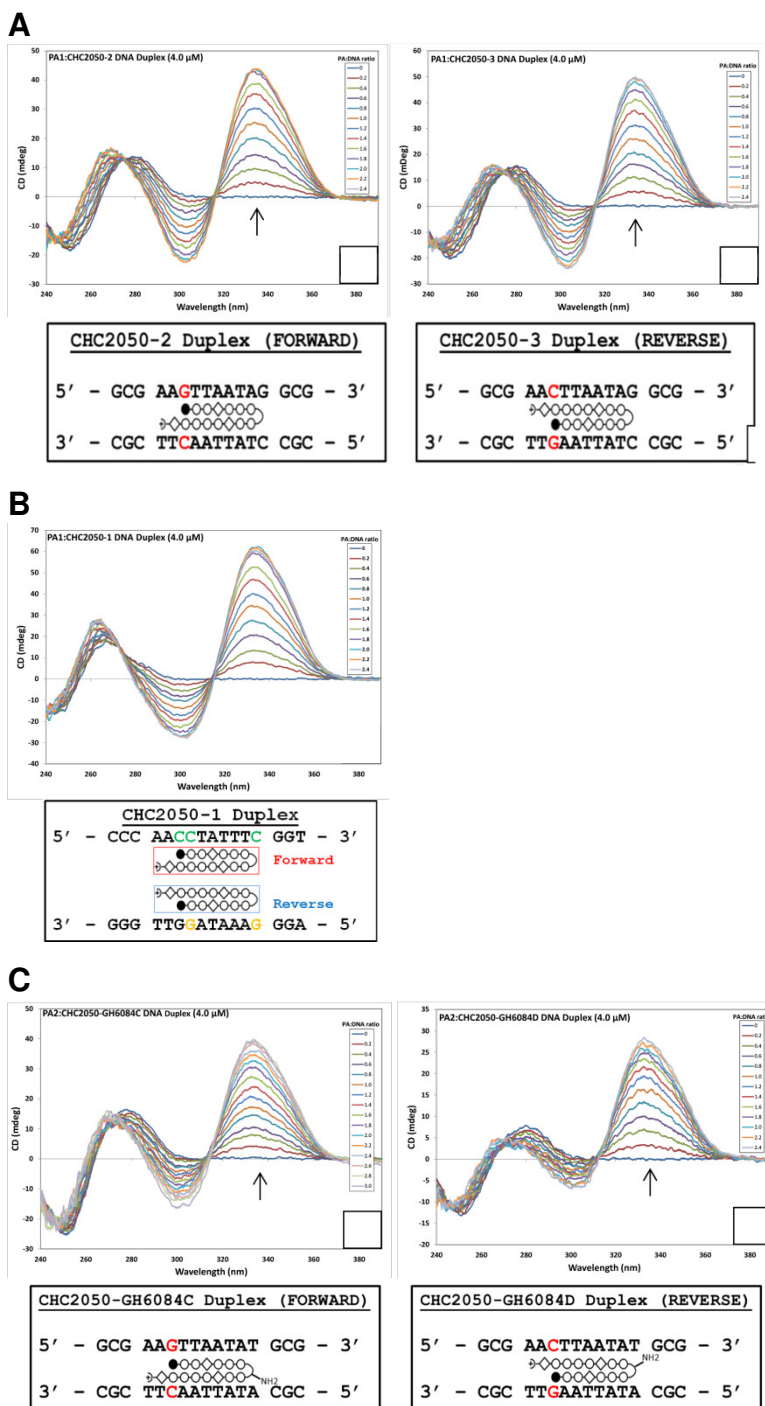
A



**Figure 1.20** Additional CD spectra from the Wilson group reproduced with permission from (Wang, S. et al., 2014) Copyright 2014 Royal Society of Chemistry, license number 4316400130362.



**Figure 1.21** Additional CD spectra from the Wilson group reproduced with permission from (Liu, B. et al., 2017) Copyright 2017 Royal Society of Chemistry, license number 4316400655823.



**Figure 1.22** CD titration experiments of NV1028 and NV1028 with chiral  $\gamma$ -(R)NH<sub>2</sub> by (Castaneda, 2017). A, CD titrations of NV1028 with a forward perfect match duplex, CHC2050-2 (left), and single mismatch forward (reverse perfect) (right). B, CD titrations of NV1028 with a forward triple mismatch (reverse double). C, NV1028 with chiral  $\gamma$  turn and forward perfect duplex, CHC2050-GH6084C (left), forward single mismatch (reverse perfect). Reproduced with permission from (Castaneda, 2017) Copyright 2017 Carlos H. Castaneda.



Two things are clear from looking at the published CD spectra: first, it is only some linked PAs—not all—that are capable of binding to DNA greater than 1:1; second, greater than 1:1 binding is observed for sequences where a single binding site is isolated, thus, the PAs are likely binding in a different orientation than the Dervan rules suggest.

### 1.7.2. Multiple binding events and cooperativity

While we tend to think of intermolecular events as a simple 1:1 model,  $R + L = RL$  (where R is the receptor with only 1 binding site, and L is the ligand), in nature this is not always the case and high DNA binding stoichiometric values are not uncommon. In these cases, R has more than 1 binding site and ligands can bind independent from or interacting on each other (**Fig. 1.23**) (Voet et al., 2013). For the purpose of this discussion receptor binding sites are limited to only 2 and will be discussed in terms of how binding relates to their association rate constants ( $k_1$  and  $k_2$ ), dissociation rate constants ( $k_{-1}$  and  $k_{-2}$ ), and binding affinities.

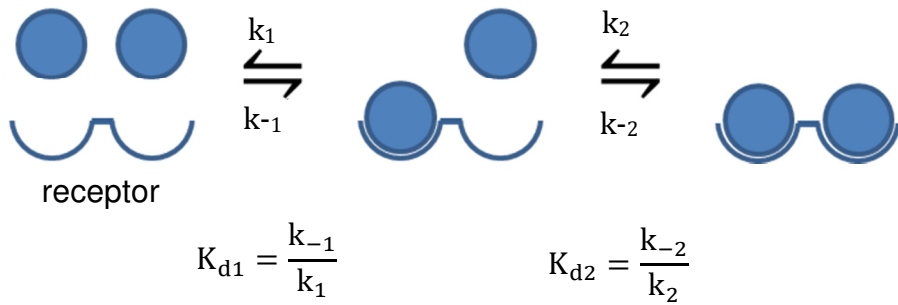
If 2 ligands bind to a receptor independently (**Fig. 1.23A**),  $k_1$ ,  $k_2$ ,  $k_{-1}$ , and  $k_{-2}$  are microscopic rate constants that all correspond to individual binding events and do not influence or affect one another. In this case the binding rate constants can be equal to each other ( $k_1=k_2$ ,  $k_{-1}=k_{-1}$ ) or different from each other ( $k_1 \neq k_2$ ,  $k_{-1} \neq k_{-2}$ ) as the binding of each ligand is an independent event.

In the interacting pathway, the ligands bind in a cooperative manner that has been described by 2 models: the concerted model (**Fig. 1.23B**) and the

independent

$$\begin{aligned}
 &k_1 = k_2 \text{ or } k_1 \neq k_2 \\
 &k_{-1} = k_{-2} \text{ or } k_{-1} \neq k_{-2} \\
 &K_{d1} = K_{d2} \text{ or } K_{d1} \neq K_{d2}
 \end{aligned}$$

A

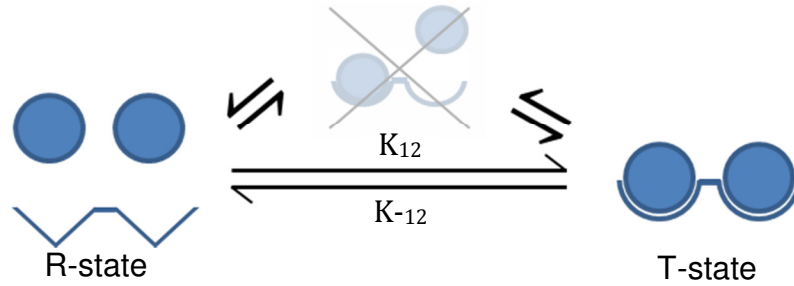


interacting

$$\begin{aligned}
 &k_1 \neq k_2 \\
 &k_{-1} \neq k_{-2} \\
 &K_{d1} \neq K_{d2}
 \end{aligned}$$

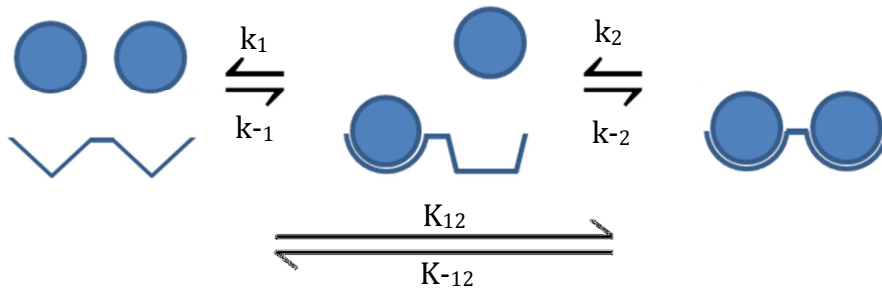
B

concerted



C

sequential



**Figure 1.23** Multisite receptor-ligand binding diagrams. A, 2 ligands bind the receptor independently with equal affinity. B, concerted interacting model. C, sequential interacting model.

sequential model (**Fig. 1.23C**). These models have been most studied and used to describe multimeric enzyme-ligand binding interactions, specifically hemoglobin, but the principles can be applied to describe multisite receptor-ligand binding interactions. In both these models, the receptor sites exist in 2 states: the relaxed state (R-state) which is the reactive species, and the tense state (T-state) which is the less reactive species. In the concerted model, the receptor contains multiple binding sites that all exist at equilibrium in either the T- or the R-state (Changeux, 2012; Monod et al., 1965). In a multisite model, such as hemoglobin, when no ligands are present, equilibrium favors the T-state. However, as the number of ligands in the binding sites increases, the equilibrium shifts to prefer the R-state. In this model, the receptor transitioning from R- to T-state involves multiple binding events, but intermediate complexes are disfavored. Thus, the binding affinities and binding rate constants are referred to as macroscopic, represented by  $K_{d12}$ , and  $K_{12}$  and  $K_{-12}$  as they are composites of the microscopic binding affinities and binding rate constants represented by  $K_{d1}$  and  $K_{d2}$ , and  $k_1$ ,  $k_{-1}$ ,  $k_2$ , and  $k_{-2}$ , respectively (**Fig. 1.23**).

In the sequential model, also referred to as the allosteric model, each of the binding sites are interconnected, but can exist in a different conformation than the other binding sites (**Fig. 1.23C**) (Changeux, 2012; Koshland et al., 1966). In this model, the first ligand binds to the receptor resulting in a conformational or chemical change of the second binding sites that makes binding of the second ligand more favorable. Unlike the concerted model, conversion from the R- to the T-state occurs as each binding event results in

changes to the adjacent binding sites in a sequential manner. Similar to the concerted model, the overall binding affinities and binding rate constants are macroscopic since the binding affinity of subsequent ligands is interacting with the preceding binding event. In this scenario, microscopic binding affinities and rate constants can be determined if a partially bound intermediate can be isolated.

Cooperativity can be either positive or negative. When cooperativity is positive, cooperativity is favorable and the complex will favor the cooperative binding mode; when cooperativity is negative then cooperativity is less favorable and the complex will favor a non-cooperative binding model. In some cases molecules are capable of binding in multiple modes and populations of complex species are observed, as in the case of Dst, and some other Dst-inspired small molecules (Pelton & Wemmer, 1990). It was shown by Pelton and coworkers via NMR that even below 1:1 Dst:PA ratios, both the 1:1 and 2:1 complexes were observed suggesting that Dst binds cooperatively. Favorability of the 2:1 was sequence dependent given that a sequence of  $A_3T_2$  favored 1:1 over 2:1, but  $A_3T_3$  favored 2:1 over 1:1. Wemmer also determined that 2:1 binding was preferred over 1:1 for mixed AT sequences compared to A-tract (Wemmer et al., 1994).

Using SPR, Wilson's group has determined sequence-dependent cooperativity factors for the 2:1 Dst-DNA complex using the microscopic binding constants of the different binding modes (Liu, Y. & Wilson, 2009; Peixoto et al., 2008). The 2:1 binding mode of Dst was determined to be highly favorable for

ATATA and ATATAT sequences, but not for ATAT and AAAA which they attribute to a difference in curvature in the DNA which in turn results in different Dst-DNA interactions.

While Dst-DNA cooperativity has been studied, cooperative binding has not been studied for linked PA. This is surprising given the previously discussed stoichiometry results in **section 1.7.1**.

## **1.8. Overview of this dissertation**

This thesis reports the findings of a study on PA-DNA binding kinetics and stoichiometry using small hairpin oligonucleotides. Fluorescence and CD spectroscopy were the methods used to investigate these PA-DNA binding interactions. The results are reported in three parts: **Chapter 3** presents the results of our study of PA-DNA A/T pattern dependence; **Chapter 4** examines size-dependent PA-DNA binding; **Chapter 5** discusses our study on the effects of N-terminal functional groups (TMG and Guan) on the DNA binding kinetics of 2 large antiviral PAs (NV1028 and NV1042).

## CHAPTER II

### MATERIALS AND METHODS

#### 2.1. Polyamide preparation and storage

Hairpin polyamides (PA) were prepared by solid-phase synthesis or by a combination of solid-phase and solution methods, and purified by HPLC, as described in (Edwards et al., 2011). Verification of purity was confirmed by UV/VIS and/or MALDI MS. PAs were stored as dried solids at 4 °C.

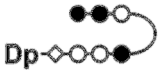
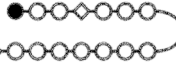
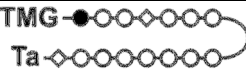
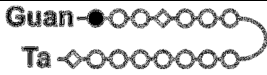
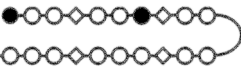


PAs used in this study were reconstituted by dissolving a speck of dried PA in 50-100  $\mu$ L 100% DMSO and quantitated via UV/VIS using the extinction coefficients ( $\epsilon_{305}$ ) in **Table 2.1**. Stocks were prepared weekly and stored under light-sensitive conditions at room temperature on the benchtop. PA stocks were kept below 100  $\mu$ M to minimize PA aggregation.

#### 2.2. DNA hairpins

##### 2.2.1. DNA hairpin design

DNA hairpins (Integrated DNA Technologies, Skokie, IL) were designed based on the pairing rules discussed in **Chapter 1**. PAs can bind in a forward or reverse orientation where the N- and C-termini align 5' to 3' or 3' to 5', respectively, (**Fig. 2.1**) based on most hairpin PA literature (Herman et al., 1998; Vasilieva et al., 2016; White, S. et al., 1997b). All hairpins were designed with the assumption that PAs prefer to bind in the forward orientation since the reverse orientation is believed to be less favorable (Herman et al., 1998).

**Table 2.1** Polyamides characterized in this dissertation

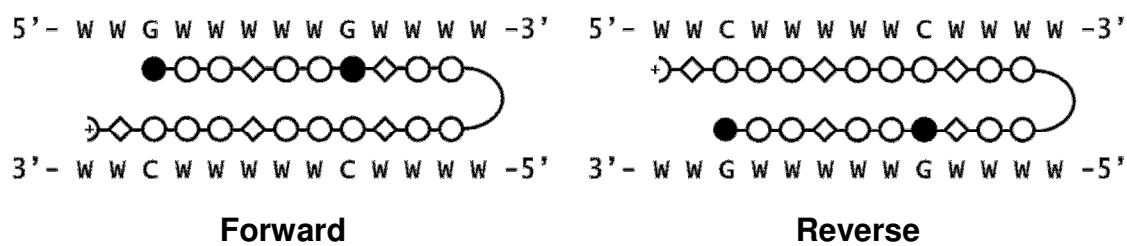
PA	Structure	Size (Rings)	HPV16 IC <sub>50</sub> <sup>a</sup> (μM)	HPV16 IC <sub>90</sub> <sup>a</sup> (μM)	ε <sub>305</sub> (M <sup>-1</sup> cm <sup>-1</sup> )
KA1039		6	NOT ACTIVE		80,500 <sup>b</sup>
NV1028		14	0.100±0.02 (4)	1.113	92,600 <sup>c</sup>
NV1113			0.304	>10	
NV1115			0.103	0.378	
NV1042		20	0.036±0.0004 (3)	0.351	91,700 <sup>c</sup>
NV1114			0.035	0.411	
NV1116			0.038	0.340	

Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop, Guanadinium (Guan), tetramethylguanidinium (TMG), 3-(dimethylamino)propylamine (Dp), 3,3'-diamino-N-methyldipropylamine (Ta).

<sup>a</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011). IC<sub>50</sub> and IC<sub>90</sub> are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

<sup>b</sup> not published. Determined using the method described in (Dupureur et al., 2012)

<sup>c</sup> (Vasilieva et al., 2016)



**Figure 2.1** Binding orientations of hairpin polyamides in the minor groove of DNA. NV1042 shown.



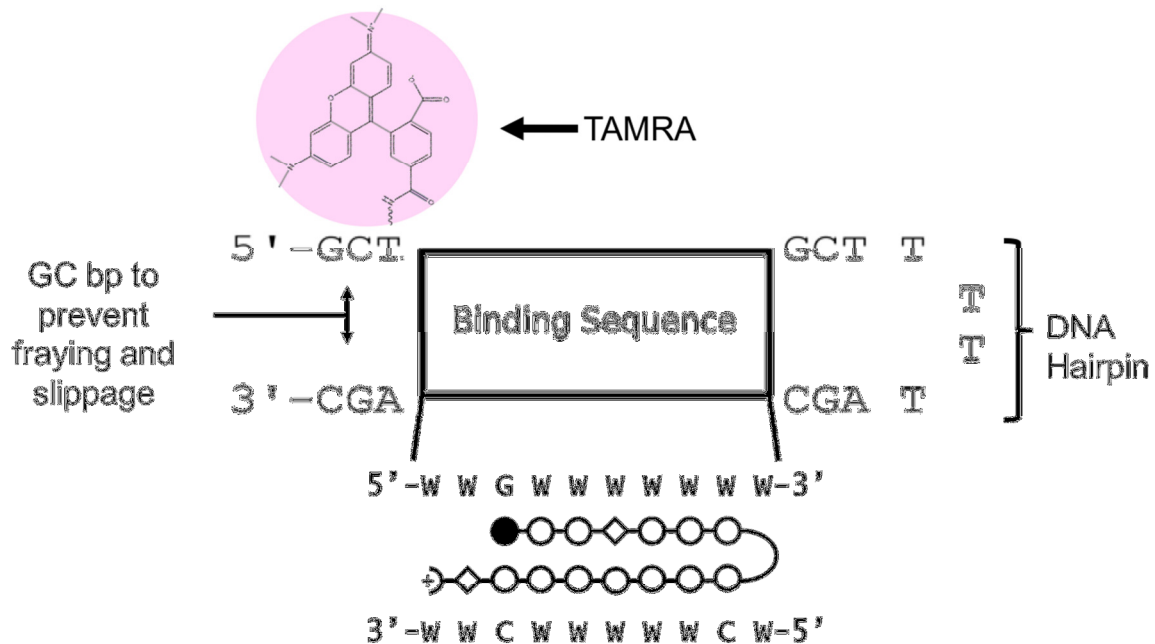
Perfect match binding sequences were isolated in a hairpin and flanked with 2-5 base pairs (bp) at the 5' and 3' end of the sequence. At least two GC or CG pairs were included in the flanks to prevent fraying and to discourage alternate binding positions. See **Fig 2.2** for DNA hairpin design details.

DNA hairpins were labeled with an extrahelical fluorophore, 6-carboxytetramethylrhodamine (TAMRA), via conjugation to an internal thymine by NHS ester linkage (**Fig. 2.2**). TAMRA was placed at either the PA N- or C-termini, or at the  $\gamma$  turn binding positions on the DNA. When possible, TAMRA was preferentially placed at the N-terminus binding position on the DNA since DNA hairpin synthesis proceeds 5' to 3'; thus, ensuring higher yields.

Unlabeled hairpins with identical sequences were used as competitors for dissociation kinetics experiments and for competition experiments to examine dye-related interference. Additionally, unlabeled hairpins and 2 unlabeled duplexes were used by Kristin Bales for circular dichroism spectroscopy. A complete list of hairpins and duplexes used in this dissertation can be found in **Table 2.2**.

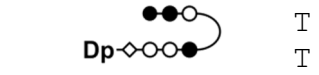
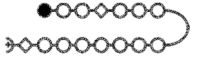
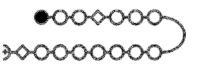
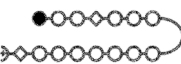
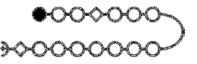
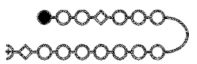
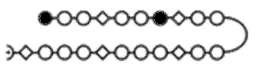
### **2.2.2. DNA preparation and storage**

DNA were shipped as dried solids and were dissolved in 2 mL of Milli-Q. DNA solutions were washed and concentrated to 200  $\mu$ L twice using a Centricon centrifugal filter with an Ultracel YM-3 filter (ThermoFisher, Waltham, MA). DNA were centrifuged at 1,500 rpm using a Sorvall ST 16 centrifuge with a TX-400 round bucket rotor.



**Figure 2.2** DNA hairpin overall design. The example shows the binding site and cartoon structure of NV1028, where W stands for A or T.

**Table 2.2** DNA hairpins discussed in this dissertation

DNA	Lab Designation	Sequence
ODN-6 ODN-6-T3	YS06125 YS04109	5' - CCT <b>TGGCT</b> TCT  3' - GGAACCGAAGT
ODN-14-SFAT ODN-14-SFAT T3 ODN-14-SFAT T23	YS07107 JN02075 /CD26043 JN02108	5' - GC <b>TAGATATATA</b> GCTT  3' - CGATCTATATATCGAT
ODN-14-SFTT ODN-14-SFTT T3 ODN-14-SFTT T23	YS07143 YS08039 YS07142	5' - GC <b>TGTTTTTTA</b> GCTT  3' - CGAACAAAAATCGAT
ODN-14-LFTT ODN-14-LFTT T6	JN07075 JN02076-13	5' - TTCCA <b>TGTTTTTTA</b> CACTGT  3' - AAGGTAACAAAAATGTGACT
ODN-14-SFAT Duplex ODN-14-SFAT T3 Duplex	JN06140/ JN06141 JN06139/ JN0141	5' - GC <b>TAGATATATA</b> GCT-3'  3' - CGATCTATATATCGA-5'
ODN-14-GH Duplex	JN07097A/ JN07097B	5' - GCG <b>AAGTTAATAT</b> GCG-3'  3' - CGCTTCAATTATACGC-5'
ODN-20 ODN-20-T3 ODN-20-T16	YS07108 CD26072 JN04058	5' - GC <b>TATGTTTAAGATA</b> GCT  3' - CGATACAAATTCTATACGA

Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop. Expected binding sequences are **bold and underlined**. Colors denote dye position: primary dye positions are labeled in **red**, alternate dye positions are shown in **green**. **Black** ODNs are unlabeled and used as competitors or for CD spectroscopy.

To anneal, the washed DNA was transferred to a 1.5 mL microcentrifuge tube and placed in 100 °C Milli-Q and allowed to cool for at least 4 hours (typically overnight). Annealed DNA was quantitated using the  $\epsilon_{260}$  provided by IDT. DNA stocks were then aliquoted and excess DNA was lyophilized for long term storage.

DNA stocks below 5  $\mu$ M were kept for no longer than 1 week and stored at 4 °C. DNA stocks below 1  $\mu$ M were made fresh daily.

## **2.3. Buffer preparation**

### **2.3.1. Recipes**

All experiments were performed in 10 mM HEPES, 50 mM NaCl, 1 mM EDTA at pH 7.4 (1X HNE). 2X HNE buffer was prepared in bulk and stored at room temperature. Experiments were performed using 1 mM CHAPS to prevent aggregation of larger hairpin PA in aqueous conditions. 1 mM CHAPS was prepared from a bulk 10 mM CHAPS, 1X HNE solution prepared weekly and stored at room temperature. Both stocks were filtered via CENTREX 0.45  $\mu$ m (ThermoFisher, Waltham, MA) sterile centrifuge filters prior to use to remove particulates.

### **2.3.2. Degassing**

Solutions used for stopped-flow experiments were degassed to prevent the formation of bubbles by dissolved gasses. Milli-Q and 1X HNE were chilled to 4 °C. Chilled solutions were then transferred to a side-arm flask and stirred with a

magnetic stir bar under vacuum. Ice packs were used to prevent warming of the solutions while degassing. 1 mM CHAPS was prepared using degassed 1X HNE and then degassed after CHAPS was dissolved.

Degassed solutions were filtered via a glass filtration system using hydrophilic PVFD membrane filter with 0.45  $\mu\text{m}$  pore size (Millipore Sigma, St. Louis, MO) and then transferred to a plastic bottle for storage at 4 °C. If solutions were not used within 24 hours, then degassing was repeated.

## **2.4. Fluorescence assays (equilibrium)**

### **2.4.1. Direct titration to determine PA-DNA binding affinities ( $K_d$ )**

The fluorescence assay used to determine binding affinities ( $K_d$ ) of PA to DNA was developed by Dr. Dupureur specifically for characterizing PA-DNA binding interactions (Dupureur et al., 2012).  $K_d$ s were determined directly by monitoring the change in fluorescence intensity of the TAMRA dye upon binding with PA using a Fluorolog 3 spectrofluorometer (Horiba). TAMRA was excited at 559 nm. Emission was detected at 580 nm via S-channel, or with a 592 nm bandpass filter (Edmund Optics) via T-channel, depending on hairpin concentration. All experiments were performed using a 3 mL quartz cuvette with a 1 cm path length and stirred continuously with a magnetic stir bar. Cuvettes and magnetic stir bars were washed and stored in a 20%  $\text{HNO}_3$  solution. Experiments were performed at 25°C and temperature was maintained by a circulating water bath.

When PA binds to the DNA it causes an unspecified change in the DNA and/or the environment of TAMRA which results in a change of fluorescence intensity (**Fig. 2.3**). The direction of the intensity change remained consistent for each PA. To determine  $K_d$ 's, PA was titrated into TAMRA-labeled hairpin held at concentrations 3-10-fold below the  $K_d$ . The reaction was mixed using a magnetic stir bar in the cuvette and allowed to equilibrate before fluorescence intensity was recorded at each concentration of PA. Equilibration times were determined for each PA and DNA pair but equilibration typically occurred in around 5-7 minutes.

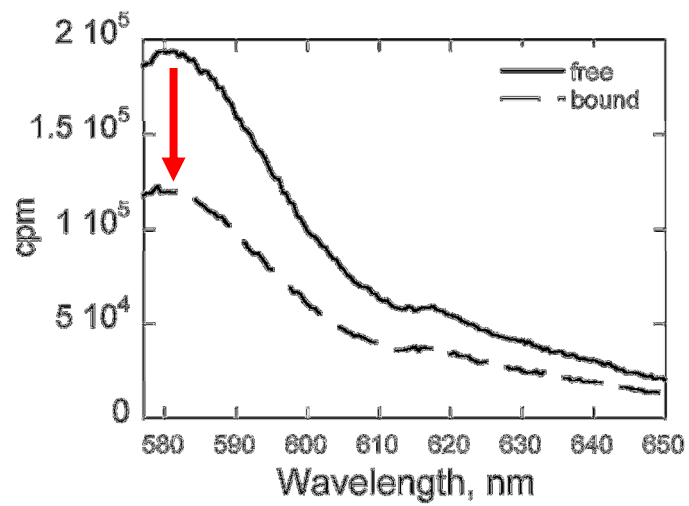
Intensities were averaged and normalized to 1 and fitted in *Kaleidagraph* (Synergy Software, Reading, PA) using the Langmuir (**eqn. 2.1**) and Hill (**eqn. 2.2**) equations, where  $\theta$  is the fraction of bound ligand,  $[L]$  is the concentration of free ligand, and  $n$  is the Hill coefficient. See **Fig 2.4A** for an example of fitted data.

$$\theta = \frac{K_a[L]}{1 + K_a[L]} \quad (2.1)$$

$$\theta = \frac{[L]^n}{K_d + [L]^n} \quad (2.2)$$

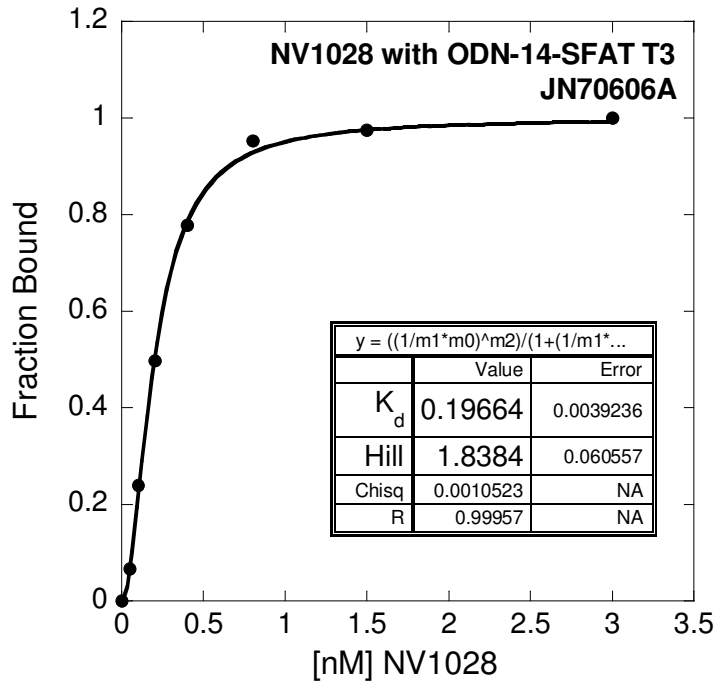
When DNA concentrations could not be at least 3-5-fold lower than the  $K_d$ , the  $K_d$  was determined using *Scientist* (Micromath, St. Louis, MO) where the model used to calculate the  $K_d$  does not rely on the assumptions required for the Hill and Langmuir equations (Voet et al., 2013).

For KA1039, NV1028, and NV1042, binding affinities were determined in triplicate and averaged. Relative errors are shown. Stocks of TMG and Guan

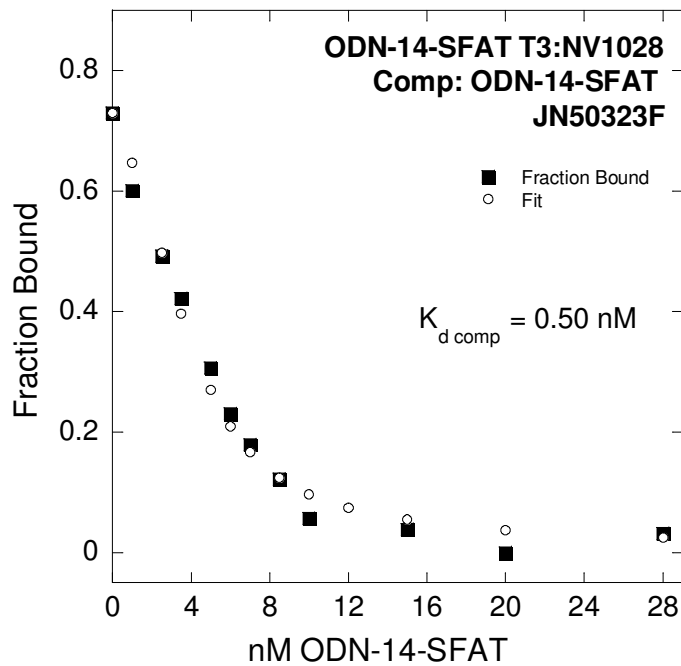


**Figure 2.3** Fluorescence spectra of free versus DNA bound to PA. The red arrow indicates the direction of the intensity change. Adapted with permission from (Dupureur et al., 2012) Copyright 2012 Elsevier. License number 4315740775723.

A



B



**Figure 2.4** Example of fitted equilibrium data. A, directly determined  $K_d$  of NV1028 and ODN-14-SFAT T3. 0.1 nM ODN-14-SFAT T3, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. B, competition equilibrium of NV1028 with ODN-14-SFAT T3. 5 nM ODN-14-SFAT T3:NV1028, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 10 mM CHAPS, pH 7.4, 25 °C



analogs were limited so binding affinities were typically collected only once. Fit errors are shown for single replicates. Each figure and table is annotated with the number of replicates and error type for the results shown.

#### **2.4.2. Indirect determination of PA-DNA $K_d$ s via competition**

Indirect determination of  $K_d$  via competition experiments was used to examine dye-related interference. To do this, a 1:1 complex of PA and TAMRA-DNA was formed at concentrations 8-10-fold higher than the directly determined  $K_d$ . Aliquots of unlabeled DNA competitor were titrated into the 1:1 mixture and allowed to equilibrate, typically around 7-10 min. Fluorescence intensities were collected after each addition of unlabeled competitor.

Collected intensities were normalized and fit using *Scientist* to determine the binding affinities of the labeled hairpin and unlabeled hairpin (**Fig 2.4B**) (Dupureur et al., 2012). If the binding affinities were the same for both labeled and unlabeled hairpins, then no dye interference was present. If binding affinities were different, then a new DNA hairpin was designed with the dye moved to a different location.

#### **2.4.3. PA-DNA binding stoichiometry**

##### **2.4.3.1. PA-DNA binding stoichiometry via fluorescence spectroscopy**

To determine stoichiometric ratios of PA-DNA binding, a titration was performed under conditions at which the concentration of labeled DNA was 5-10-fold higher than the directly determined  $K_d$ . Intensities were collected after each

aliquot of PA was added until saturation of the DNA was observed.

Stoichiometric ratios were determined at 25 °C for all PAs. See **Fig 2.5** for an example of a stoichiometry plot.

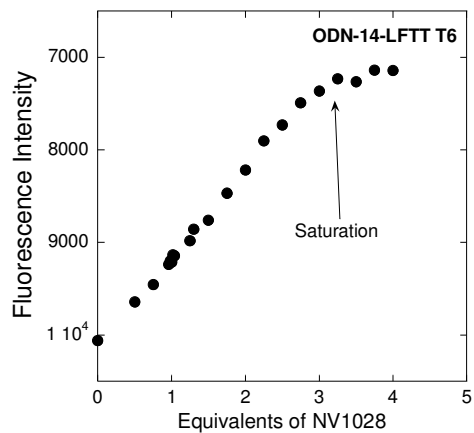
#### **2.4.3.2. PA-DNA binding stoichiometry via circular dichroism spectroscopy**

PA-DNA binding stoichiometry was also determined via circular dichroism spectroscopy by Kristin Bales using a Jasco J-1500. Equivalents of PA were titrated into 5  $\mu$ M of unlabeled DNA in 1X HNE, 1 mM CHAPS and thermostatted at 20 °C. All experiments were performed in a 1 mL quartz cuvette with 1 cm path length. The reaction mixture was pipet mixed after each aliquot of PA was added. Spectra were collected using the parameters shown in **Table 2.3**. Spectra were collected from 225-400 nm (**Fig. 2.6A**). Hairpin PA binding in the minor groove of DNA is observed around 300-370 nm with a maximum around 330 nm (Lyng et al., 1992). CD mdeg at around 330 nm were collected and plotted with respect to PA equivalents to determine stoichiometric ratios (**Fig. 2.6B**).

### **2.5. Fluorescence assays (PA-DNA binding kinetics)**

#### **2.5.1. PA-DNA association kinetics via steady state instrument**

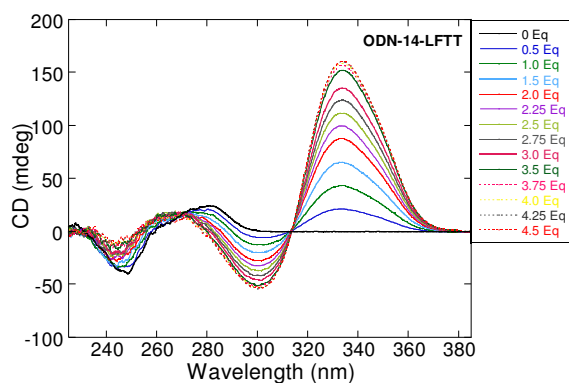
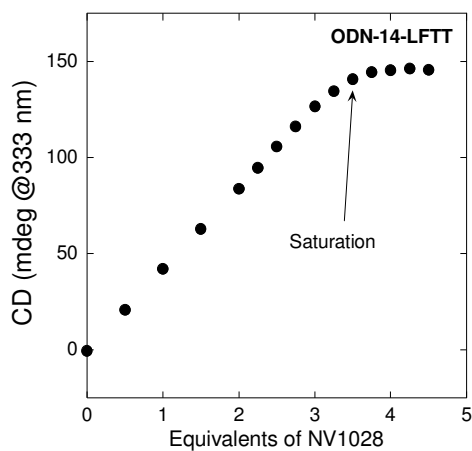
PA-DNA association rate constants ( $k_{on}$ ) were determined by performing a series of reactions at a fixed concentration of labeled DNA while varying PA concentrations. Changes in intensities were monitored in real-time. Collected intensities were then normalized and fitted to **eqn. 2.3**, where  $k_{obs}$  is the observable rate constant and  $t$  is time (**Fig. 2.7A**). When  $k_{obs}$  is plotted as a



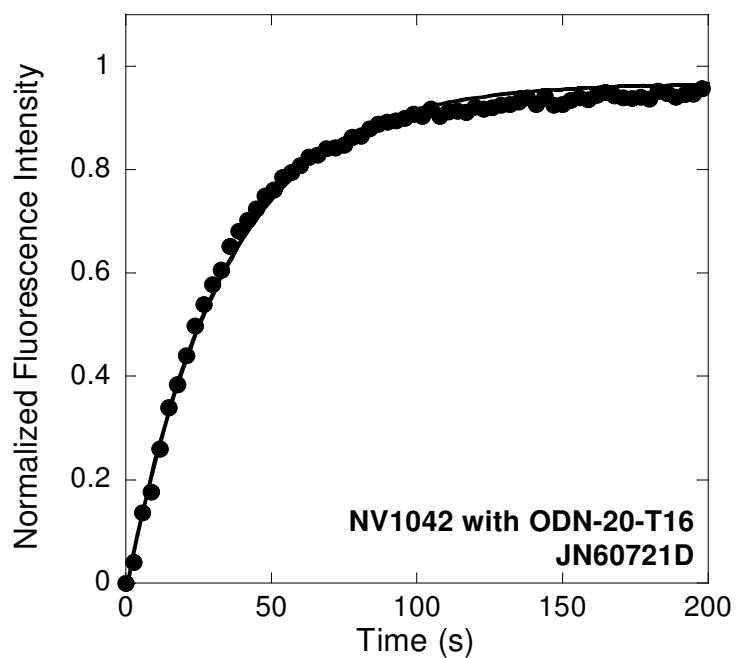
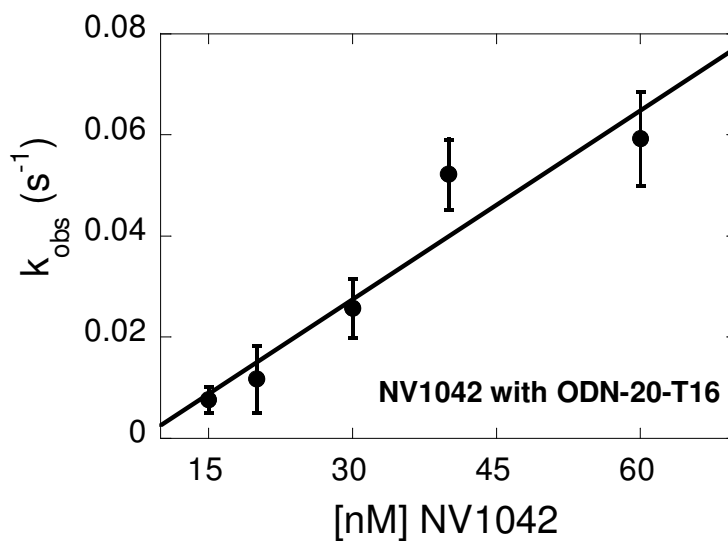
**Figure 2.5** Stoichiometry data of NV1028 with ODN-14-LFTT T6 collected via fluorescence. The black arrow indicates the saturation point. 5 nM ODN-14-LFTT, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. See **Chapter 3** for discussion of data.

**Table 2.3** Settings used for CD experiments

<b>Parameter</b>	<b>Setting</b>
Wavelength Range	225 nm – 400 nm
Scan Speed	50 nm/min
Data Pitch	0.2 nm
Slit Width	200 $\mu$ m
Accumulations	5
Cell Type	Quartz
Cell Volume	1 mL
Path length	1 cm

**A****B**

**Figure 2.6** Stoichiometry data collected by Kristin Bales via circular dichroism for KA1039 with ODN-6. A, overlay of spectra collected; B, stoichiometry plot using intensities collected at 326.6 nm. The black arrow indicates the saturation point. 5  $\mu$ M ODN-14-LFTT, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 20  $^{\circ}$ C. See **Chapter 3** for discussion of data.

**A****B**

**Figure 2.7** Fitted association kinetics data for NV1042 with ODN-20 T16. A, experiment to determine  $k_{\text{obs}}$ ; B, secondary plot to determine  $k_{\text{on}}$  and  $k_{\text{off}}$ . 2 nM ODN-20-T16, 30 nM NV1042 (A), 10 mM HEPES, 50 nM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C.

function of PA concentration it can be fit linearly to **eqn. 2.4**, where the slope is  $k_{on}$  and the y-intercept is the dissociation rate constant ( $k_{off}$ ) (**Fig. 2.7B**) (Conlan & Dupureur, 2002). See **section 3.3.3.2** for fits to alternate models.

$$y = 1 - e^{-k_{obs}t} \quad (2.3)$$

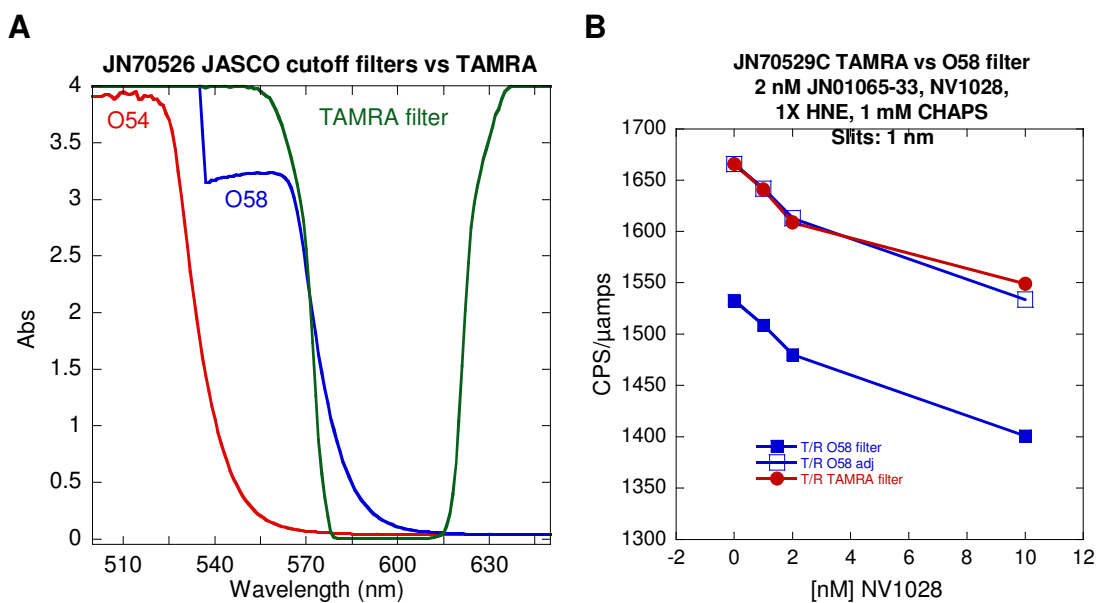
$$k_{obs} = k_{on}[PA] + k_{off} \quad (2.4)$$

PA-DNA association kinetic traces were collected at least once. Where only a single replicate was obtained, the error bars shown correspond to the fit errors of the individual association curves collected for each concentration of PA. Where duplicates were obtained, the error bars shown correspond to relative errors determined for each concentration of PA.

If curves were only determined once, the errors shown are fit errors of the linear fits. If curves were determined in duplicate, the slopes were determined for each set of curves and then the slopes were averaged and the error shown is the relative error of the slopes. Each figure and table is annotated with the number of replicates and error type for the results shown.

### **2.5.2. PA-DNA association kinetics via stopped-flow fluorescence**

Stopped-flow kinetics experiments were performed using the Jasco J-1500 configured for fluorescence spectroscopy with a filter. Cutoff filter O-58 was purchased from Jasco and provided similar sensitivity to the TAMRA bandpass filter used on the steady state instrument (**Fig. 2.8**).



**Figure 2.8** Comparison of Jasco cutoff filters versus TAMRA bandpass filter. A, absorption spectra of cutoff filters O54 and O58 absorbance spectra versus TAMRA bandpass filter; B, fluorescence spectroscopy experiment to determine sensitivity and possible loss of signal when using O58 cutoff filter vs the TAMRA filter with a quadruple mismatch DNA hairpin and NV1028.



The Jasco SFS-562T 2-syringe stopped-flow attachment was used with 10 mL syringes. The stopped-flow cell used had a 0.5 cm path length and a total volume of 57.6  $\mu\text{L}$ . Mixer volume was 8.8  $\mu\text{L}$ . Flow rates were calculated by the instrument. Dead time was calculated manually using **eqn. 2.5**. The system and cuvettes were washed with water and stored in 20% ethanol.

$$Dead\ time\ (ms) = \frac{Cell\ Volume\ (\mu L) + Mixer\ Volume(\mu L)}{total\ flow\ rate\ \left(\frac{mL}{s}\right)} \quad (2.5)$$

KA1039 stocks were prepared daily in degassed 1X HNE. DNA stocks were prepared in degassed 1X HNE and degassed 2 mM CHAPS. For experiments, DNA and KA1039 solutions were mixed 1:1 so that experiments were performed in 1X HNE, 1 mM CHAPS, like those performed on the steady-state instrument.

Experiments were performed at 5 °C. The cell was cooled by a water bath and solutions and syringes were kept cool using ice packs. Solutions were pre-chilled prior to making stocks.

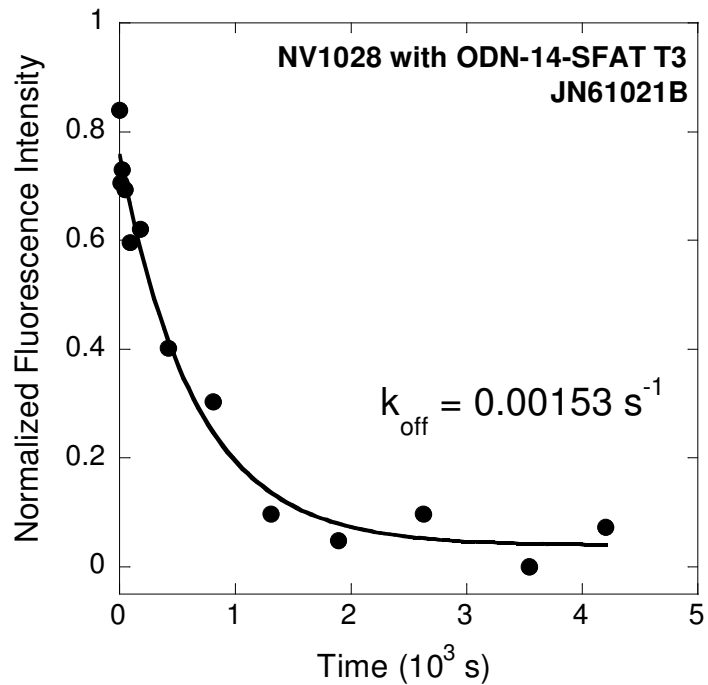
### **2.5.3. PA-DNA dissociation kinetics**

To directly determine PA-DNA dissociation rate constants ( $k_{\text{off}}$ ), an excess of unlabeled DNA was added to a pre-formed 1:1 PA-DNA. Complex concentrations were at least 10-fold above the directly determined  $K_d$ . Single point intensities were collected both before and after PA was added to ensure

that PA bound and that the system was at equilibrium. Recovery of fluorescent intensity (i.e. a change in intensity opposite that observed when PA is added) upon the addition of unlabeled competitor was monitored in real-time. Averaged intensities were normalized and fit using **eqn 2.6**. See **Fig. 2.9** for an example of fitted data.

$$y = e^{-k_{off}t} \quad (2.6)$$

PA-DNA dissociation kinetic traces were collected at least once. Where only a single replicate was obtained, the error shown correspond to the fit errors of the curves. Where duplicates were obtained, the  $k_{off}$  values were averaged and the errors presented are fit errors. Each figure and table is annotated with the number of replicates and error type for the results shown.



**Figure 2.9** Dissociation kinetics data for NV1028 with ODN-14-SFAT T3. 5 nM ODN-14-SFAT T3:NV1028, 100 nM ODN-14, 10 mM HEPES, 50 nM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C.

## CHAPTER III

### DNA BINDING KINETICS OF A LARGE ANTIVIRAL POLYAMIDE

#### 3.1. Introduction and background

##### 3.1.1. History and motivation

NV1028 is a 14-ring hairpin polyamide (PA) designed to bind to the E1 and E2 binding sites of HPV16. Its binding site is WWGWWWWWWW, where W stands for A or T (**Fig. 1.5**). When compared to Dst, NV1028 has been shown to be effective against 3 of the most problematic strains of HPV—HPV16, HPV18, HPV31—with no observed cytotoxicity in cell culture (Edwards et al., 2011). It has been shown that HPV DNA concentrations in cells are decreased relative to total DDNA and housekeeping gene DNA via activation of the DNA damage response system (DDR); however, it is unclear how activation occurs (Edwards et al., 2013a; Edwards et al., 2013b).

Of the large antiviral hairpin polyamides, NV1028 is well-characterized. Previous studies of the interactions of NV1028 with large duplex fragments by DNase I footprinting and affinity cleavage revealed binding was not limited to the binding sequence for which it was designed (He et al., 2014; Vasilieva et al., 2016). Binding sequences for the footprints and cleavage patterns observed were assigned by giving preference to sites with the fewest number of mismatches in a forward binding orientation. However, there are sites within the large duplex fragments where the best possible match corresponds to a

sequence of 3 or more mismatches in either binding orientation. Additionally, the binding affinities determined for NV1028 to these mismatched sites were very tight (low nM).

As discussed in **Chapter 1**, drug binding kinetics can often be better than binding affinity as an indicator of drug efficacy (Copeland, 2016; Copeland et al., 2006; Tummino & Copeland, 2008). Previous studies have characterized the differences in DNA binding kinetics with respect to perfect versus mismatched sites of 6, 8, and 10-ring hairpin PAs (Baliga et al., 2001; Henry et al., 2004). Other related studies have focused on characterizing the binding kinetics of hairpin PAs with modified building blocks designed to discriminate between A and T (Zhang et al., 2012; Zhang et al., 2006). This area is discussed in detail in **Chapter 1**. However, there have been no studies characterizing the binding kinetics of NV1028 with DNA.

It has been established that minor groove widths vary with respect to A/T pattern (Burkhoff & Tullius, 1988). However, despite it being shown that Dst binds to mixed AT and A-tract sequences of DNA via different binding modes (Chen & Sha, 1998; Fagan & Wemmer, 1992; Wang, S. et al., 2011), NV1028 is designed to bind degenerately to mixed AT and A-tract sequences. Yang Song has studied the thermodynamics of NV1028 binding to mixed AT and TT sequences and determined that while NV1028 binding is entropically driven for both sequences, enthalpic contributions differ: enthalpic contributions are favorable for AT sequences, but unfavorable for TT. See Yang Song's thesis. The differences in enthalpy could be related to NV1028 binding in a different

orientation, or via multiple binding modes and if, so, it stands to reason that the DNA binding kinetics of NV1028 may also be different for AT versus TT.

Here I report on a study of the DNA binding kinetics of NV1028. Using the fluorescence assay developed by our lab and described in **Chapter 2**, I determined kinetic rate constants and DNA binding stoichiometric ratios for NV1028 with DNA oligomers which contained varied A/T patterns in the recognition sequence.

### **3.1.2. Design of DNA hairpins used for this study**

A detailed discussion of DNA oligomer design can be found in **Chapter 2**. Two perfect match sequences were selected for this study (**Table 3.1**). The first binding sequence was a designed sequence with repeating AT base steps and short flanks (2 bp at 5' and 3 bp at 3') on either side of the binding site (ODN-14-SFAT). ODN-14-SFAT was labeled with TAMRA, a rhodamine-based fluorescent dye, at either T3 (ODN-14-SFAT T3) or T23 (ODN-14-SFAT T23). Additionally, a duplex was designed with this sequence to test the influence of the hairpin loop and was labeled at T3 (ODN-14-SFAT T3 duplex).

The second binding sequence is the perfect match sequence located at 7600-7609 in the HPV16 genome. This sequence was previously characterized by DNase I footprinting and affinity cleavage and is also referred to as NV1028 AC site #13 in previous publications (Vasilieva, 2014; Vasilieva et al., 2016). Short flanks identical to SFAT were used for this hairpin, ODN-14-SFTT. This oligonucleotide was labeled at T3 or T23 (ODN-14-SFTT T3 and ODN-14-SFTT

**Table 3.1** NV1028 and the DNA hairpins relevant to **Chapter 3**<sup>a</sup>

DNA	Lab Designation	Sequence
<b>ODN-14-SFAT</b> <b>ODN-14-SFAT T3</b> <b>ODN-14-SFAT T23</b>	<b>YS07107</b> <b>JN02075/CD26043</b> <b>JN02108</b>	5' -GC <b><u>TAGATATATA</u></b> AGCTT 3' -CGATCTATATAT <b>TCGAT</b> T T
<b>ODN-14-SFTT</b> <b>ODN-14-SFTT T3</b> <b>ODN-14-SFTT T23</b>	<b>YS07143</b> <b>YS08039</b> <b>YS07142</b>	5' -GC <b><u>TTGTTTTTTA</u></b> AGCTT 3' -CGAACAAAAAAT <b>TCGAT</b> T T
<b>ODN-14-LFTT</b> <b>ODN-14-LFTT T6</b>	<b>JN07075</b> <b>JN02076-13</b>	5' -TTCCA <b><u>TTGTTTTTTA</u></b> CACTGT 3' -AAGGTAACAAAAAATGTGACT T T
<b>ODN-14D-SFAT Duplex</b> <b>ODN-14-SFAT T3 Duplex</b>	<b>JN06140/</b> <b>JN06141</b> <b>JN06139/</b> <b>JN0141</b>	5' -GC <b><u>TAGATATATA</u></b> AGCT-3' 3' -CGATCTATATATCGA-5'

<sup>a</sup> Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop. Expected binding sequences are **bold and underlined**. Colors denote dye position: primary dye positions are labeled in **red**, alternate dye positions are shown in **green**. **Black** ODNs are unlabeled and used as competitors or for CD spectroscopy.

T23, respectively). Another hairpin, ODN-14-LFTT, contains an identical binding site to ODN-14-SFTT, but with expanded flanking sequences. This oligonucleotide contains 5bp of the natural sequence context on either end of the binding site. For fluorescence spectroscopy, ODN-14-LFTT was labeled at T6 (ODN-14-LFTT T6).

## 3.2. Results

### 3.2.1. AT versus TT-dependent NV1028-DNA binding affinities

To determine DNA binding affinity ( $K_d$ ), the change in fluorescence was monitored as PA was titrated into DNA using the method and fitting protocol described in **Chapter 2**.  $K_d$ s for NV1028-DNA were determined for ODN-14-SFAT T3, ODN-14-SFAT T23, ODN-14-LFTT T6, and ODN-14-SFTT T3 and ranged from 0.16 nM to 0.26 nM with the average  $K_d$  of  $0.20 \pm 0.05$  nM (**Table 3.2**). A/T pattern and flank length did not have an observable effect on the NV1028  $K_d$ s. These  $K_d$ s are 2-10-fold tighter than the  $K_d$ s of perfect match sites determined using a large fragment duplex via DNase I footprinting, 0.7-1.9 nM (Vasilieva et al., 2016). This is likely due to the difference in DNA substrates used in these assays. Given that NV1028 is capable of binding to mismatched sequences, it is likely that the large fragment duplex used in DNase I footprinting acts as a bulk DNA competitor resulting in the weaker observed  $K_d$ s for NV1028.

$K_d$  characterization was not possible for ODN-14-SFTT T23 nor ODN-14-SFAT T3 duplex due to poor dye behavior (isotherms resembled a parabola, not a rectangular hyperbola) for ODN-14-SFTT T23 and incompatible duplex



**Table 3.2** DNA binding affinities for NV1028

<b>DNA</b>	<b>Direct <math>K_d^a</math> (nM)</b>	<b>Competition <math>K_d^b</math> (nM)</b>
ODN-14-LFTT T6	0.20±0.06	0.03±0.01
ODN-14-SFAT T3	0.26±0.16	0.9±0.5
ODN-14-SFAT T23	0.16±0.02	0.35±0.08
ODN-14-SFTT T3	0.31±0.01 <sup>c</sup>	0.45 <sup>c</sup>

<sup>a</sup> 0.1 nM labeled DNA, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 25 °C. This work. Unless otherwise specified,  $K_d$  was determined in duplicate and relative error is shown.

<sup>b</sup> collected by Yang Song, 1:1 NV1028-labeled DNA complex (1-2 nM), 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 25 °C.

<sup>c</sup> single replicate at this writing

thermodynamics for ODN-14-SFAT T3 duplex.

### **3.2.2. DNA binding affinities determined indirectly via competition assays**

To determine if the extrahelical dye, TAMRA, has any effect on PA binding, DNA binding affinities were also determined indirectly by Yang Song via a competition assay described in **Chapter 2 (Table 3.2)**. In general,  $K_d$ s determined indirectly were slightly weaker, but still in the low nM-high pM range with the exception of NV1028 with ODN-14-LFTT T6, which exhibits a much tighter  $K_d$  than that for labeled DNA. Weaker  $K_d$ s may be the result of sensitivity issues experienced while trying to collect competition data. When NV1028 binds to DNA, fluorescence intensity decreases; when competitor is added, fluorescence intensity increases (recovers). In NV1028 competition experiments, not all of the fluorescence intensity lost was recovered (see **3.2.7**). When directly observing NV1028 binding to DNA, the changes in intensity were large and easily observed; when indirectly observing NV1028 binding to DNA, the changes in intensity were smaller and, thus, data tended to be noisier.

Regarding the unusually low indirect  $K_d$  for ODN-14-LFTT (**Table 3.2**), 25 °C is the only temperature at which  $K_d$ s determined directly and via competition were significantly different. See Yang Song's thesis.

### **3.2.3. Circular dichroism spectroscopy**

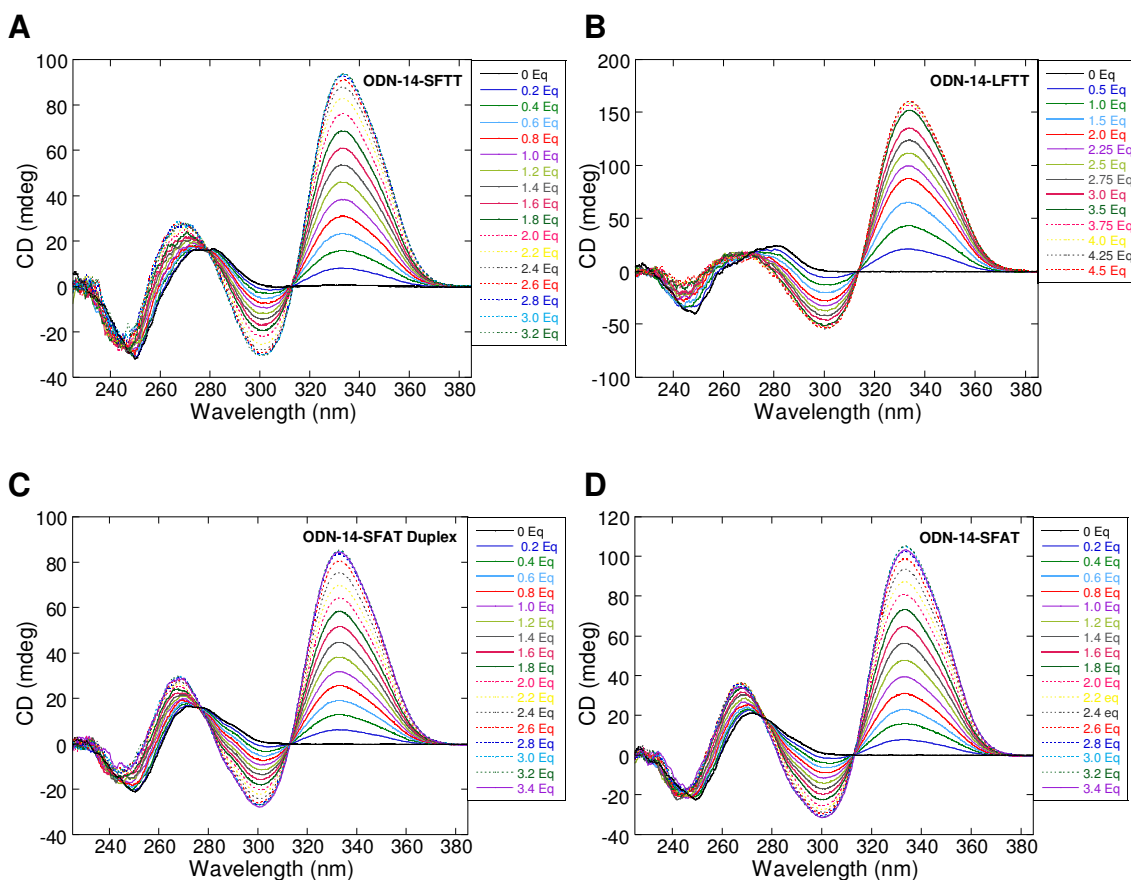
Circular dichroism (CD) spectroscopy can be used to probe DNA conformation in response to PA binding (Lyng et al., 1992). Using the method

described in **Chapter 2**, CD spectra were collected in duplicate by Kristin Bales for ODN-14-SFAT, ODN-14-LFTT, ODN-14-SFTT, and ODN-14-SFAT duplex at increasing equivalents of PA while DNA remained fixed at 5  $\mu$ M (**Fig. 3.1**). A strong response was observed at 333 nm, as expected for hairpin PAs in the minor groove of DNA (Lacy et al., 2002; Lyng et al., 1992) (**Fig. 3.2**).

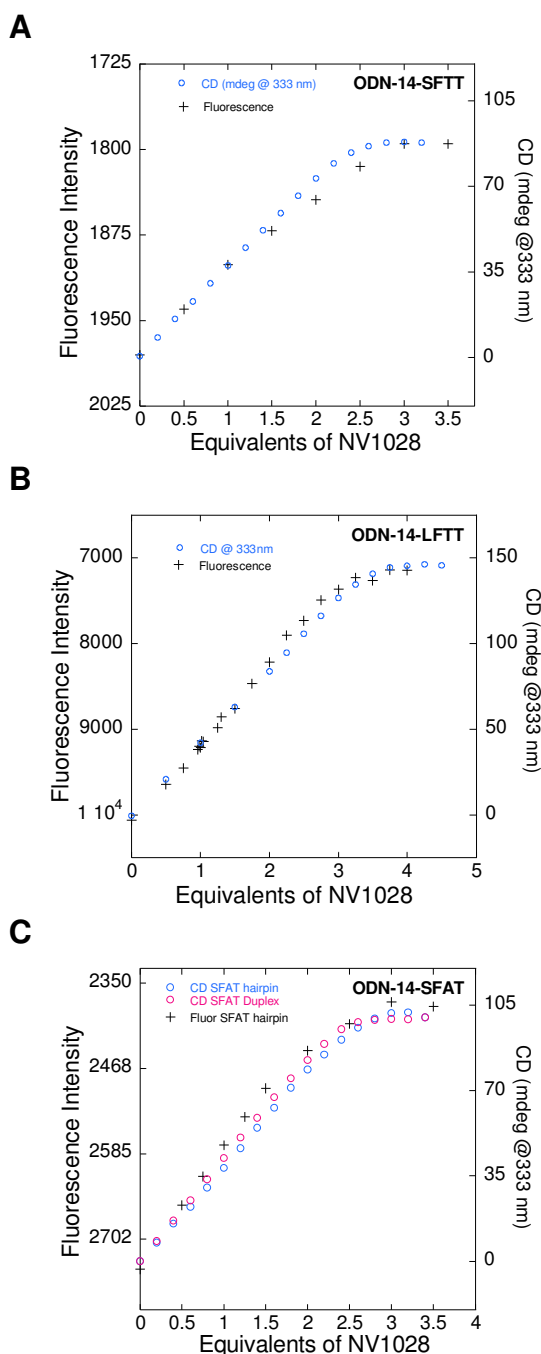
Unlike the conditions used for the fluorescence assay, conditions for CD spectroscopy were thermodynamically favorable for a stable ODN-14-SFAT duplex. CD analysis was primarily performed on ODN-14-SFAT duplex to determine the impact of the hairpin loop. CD spectra are indistinguishable from that of the corresponding DNA hairpin, indicating the hairpin loop had no significant impact on the DNA conformational response (**Fig. 3.1CD**).

#### **3.2.4. Comparing PA-DNA binding stoichiometries obtained via fluorescence and CD spectroscopy**

NV1028-DNA binding stoichiometries were determined via both CD and fluorescence spectroscopy for ODN-14-SFAT T3 and T23, ODN-14-LFTT T6, and ODN-14-SFTT T3 via the methods described in **Chapter 2**. Using both methods allowed us to investigate stoichiometries under low nM (fluorescence) and  $\mu$ M (CD) conditions. Saturation of the DNA was not observed until at least 2 equivalents—typically saturation was observed between 2 and 3 equivalents. Fluorescence titrations were very similar to those determined via CD (**Fig. 3.2**). These results indicate that high stoichiometric ratios are not due to weak binding interactions only observed at high concentrations, nor are they artifacts related to dye interference or the hairpin loop. Additionally, these results indicate that



**Figure 3.1** CD spectra collected by Kristin Bales for NV1028 with different DNA oligonucleotides. PA equivalents ranged from 0-3.4 for ODN-14-SFTT (A), ODN-14-SFAT duplex (C), and ODN-14-SFAT (D), and from 0-4.5 for ODN-14-LFTT (B). Conditions: 5  $\mu$ M DNA, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 20  $^{\circ}$ C. CD spectra were averaged from 5 accumulations.



**Figure 3.2** Overlay of stoichiometry data collected by CD (O) at 333 nm and fluorescence (+) spectroscopy. A, ODN-14-SFTT; B, ODN-14-LFTT; C, ODN-14-SFAT and ODN-14-SFAT duplex, bottom. 5  $\mu$ M DNA (CD) or 2-30 nM DNA (Fluor), 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 20  $^{\circ}$ C (CD) or 25  $^{\circ}$ C (Fluor). CD: spectra were averaged from 5 accumulations, each point is the average mdeg of 2 spectra at a specific wavelength. Fluor: each point is the average of 3 intensities.

NV1028 interacts with the DNA similarly regardless of A/T pattern.

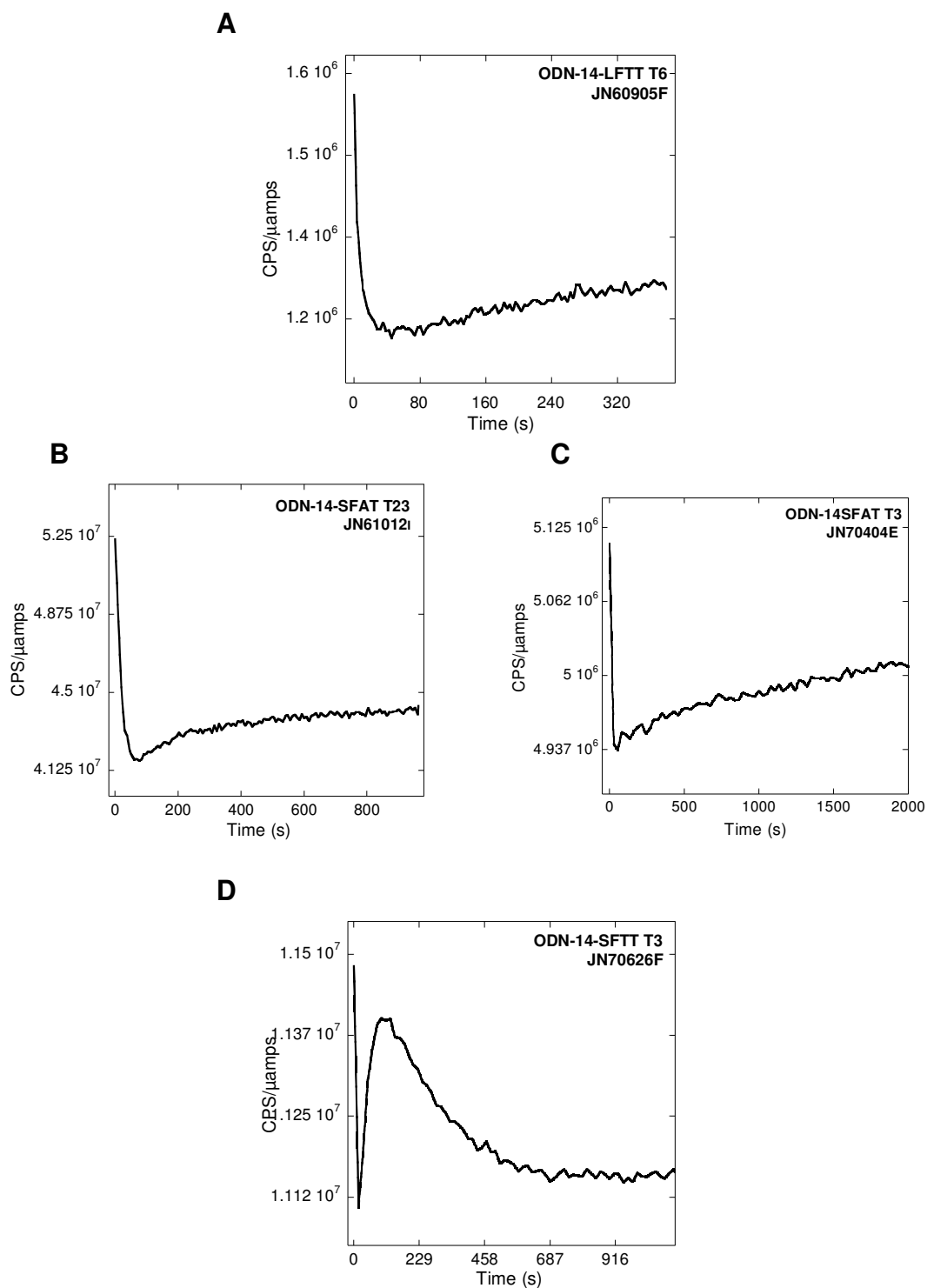
### **3.2.5. NV1028-DNA association kinetics**

#### **3.2.5.1. NV1028-DNA association rate constants**

Given the at least 2:1 PA-DNA binding stoichiometries discussed above, we were interested to know whether multiple binding events could be distinguished kinetically. As shown in **Fig 3.3**, for all DNA hairpins, at least 2 phases were clearly observed: an initial fast phase where the intensity decreased, and a second much slower phase where the intensity recovered. As discussed in **Chapter 1** biphasic binding kinetic traces have been observed via other fluorescent methods for the polyamide parent molecule Distamycin (Baliga & Crothers, 2000b). However, to my knowledge, biphasic binding has not been previously reported for other polyamides studied via SPR.

Despite greater than 2:1 PA-DNA binding stoichiometries for NV1028 with all the oligonucleotides tested, a unique third phase was not observed for ODN-14-LFTT T6, ODN-14-SFAT T3, and ODN-14-SFAT T23. However, three phases are observed for NV1028 with ODN-14-SFTT T3 (**Fig. 3.3C**). This kinetic trace has an initial fast phase where intensity decreases, a slow phase where the intensity recovers, then a third phase where the intensity decreases again.

To distinguish between an intermolecular event (concentration dependent) and a rearrangement (concentration independent), we performed quantitative association kinetics experiments by monitoring the change in fluorescence when varying concentrations of excess NV1028 were added to a fixed concentration of



**Figure 3.3** Representative examples of association kinetics curves. A, ODN-14-LFTT T6; B, ODN-14-SFAT T23; C, ODN-14-SFAT T3; D, ODN-14-SFTT T3. 0.5-2.5 nM DNA, excess NV1028, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C.

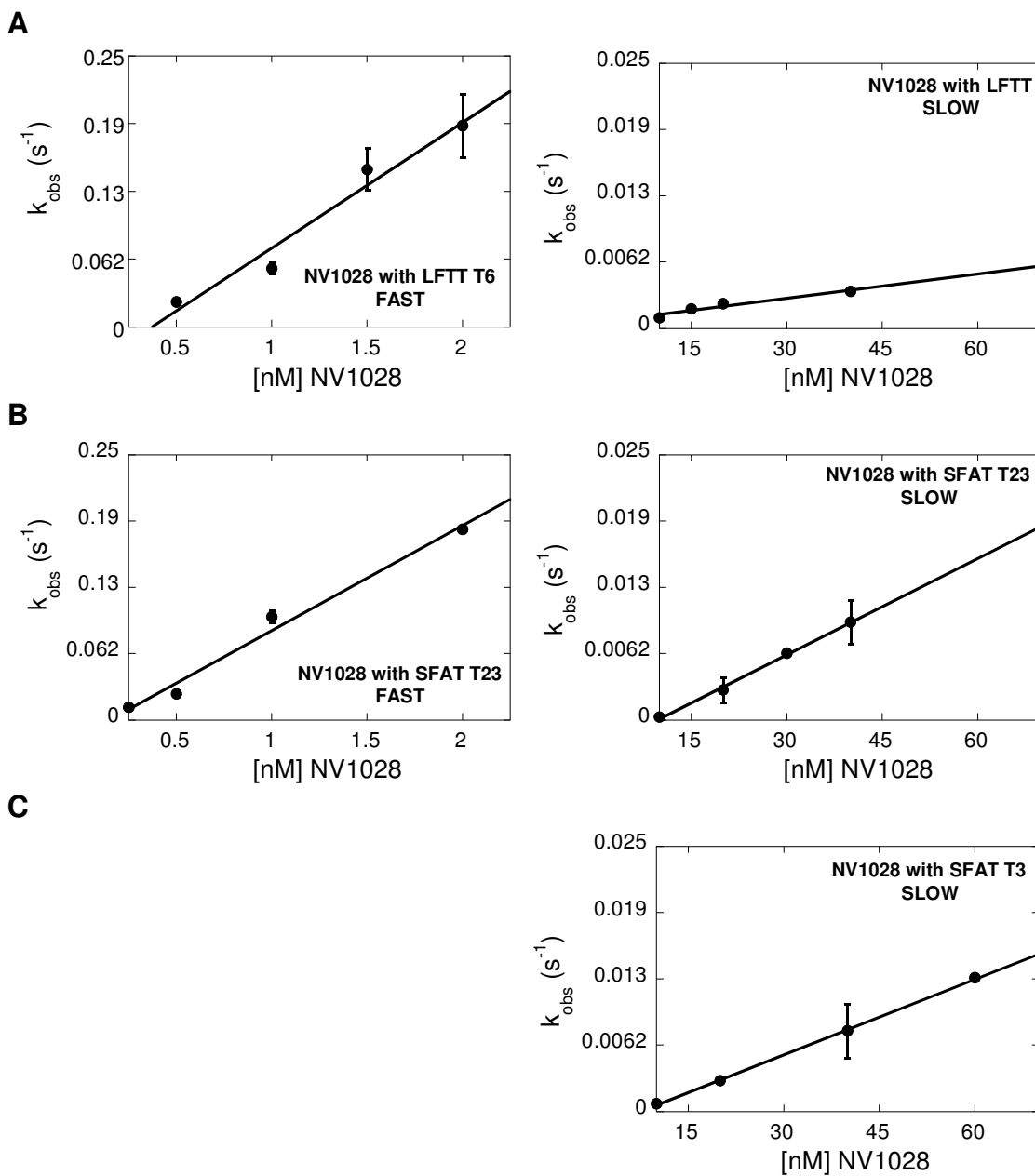
labeled DNA. The data from these experiments were separated into “fast phase” or “slow phase” and each phase was normalized and fitted independently as described in **Chapter 2**.

It was possible to obtain an association rate constant ( $k_{on}$ ) for the fast phases of ODN-14-LFTT T6 and ODN-14-SFAT T23 (**Fig. 3.4, Table 3.3**). These values indicate that the observed fast phase is concentration dependent and that  $k_{on\ fast}$  is similar regardless of the A/T pattern. These values also indicate that this binding interaction is diffusion-controlled. The fast phase for ODN-14-SFAT T3 was too fast at the concentrations required to achieve adequate signal-to-noise ratios.

The slow phase  $k_{on}$  values were also characterized for ODN-14-LFTT T6, ODN-14-SFAT T3, and ODN-14-SFAT T23 (**Fig. 3.4, Table 3.3**). These values are both exponentially ( $\sim 10^3$ ) slower than the initial phase and independent of A/T pattern. Furthermore, PA concentration dependence suggests the observed slow phase is likely a binding event and not related to a rearrangement after PA binding.

The lack of an observed third binding event for ODN-14-SFAT T3 and ODN-14-LFTT T6 despite binding stoichiometry greater than 2:1 may be due to unfavorable kinetics or a dye-related issue, especially given that a third phase is observed for NV1028 and ODN-14-SFTT T3. This is discussed in further detail in **section 3.3.5**.





**Figure 3.4** Secondary plots for NV1028 with ODN-14-LFTT T6 (A), ODN-14-SFAT T23 (B), and ODN-14-SFAT T3 (C). ODN-14-SFAT T3 fast phase kinetics not feasible via fluorescence. SFTT kinetics were not characterized. Conditions: 0.1 nM ODN-14-SFAT T23 or 0.5 nM ODN-14-SFAT LFTT T6 (fast phase), 2.5 nM DNA (slow phase), 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. The fast phase of NV1028 with ODN-14-LFTT T6 and both phases of NV1028 with ODN-14-SFAT T23 were determined in duplicate, relative error is shown; all others were determined once, fit error shown.

**Table 3.3** Association rate constants for NV1028<sup>a</sup>

<b>DNA</b>	<b><math>k_{\text{on fast}}</math> (<math>10^6 \text{ M}^{-1} \text{ s}^{-1}</math>)</b>	<b><math>k_{\text{on slow}}</math> (<math>10^6 \text{ M}^{-1} \text{ s}^{-1}</math>)</b>
ODN-14-LFTT T6	120±20	0.077±0.01
ODN-14-SFAT T23	98±0.2	0.28±0.08
ODN-14-SFAT T3	(160) <sup>b</sup>	0.16±0.08
ODN-14-SFTT T3	-- <sup>c</sup>	-- <sup>c</sup>

<sup>a</sup> The fast phase of NV1028 with ODN-14-LFTT T6 and both phases of NV1028 with ODN-14-SFAT T23 were determined in duplicate, relative error is shown; all others were determined once, fit error shown.

<sup>b</sup> calculated value based on experimental data. Too fast to be determined using a steady state instrument. See discussion in **section 3.3.2**.

<sup>c</sup> unable to characterize due to unusual dye behavior.

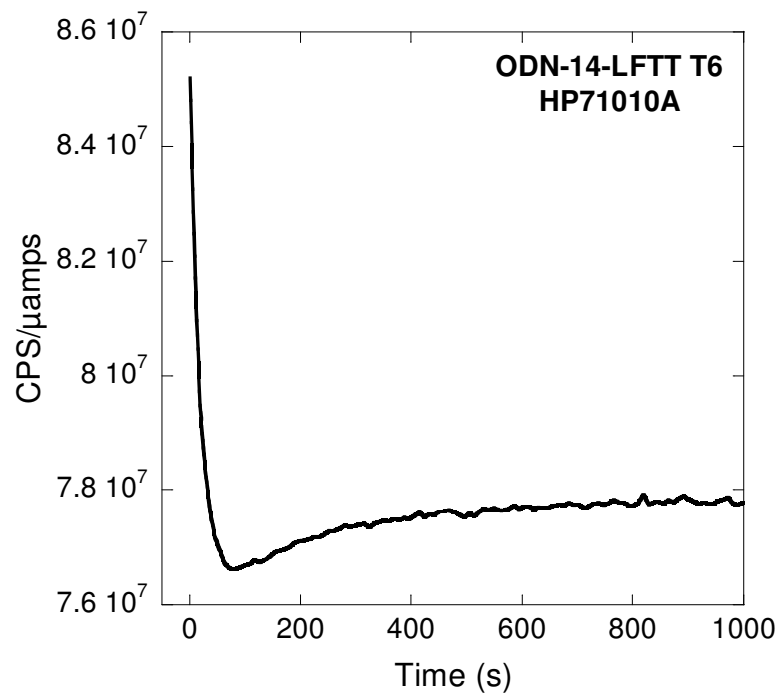
### **3.2.5.2. Multiple phases are observed when PA and DNA are 1:1**

Since association experiments were performed when PA was in excess of DNA, it was unclear if NV1028 preferred to bind to DNA 1:1 and the high stoichiometry was an artifact of high PA:DNA ratios, or if NV1028 preferred to bind greater than 1:1 regardless of PA:DNA ratios. Given that we determined the second phase was concentration dependent and, therefore, related to an intermolecular event, a qualitative association kinetics experiment was performed to see whether the second phase was still observed at 1:1 PA-DNA concentrations.

The experiment was performed with NV1028 and ODN-14-LFTT T6 by monitoring the change in intensity over time when 1 equivalent of NV1028 was added to DNA at concentrations at least 10-fold above the directly determined  $K_d$ . Under these conditions we would expect close to 100% saturation. If NV1028 were binding sequentially (i.e. prefers 1:1 binding mode) then the second phase should not be observed. However, at least 2 phases were observed at 1:1 (**Fig. 3.5**) which suggests that multiple NV1028 molecules bind to a single DNA hairpin molecule even when NV1028 is not in excess. This result suggests some type of cooperativity between the PA molecules. Additionally, this means that even at 1:1, excess free DNA is present which can negatively affect signal-to-noise ratios.

### **3.2.6. Macroscopic NV1028-DNA binding affinities ( $K_d$ )**

The observation of two unique association phases raised concerns about



**Figure 3.5** Association kinetics curve of NV1028 with ODN-14-LFTT T6 at 1:1. Conditions: 5 nM ODN-14-LFTT T6:NV1028, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. Data collected by Hyung Park.

equilibrium measurements. Specifically, since the association signal changes direction, the incubation time after each addition of PA becomes critical.

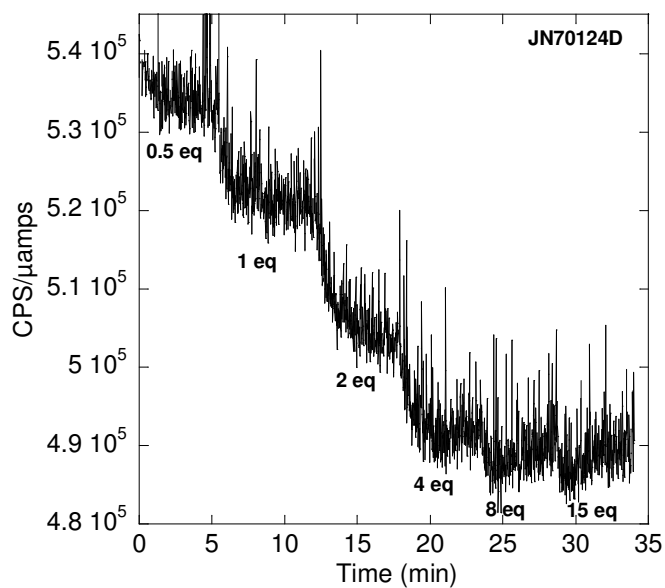
Additionally, it was unclear whether the  $K_d$ s determined were microscopic  $K_d$ s and related to a singular binding event—and, if so, which binding event, or if the  $K_d$ s were macroscopic  $K_d$ s where the number is a composite of multiple binding events and therefore multiple microscopic  $K_d$ s.

Equilibrium titration data were collected while monitoring the change in intensity of ODN-14-SFAT T3 in real time (**Fig. 3.6A**). With each addition of NV1028 added, the intensity decreases then a plateau in signal intensity is observed, followed by a slow recovery of fluorescence intensity (more easily observed at higher equivalents). The plateau is observed at around 5 min after each addition, the equilibration time used when collecting  $K_d$  for NV1028. The plateau region represents the stage at which the contributions to the dye intensity from both phases are at equilibrium. Thus, the binding affinities collected do not correspond to either phase and instead should be treated as macroscopic  $K_d$ .

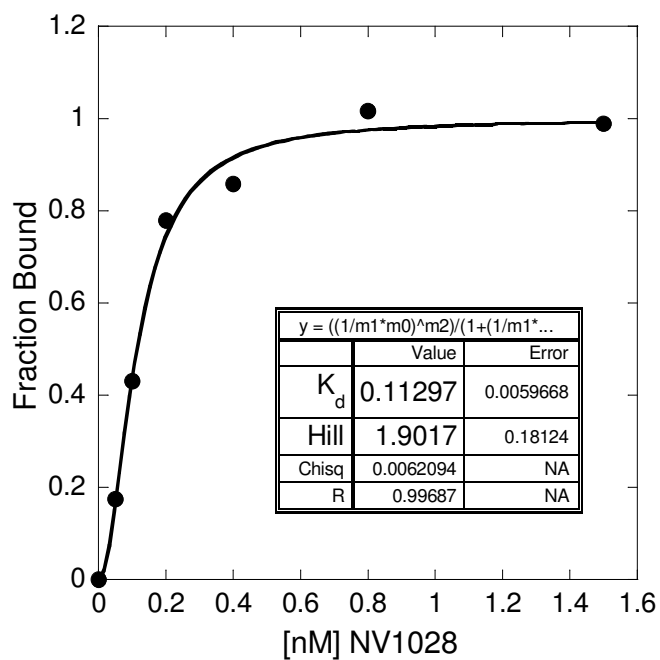
The average intensities of this plateau region were plotted, normalized, and fit per the protocol discussed in **Chapter 2 (Fig. 3.6b)**. The  $K_d$  was determined to be 0.11 nM, consistent with  $K_d$  collected for ODN-14-SFAT T3 via direct titration therefore incubation times are appropriate for collecting macroscopic binding affinities and we assume this is true of the other ODNs.

### **3.2.7. NV1028-DNA dissociation kinetics**

To characterize PA-DNA dissociation kinetics, the change in fluorescence



**A**



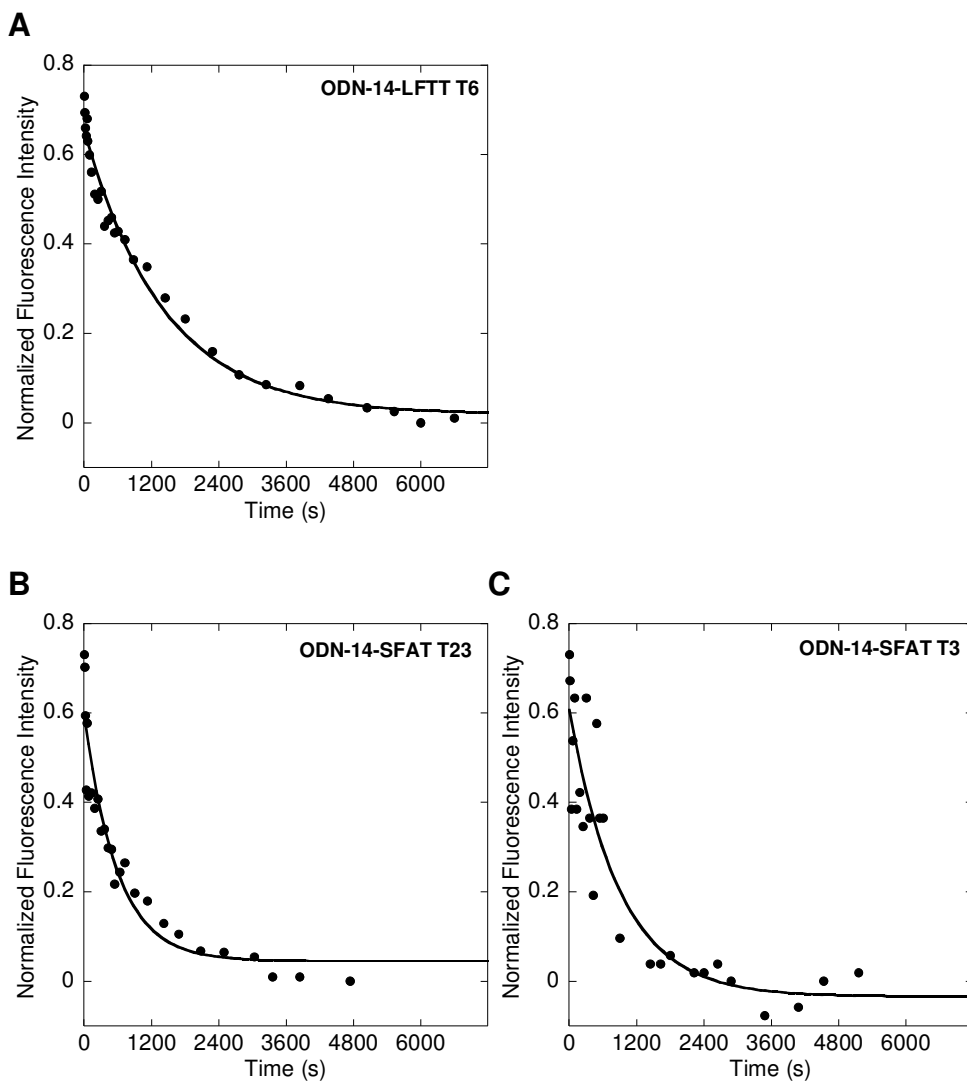
**B**

**Figure 3.6**  $K_d$  collected while monitoring real-time changes in fluorescence of ODN-14-SFAT T3. Real-time trace (A),  $K_d$  determined using average intensity in plateau region per each equivalent added (B). Conditions: 0.5 nM ODN-14-SFAT T3, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS.

of a 1:1 PA-DNA complex was monitored with respect to time after an excess of unlabeled competitor was added (see **Chapter 2**) (**Fig. 3.7**). Like with NV1028 competition experiments, low recovery of the dye intensity was observed in DNA dissociation kinetics experiments. Additionally, recovery of dye intensity was extremely slow: the typical dissociation experiment required over an hour to complete. Different than what was observed for association kinetics, fluorescence intensity during dissociation experiments was uniform in the direction of change resulting in a curve that appeared to be monophasic.

Given that more than one NV1028 binds to DNA and that high stoichiometric binding is observed even at 1:1, from the lack of recovery in our experiments we infer that not all NV1028 is dissociating from DNA. Furthermore, this implies that the dissociation rate constants determined here can only be treated as macroscopic dissociation rate constants ( $K_{\text{off}}$ ) since we currently have no way of knowing exactly how many NV1028 are dissociating during these experiments. Additionally, since we are unable to determine microscopic dissociation rate constants, and microscopic binding affinities, we are unable to characterize cooperativity at this time.

Raw intensities were normalized and fitted for comparison. Much like  $K_d$ , the  $K_{\text{off}}$  were similar among the different DNA hairpins. The average  $K_{\text{off}}$  value was  $1.2 \pm 0.5 \text{ s}^{-1}$  with a range of 0.6-1.6  $\text{s}^{-1}$  (**Table 3.4**). These  $K_{\text{off}}$  values tend to be at least 10-fold slower than most of the  $K_{\text{off}}$  of smaller PAs reported in **Chapter 1**. The  $K_{\text{off}}$  value determined with ODN-14-LFTT T6 is 2-fold slower than the  $K_{\text{off}}$  determined with ODN-14-SFAT T3 and T23. This difference may be due



**Figure 3.7** Representative examples of  $K_{off}$  data collected via fluorescence spectroscopy. ODN-14-LFTT T6 (A), ODN-14-SFAT T23 (B), ODN-14-SFAT T3 (C). Conditions: 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. 5 nM NV1028-Labeled DNA complex, 100 nM unlabeled competitor, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C.



**Table 3.4** DNA dissociation rate constants ( $K_{off}$ ) for NV1028

DNA	$K_{off}$ ( $\times 10^{-3} \text{ s}^{-1}$ )
ODN-14-LFTT T6	0.6 $\pm$ 0.07
ODN-14-SFAT T23	1.6 $\pm$ 0.05
ODN-14-SFAT T3	1.3 $\pm$ 0.2
Average	1.2 $\pm$ 0.5

Conditions: 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. 5 nM NV1028-Labeled DNA complex, 100 nM unlabeled competitor, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. Dissociation rate constants were determined in duplicate and relative error is shown.

to the A/T pattern. However, the larger flanks of ODN-14-LFTT T6 may be involved in non-specific binding interactions resulting in an observable difference in  $K_{off}$  values.

### **3.3. Discussion**

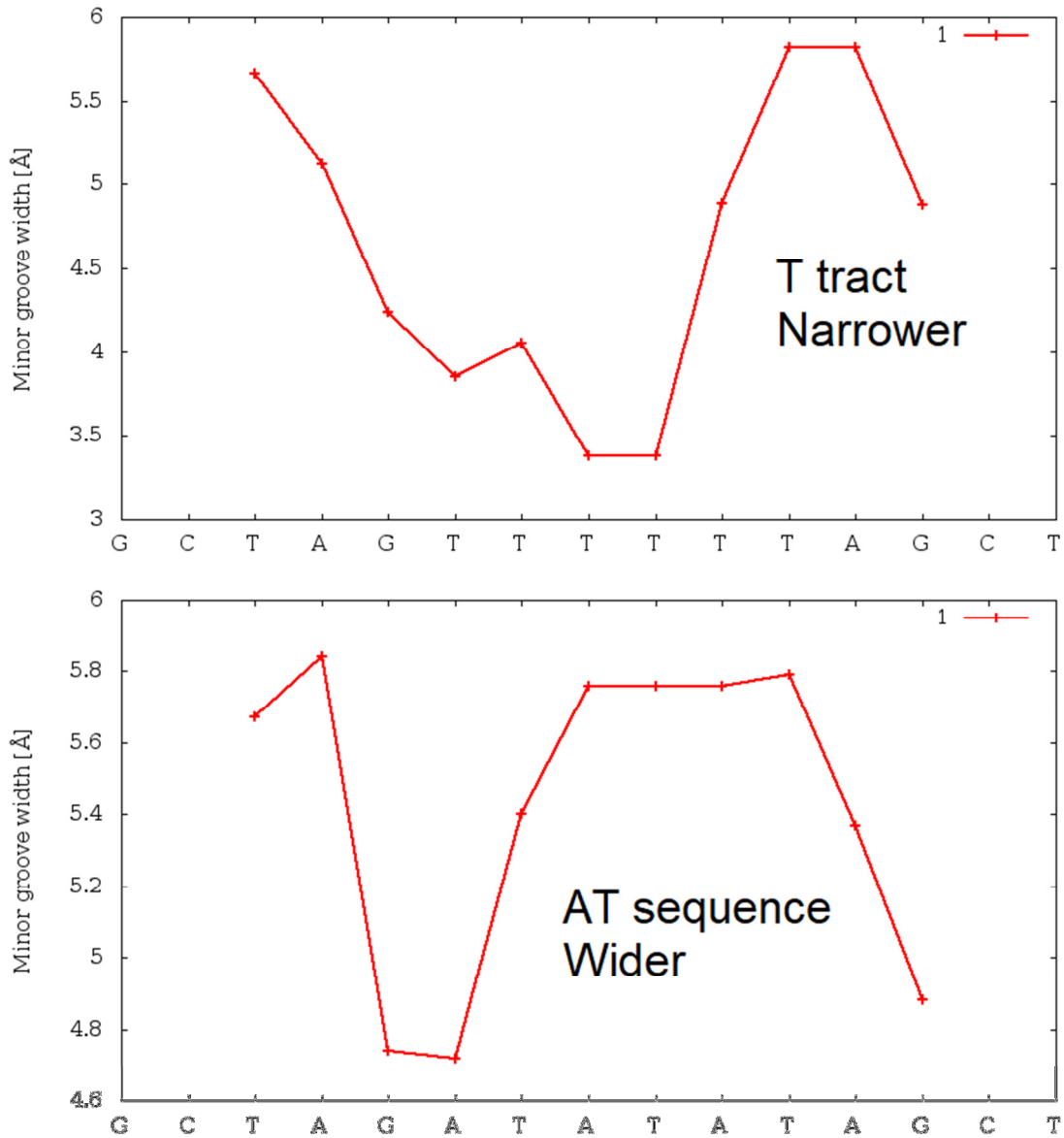
#### **3.3.1. Minor groove widths and NV1028-DNA binding kinetics**

It has been shown that minor groove widths are sequence dependent (Burkhoff & Tullius, 1988; Crothers & Shakked, 1999; Wu & Crothers, 1984). Additionally, it has been shown that Distamycin A (Dst) does not bind degenerately to all A/T pattern sequences and that the formation of the 2:1 complex is sequence dependent and tends to form preferentially in mixed AT sites (Chen & Sha, 1998).

The predicted minor groove widths of ODN-14-SFTT and ODN-14-SFAT are consistent with the literature (**Fig. 3.8**). However, despite the differences in minor groove width between the binding sequences used in this study, there is no difference in NV1028 DNA binding affinity or DNA binding kinetics with respect to the A/T pattern. Additionally, binding stoichiometry is conserved, as well.

#### **3.3.2. Reconciliation of kinetics and binding affinity data for AT and TT-dependence**

Since more than one NV1028 binds to each hairpin studied here, and multiple phases have been observed, the measured  $K_d$ s and  $K_{off}$ s of 1:1 complexes are by necessity macroscopic. Coupled with the microscopic  $k_{on}$



**Figure 3.8** DNA minor groove width prediction for ODN-14-SFTT (top), and ODN-14-SFAT (bottom). Generated by Yang song using (<http://rohslab.cmb.usc.edu/DNAshape/>).

values determined here, these data provide an opportunity to reconcile parameters obtained independently and associated with different events.

Given the presence of two discernible microscopic concentration-dependent association phases for ODN-14-LFTT T6 and ODN-14-SFAT T23, a macroscopic association rate constant ( $K_{on}$ ) was calculated for these DNAs using **eqn. 3.2 (Table 3.5)**.  $K_{on}$  values determined with ODN-14-LFTT T6 and ODN-14-SFAT T23 were not only very similar to each other, but also to  $k_{on}$  values reported for 8-ring hairpins ( $\sim 3 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ , averaged  $k_{on}$  values of 8-ring hairpin PAs discussed **Chapter 1**).

$$K_{on} = \sqrt{k_{on \text{ fast}} * k_{on \text{ slow}}} \quad (3.2)$$

Using  $K_{on}$  and  $K_d$ ,  $K_{off}$  was calculated using **eqn. 3.3** and compared to the directly determined  $K_{off}$  values reported in **section 3.2.7** and were found to be internally consistent with directly determined  $K_{off}$ , as shown in **Table 3.5**.

$$K_d = \frac{K_{off}}{K_{on}} \quad (3.3)$$

If we assume the fast phase  $k_{on}$  for NV1028 with ODN-14-SFAT T3 is concentration dependent, then experimentally determined  $K_d$ ,  $K_{off}$ , and slow phase  $k_{on}$  can be used to calculate the fast phase  $k_{on}$  to be around  $160 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$  (**Table 3.3**). This is consistent with the experimentally determined fast phase  $k_{on}$  for ODN-14-LFTT T6 and ODN-14-SFAT T23.

**Table 3.5** Comparison of Experimental and Computed NV1028-DNA binding constants<sup>a</sup>

DNA	Direct $K_d$ (nM)	$k_{on\ fast}$ ( $10^6\ M^{-1}\ s^{-1}$ )	$k_{on\ slow}$ ( $10^6\ M^{-1}\ s^{-1}$ )	Calculated $K_{on}^b$ ( $10^6\ M^{-1}\ s^{-1}$ )	$K_{off}$ ( $10^{-3}\ s^{-1}$ )	Calculated $K_{off}^c$ ( $10^{-3}\ s^{-1}$ )
ODN-14-LFTT T6	0.23±0.02	120±19	0.077±0.01	3.0	0.6±0.07	0.69
ODN-14-SFAT T23	0.16±0.02	98±0.2	0.28±0.08	5.2	1.3±0.2	0.83

<sup>a</sup> The fast phase of NV1028 with ODN-14-LFTT T6 and both phases of NV1028 with ODN-14-SFAT T23 were determined in duplicate, relative error is shown; all others were determined once, fit error shown.

<sup>b</sup> computed using  $K_{on} = \sqrt{k_{on\ fast} * k_{on\ slow}}$

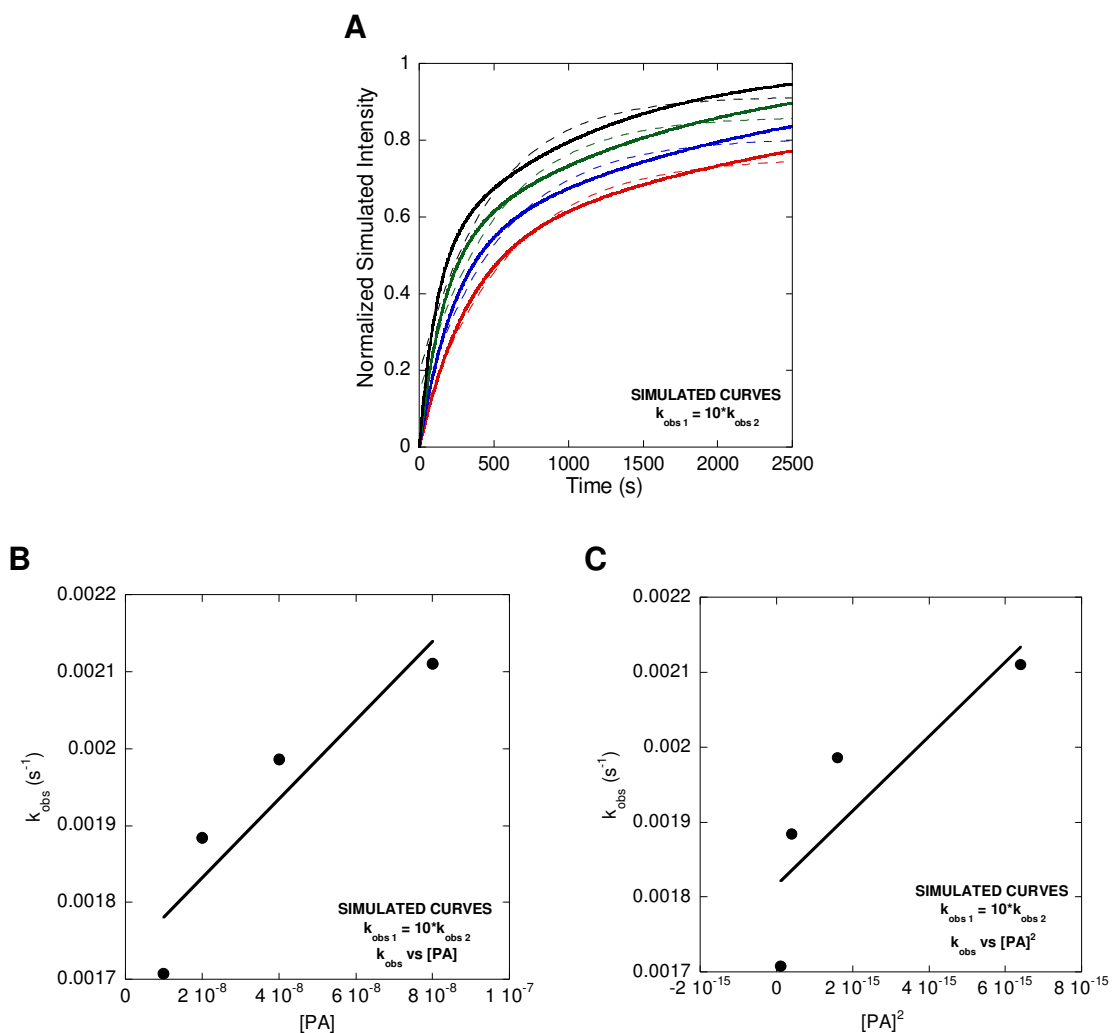
<sup>c</sup> computed using  $k_{off} = K_{on} * K_d$

### 3.3.3. Probing the slow phase for an additional binding event

Given that a binding stoichiometry of at least 2:1 PA:DNA was observed via fluorescence and CD spectroscopy, the observation of only 2 phases for NV1028 binding ODN-14-LFTT T6, ODN-14-SFAT T23, and ODN-14-SFAT T3 was initially puzzling. Since we could observe the third binding event under nM conditions used for fluorescence spectroscopy stoichiometry experiments, we assumed the third event happened on a timescale reasonable enough for us to observe during DNA association experiments. Further supporting this assumption, association traces collected for ODN-14-SFTT T3 indicated the third binding event happened within the timeframe used for our DNA association experiments. Thus, we wondered whether the slow phase did, in fact, correspond to only a single binding event.

#### 3.3.3.1. Biphasic curves may fit well to monoexponential equations

We performed simulations to determine if 2 events with different rate constants are well represented by a monoexponential model and appear to have linear concentration dependence with respect to a single equivalent of PA. Association curves were simulated using a biexponential kinetics equation (**eqn. 3.1**) and then fit using a monoexponential kinetics equation (**eqn. 2.3**). The resulting monoexponential  $k_{\text{obs}}$  were then plotted with respect to PA concentration and fit linearly (**eqn. 2.4**). Simulations indicated that if  $k_{\text{obs } 1}$  and  $k_{\text{obs } 2}$  were 10-fold different from each other, association were represented by a monoexponential equation with favorable goodness of fit (R-values) (**Fig. 3.9A**).



**Figure 3.9** Simulated association curves generated using a biexponential equation. A, simulated association curves generated with a biexponential equation (**eqn. 3.1**) (solid), then fit with a monoexponential equation (**eqn. 2.3**) (dashed) to determine  $k_{obs}$ ; secondary plots of simulated  $k_{obs}$  versus  $[PA]$  (B), and  $[PA]^2$  (C) (**eqn. 2.4**).

Secondary plots of the monoexponential  $k_{obs}$  with respect to [PA] and [PA]<sup>2</sup> both have some curvature. However, given the goodness of fit of the biphasic curves by a monoexponential equation and that more linearity is observed when  $k_{obs}$  is plotted versus [PA] it would be easy to dismiss the curvature as noise and any deviation from the linear fit as experimental error.

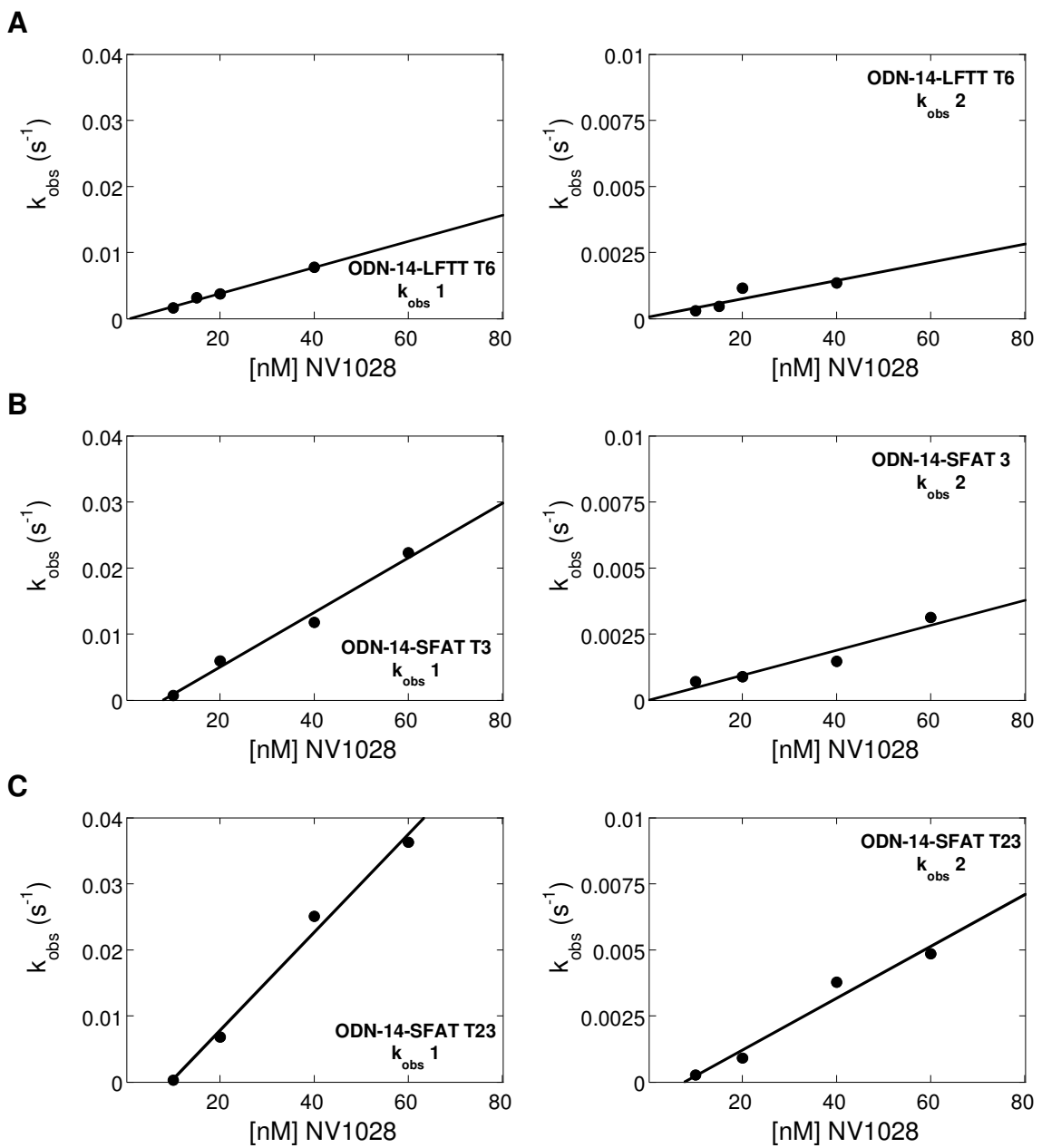
$$y = 1 - (e^{-k_{obs1}t} + e^{-k_{obs2}t}) \quad (3.1)$$

Simulations were also performed where  $k_{obs1} = k_{obs2}$  and when  $k_{obs1}$  and  $k_{obs2}$  were 100-fold different (not shown). Monoexponential fits of these lines were poor and no linear concentration dependence with respect to PA was observed. The simulation results show that 2 binding events that have association rate constants that are ~10-fold different are well-represented by a monoexponential model.

### 3.3.3.2. Biexponential fits of slow phase curves

Given the results of the simulation, biexponential fitting was performed on the slow phase association curves. Using **eqn. 3.1**, the slow phase curves were fitted and  $k_{obs1}$  and  $k_{obs2}$  were plotted with respect to PA concentration and fit linearly (**Fig. 3.10**). Consistent with simulated results,  $k_{obs1}$  and  $k_{obs2}$  were approximately 10-fold different from each other. Two linear concentration dependent lines were determined for NV1028 association with all ODN tested indicating the slow phase represents 2 binding events.





**Figure 3.10** Representative secondary plots of NV1028 slow phase association curves when fit to a biexponential equation. A, NV1028 with ODN-LFTT T6; B, NV1028 with ODN-14-SFAT T3; C, NV1028 with ODN-14-SFAT T23.

The respective concentration dependent rate constants for ODN-14-LFTT T6 and ODN-14-SFAT T23 determined via biexponential fitting of the NV1028 slow phase association curves were similar to each other (**Table 3.6**). However, rate constants determined for ODN-14-SFAT T3 were ~10-fold faster than the respective constants determined for the other 2 oligonucleotides. It is unclear if this difference is simply a fitting artifact or if this means something. Additionally, if this does mean something, it is unclear what.

If correct, this data would indicate that three concentration-dependent binding events are occurring. However, it is still unclear where and how more than one equivalent is binding to the DNA. Binding stoichiometry experiments are performed at relatively high concentrations (10-fold above the  $K_d$ ) which are kinetically favorable for observing slower interactions. It is possible that at lower concentrations (around  $K_d$ ), like those used in the association kinetic experiments reported here, the third binding event is too slow to be observed. Additionally, given that it seems the microscopic  $k_{on}$  of the third phase is very slow, it may not have a significant impact on the macroscopic  $K_{on}$  which is why data is internally consistent even without characterization of a third phase.

### **3.3.4. NV1028-DNA binding stoichiometry**

#### **3.3.4.1. High PA-DNA binding stoichiometries are consistent with previously reported data**

Isotherms collected by Elena Vasilieva using DNase I footprinting and fit to the Hill equation indicated that more than one equivalent of NV1028 may be binding to a specific sequence (Vasilieva, 2014; Vasilieva et al., 2016).

**Table 3.6** Concentration dependent rate constants for NV1028 slow phase when fit to a mono or biexponential equation<sup>a</sup>

DNA	Monoexponential	Biexponential	
	$k_{on\ slow}$ ( $10^6 M^{-1} s^{-1}$ )	$k_{on\ 1}$ ( $10^6 M^{-1} s^{-1}$ )	$k_{on\ 2}$ ( $10^6 M^{-1} s^{-1}$ )
ODN-14-LFTT T6	0.077±0.01	0.20	0.04
ODN-14-SFAT T3	0.28±0.08	1.30	0.20
ODN-14-SFAT T23	0.16±0.08	0.41	0.05

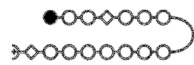
<sup>a</sup>The slow phase of NV1028 with ODN-14-SFAT T23 was determined in duplicate, relative error is shown; all others were determined once, fit error shown.

Experiments described here confirm this: NV1028-DNA binding stoichiometry was determined to be at least 2:1. But these data were inconsistent with Dervan's rules and our understanding of how hairpin PAs were designed to bind to DNA.

However, CD spectroscopy data suggesting that more than 1 equivalent of hairpin PA binds to an isolated singular binding site has been presented—but not addressed or discussed in detail—in the literature (Buchmueller et al., 2005; Castaneda, 2017; Liu, B. et al., 2017; Pilch et al., 1996; Wang, S. et al., 2014; Wang, S. et al., 2012), see **Chapter 1**. Gaofei He reported 2:1 binding for NV1028 with duplex GH6084C which contains an isolated perfect match (**Fig. 3.11A**) (not published). Kristin Bales repeated the experiments with GH6084C to higher equivalents and determined a stoichiometric ratio of greater than 2:1 (**Fig. 3.11BC**).

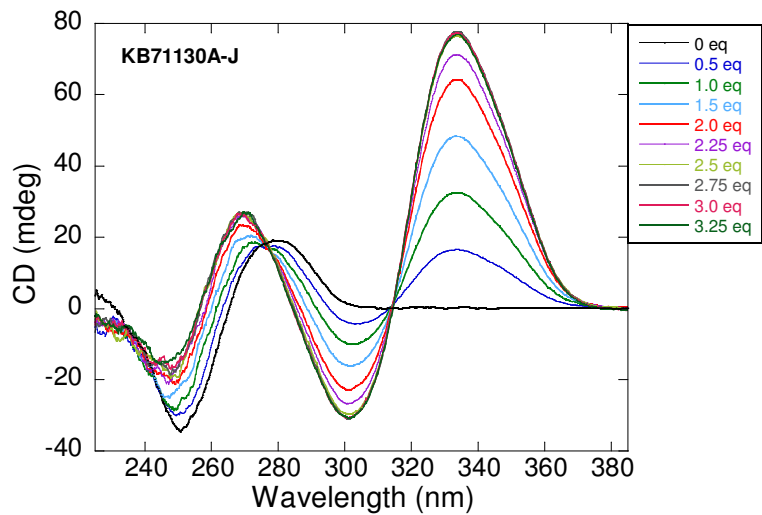
As discussed in **Chapter 1**, DNA binding stoichiometries were also determined by Carlos Castaneda for NV1028, and an NV1028 analog that contained a chiral (R)-NH<sub>2</sub> at the  $\gamma$ -turn (Castaneda, 2017). While Castaneda did observe saturation at greater than 1:1, his results for NV1028 with a perfect match site are different than the data here given that our data indicate at least 2 or more equivalents bind to the DNA. However, the differences may be due to variances in the method since Castaneda's CD titrations were performed without CHAPS.

**A**  
5'-GCGAAGTTAATATGCG-3'

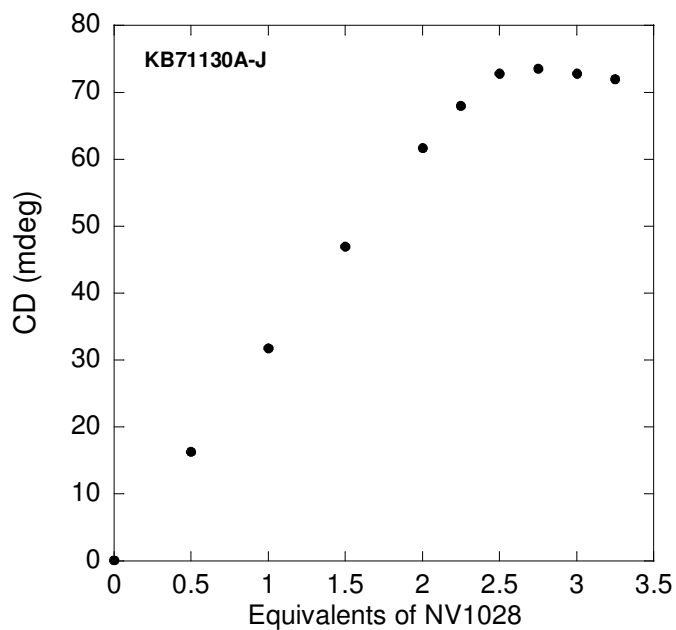


3'-CGCTTCAATTATACGC-5'

**B**



**C**



**Figure 3.11** CD spectrum (B) and titration plot (C) for NV1028 with GH6084C (A), shown with NV1028. Data collected by Kristin Bales.

### 3.3.4.2. NV1028-DNA binding modes

The high PA:DNA binding stoichiometry data begs two very important questions: on a small DNA hairpin with an isolated binding site, where and how are the additional equivalents binding? Vasilieva and He have shown that NV1028 is capable of binding to sites containing multiple mismatches with little to no difference in the  $K_d$  when characterized via DNase I footprinting on large viral fragments (He et al., 2014; Vasilieva et al., 2016). It has been suggested that because larger hairpin PAs contain more rings capable of hydrogen bonding and internal  $\pi$ -stacking interactions, unfavorable enthalpic interactions may have less impact on PA binding (Koeller et al., 2014). It should be noted that while the larger hairpin PAs have been shown to bind to mismatched sites, they do not bind indiscriminately. Unique and replicable footprinting and affinity cleavage patterns have been observed for NV1028 and NV1042, a 20-ring hairpin PA (He et al., 2014; Vasilieva, 2014; Vasilieva et al., 2016).

This suggests that not only is NV1028 capable of binding to mismatched sites, but also that not all residues of NV1028 are required for the PA to bind avidly to DNA. This also suggests that the larger antiviral PAs select their binding sites using more than just minor groove hydrogen bond topography. However, the lack of difference in both binding affinity and binding kinetics between A/T patterns means groove widths do not matter for PA binding site selection. Our data also indicate that multiple binding is preferred even at substoichiometric ratios, which indicates some type of cooperativity. While it remains unclear what exactly drives large hairpin PA binding site selection, it is apparent that Dervan's

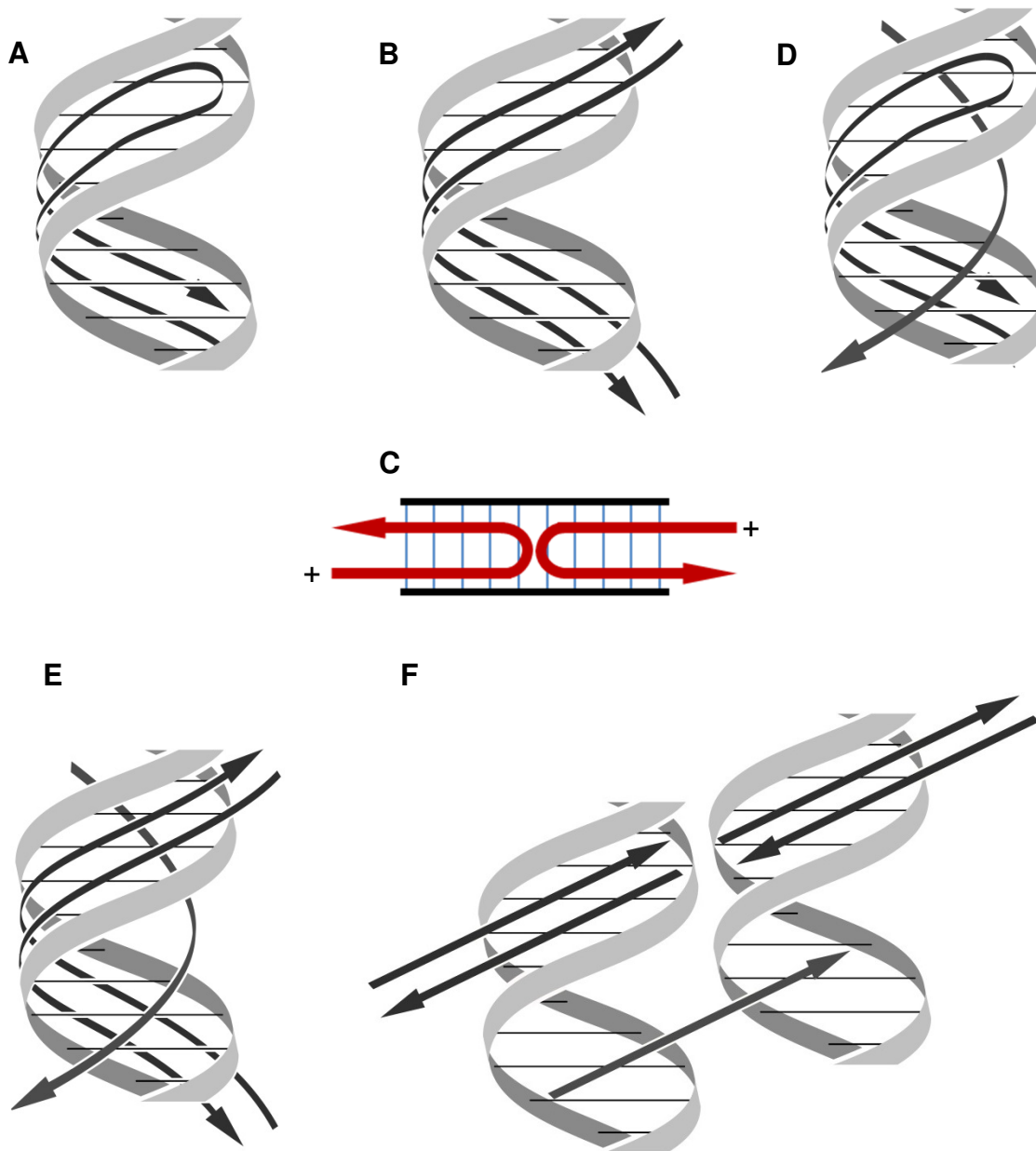
rules do not adequately address the binding interactions between larger hairpin PAs and DNA.

In a recent study by Qiao and coworkers more than one binding mode is observed via CD titrations with an 8-ring hairpin PA (Qiao et al., 2015). Through single value decomposition analysis (SVD) of the CD spectra, an additional mode is observed even at low PA-DNA ratios. This finding is consistent with our own in that a second phase is observed even when PA-DNA ratios are 1:1. Below, a number of binding models consistent with our data are discussed.

### ***Binding modes if 2:1 PA:DNA***

If we assume that the first equivalent binds per Dervan's rules (shown in **Fig. 3.12A** for reference), then the second equivalent would have to bind non-specifically and in an orientation not previously observed. The second equivalent could theoretically interact electrostatically whereby the positively charged C-terminus of the PA interacts with the negatively charged phosphate backbone of DNA (not shown). However, very little salt-dependence was observed by Yang Song in his thermodynamic characterization so an electrostatic interaction seems unlikely.

If the first equivalent is not bound per Dervan's rules, but binding is only occurring in the minor groove, then a couple models may explain how they are binding (**Fig. 3.12BC**). The first model is where the PAs are not hairpins, but linear (**Fig. 3.12B**). In order to accommodate this type of binding, not all of the PA residues would be bound to DNA—there are simply not enough base pairs in



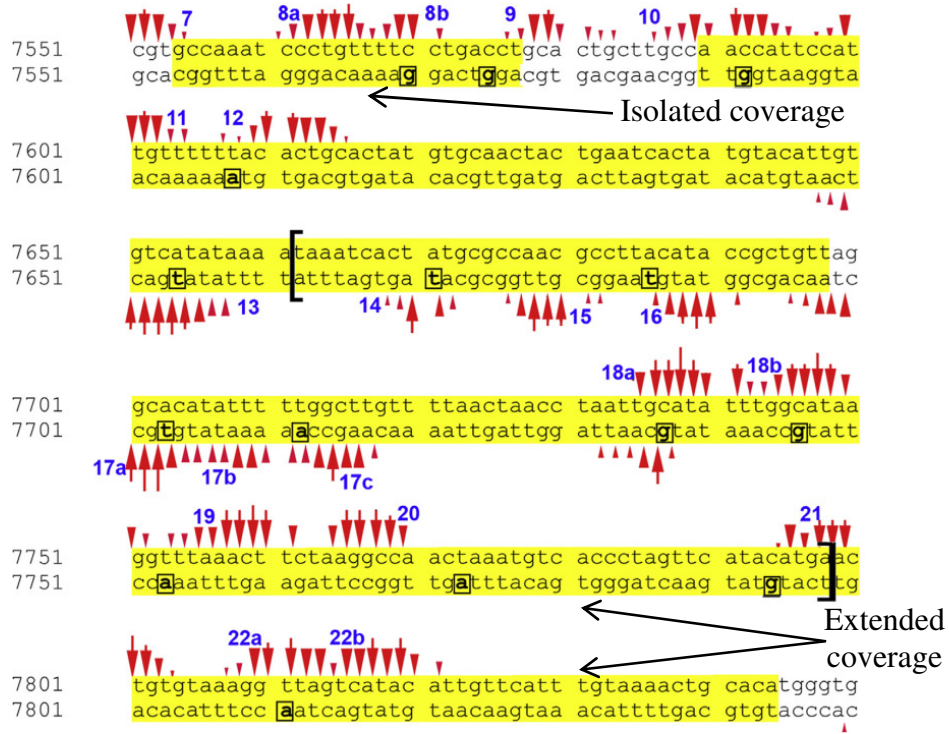
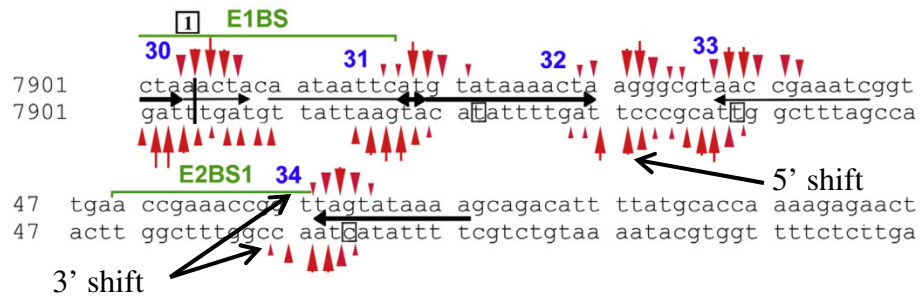
**Figure 3.12** Binding models for multiple equivalents of NV1028. A, 1:1 binding where the hairpin binds to its canonical binding sequences based on Dervan's rules; B, 2:1 binding where both equivalents bind linearly in the minor groove; C 2:1 binding where both equivalents bind linearly in the minor groove as hairpins (bottom); D, 2:1 binding where one equivalent binds in each of the grooves; E, 3:1 binding where 2 PAs bind linearly in the minor groove and the third binds in the major groove; F, 2.5:1 where a bridging interaction occurs between 2 DNA molecules.



our oligonucleotides: as a hairpin NV1028 was designed to bind to a 10 bp sequence, the DNA hairpins used in this study are only 15 bp (“short flank” sequences) or 20 bp (“long flank” sequence). At the very least, this model calls into question our understanding of how many bonds are required for a PA to avidly bind to DNA.

The second model is where the PAs both bind in the minor groove as hairpins (**Fig. 3.12C**). As discussed with the linear model, if two PA are binding as hairpins then not all of the residues can be bound to the DNA. Since NV1028 is capable of binding in either the forward or reverse orientation (**Fig. 2.1**), it is likely that due to electrostatic repulsion by the positively charged C-terminus, the hairpins would orient themselves in whatever way limited the repulsion. This model would also imply that there may be a minimal number—or ratio—of binding interactions that occur between the large hairpin PAs and DNA even when a large perfect match site is present and that not all rings of large PAs may be required to bind securely to DNA.

DNase I footprinting and affinity cleavage data could support either the extended linear or hairpin binding mode of NV1028 (**Fig. 3.13**) (Vasilieva et al., 2016). In affinity cleavage experiments, the FeEDTA moiety is placed at the C-terminus of the PA, thus orientation of the PA can be discerned based on the cleavage patterns. In DNase I footprinting cleavage occurs at areas where there is no PA bound. Taken together the data can be used to characterize binding orientation and modes. As seen in **Fig. 3.13A**, there are areas of the viral DNA fragment where coverage exceeded 10 bp without an affinity cleavage site to

**A****B**

**Figure 3.13** Sequence maps of NV1028 data collected via affinity cleavage and DNase I footprinting by Elena Vasilieva. Blue numbers correspond to affinity cleavage sites as numbered in (Vasilieva et al., 2016). Red vertical arrows indicate intensities relative to other arrows within a site. Yellow shading indicates where footprints are observed at 10 nM NV1028. Solid horizontal arrows indicate the binding site consistent with observed affinity cleavage pattern. Adapted from (Vasilieva et al., 2016), permission request is pending.

rationalize the PA binding as a hairpin. However, there are also areas of the viral fragment where footprints are isolated and the affinity cleavage sites can be used to rationalize hairpin binding. Thus, the data indicates NV1028 may be binding in either the extended linear or the hairpin binding mode or both binding modes.

Another model is also possible where the first equivalent binds per Dervan's rules, but the second equivalent interacts with the major groove (**Fig. 3.12D**). As discussed previously for Dst, in order to accommodate a second equivalent, the minor groove must widen (and, therefore, the major groove narrows) and crystal structures of 8-ring cyclic PAs have shown these same effects on DNA groove widths (Chenoweth & Dervan, 2009). Perhaps when the initial hairpin binds in the binding site, the major groove narrows enough to make binding a second equivalent favorable. Major groove binding has not been reported for hairpin PA of any size, but previous affinity cleavage data may support this hypothesis. Characterization of minor groove binders via affinity cleavage will display a 3' shift from the canonical binding sequence (**Fig. 3.13B**). However, this is not always the case for NV1028 where 5' shifts are observed with respect to some of the binding sites (Vasilieva, 2014; Vasilieva et al., 2016). This data may suggest that DNA major groove binding may be occurring. However, this data may also simply suggest that the PA has "slipped" and the PA is bound to a non-specific sequence.

### ***Binding modes if 3:1 PA:DNA***

Both fluorescence and CD spectroscopy data indicate that saturation of

the DNA occurs above 2:1. The models would not differ very much from those discussed for 2:1 binding, but it is likely there may be a combination of 2 or more binding modes occurring in order to accommodate the third equivalent. For example, 2 equivalents may be bound in the minor groove and a third equivalent may be interacting in or with the major groove (**Fig. 3.12E**).

### ***Binding modes if 2.5:1 PA:DNA***

As discussed previously, data from our group and the Bashkin group indicates saturation occurs at 2.5:1 for both CD and fluorescence spectroscopy experiments. A possible explanation for this ratio is that within the reaction mixture there may be 2 populations of PA-DNA complexes: one population where PAs bind to DNA in a 2:1 mode, and the other where PAs bind to DNA in a 3:1 mode. In order for this model to be accurate, one binding mode would not be dominant, and a PA would be equally likely to bind at 2:1 as 3:1.

But, if 2.5:1 binding is not an average, then how does a half equivalent bind to DNA? In order to properly visualize how a half equivalent can bind to DNA, it is helpful to discuss this model as a 5:2 binding model where one NV1028 molecule acts as a bridge between 2 DNA molecules (**Fig. 3.12D**). We propose this bridging interaction is not only possible *in vivo*, but also may have broader implications for the antiviral mechanism of large antiviral PAs and will be discussed further in **Chapter 4**.

## CHAPTER IV

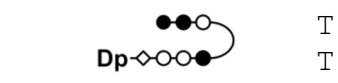
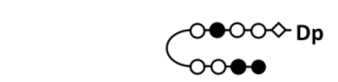
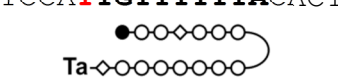
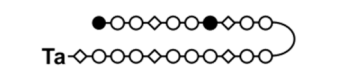
### SIZE-DEPENDENCE OF HAIRPIN POLYAMIDE-DNA BINDING AFFINITIES AND KINETICS

#### 4.1. Introduction and background

It has been shown that anti-HPV activity is only found with large hairpin PAs that are capable of binding to 10 or more bp, such as NV1028 and NV1042, 14- and 20-ring hairpin PAs (**Table 4.1**).  $K_d$ s determined via DNase I footprinting of NV1042 and NV1028 to sites with up to 2 mismatches were very similar, around 2-5 nM (He et al., 2014; Vasilieva et al., 2016). However, NV1042 is 3 times more potent than NV1028 when comparing both the  $IC_{50}$  and  $IC_{90}$ 's while smaller hairpin PAs, such as KA1039, a 6-ring hairpin PA, are inactive as antivirals (**Table 4.1**) (Edwards et al., 2011). Yet binding kinetics have only been characterized for smaller hairpin PAs, such as KA1039 and KA1002, 6- and 8-ring hairpin PAs, respectively (Bashkin, J. K. et al., 2013; Wang, S. et al., 2014; Wang, S. et al., 2013; Wang, S. et al., 2012).

By studying hairpin PAs of other sizes and potencies, we have the opportunity to determine if the unusual properties observed for NV1028 (biphasic kinetics and high DNA binding stoichiometries) extend to other PAs that differ in size and if those properties correlate to antiviral activity. Additionally, since KA1039 binding kinetics have already been characterized via SPR, we also have the opportunity to validate our method for studying DNA binding kinetics via fluorescence spectroscopy. Additionally, as discussed in **Chapter 1**, no

**Table 4.1** DNA hairpins relevant to **Chapter 4**<sup>a</sup>

PA	# of rings	DNA	Lab Designation	Sequence	HPV16 <sup>b</sup>	
					IC <sub>50</sub> (μM)	IC <sub>90</sub> (μM)
KA1039 KA1002	6 8	ODN-6 <b>ODN-6-T3</b>	YS07108 <b>YS04109</b>	5' -CC <b>TTGGCT</b> TC T  3' -GG <b>AACCGA</b> AG T	Not active	
KA1002	8	<b>ODN-8-T37</b> <sup>c</sup>	<b>CD24111b</b>	5' -CCTGGAGAGG <b>AAGCCA</b> AGTG T  3' -GGACCTCT <b>CC TTCGGT</b> TCAC T	Not active	
NV1028	14	ODN-14-LFTT <b>ODN-14-LFTT T6</b>	JN07075 <b>JN02076-13</b>	5' -TTCCA <b>TTGTTTTTTA</b> CACTG T  3' -AAGGT <b>AACAAAAA</b> ATGTGAC T	0.100	1.113
NV1042	20	ODN-20 <b>ODN-20-T3</b> <b>ODN-20-T16</b>	YS07108 <b>CD26043</b> <b>JN04058</b>	5' -GC <b>TATGTTTAAGATA</b> TGCT T  3' -CGA <b>TACAAATTC</b> TATACGA T	0.036	0.351

<sup>a</sup> Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop. Expected binding sequences are **bold**. Colors denote dye position: primary dye positions are labeled in **red**, alternate dye positions are shown in **green**. **Black** ODNs are unlabeled and used as competitors or for CD spectroscopy.

<sup>b</sup> Antiviral data for NV1028 and NV1042 determined by (Edwards et al., 2011)

<sup>c</sup> previously published  $k_{off}$  data with ODN-8-T37, referred to as HP1, in (Dupureur et al., 2012)

systematic study of hairpin PA size-dependence has been conducted thus far, and due to subtle variations in PA sequence, structure, and methods, no conclusions can be drawn from the literature

Here I report on a study of hairpin polyamide size dependence of DNA binding kinetics. Using the fluorescence assay described in **Chapter 2**, I determined the kinetic rate constants and binding stoichiometries of 3 PAs ranging in size from 6 to 20-rings with their cognate DNA sequences.

#### **4.1.1. Design of DNA hairpins used for this study**

All DNA hairpins relevant to this chapter are shown in **Table 4.1**. DNA hairpins were designed by isolating perfect match sites using Dervan's rules of recognition, as discussed in **Chapter 2** (Kielkopf et al., 1998a; Kielkopf et al., 1998b). The design of the KA1039 DNA hairpin, ODN-6-T3, was based on the biotinylated hairpin used in previous kinetics studies by SPR (Wang, S. et al., 2013; Wang, S. et al., 2012). TAMRA was added at the T3 position.

NV1028 characterization has been previously reported in **Chapter 3**. The results discussed in this chapter will be those with ODN-14-LFTT T6, unless otherwise stated, since we determined the A/T pattern in the recognition sequence had no impact on DNA binding kinetics. ODN-14-LFTT T6 contains the perfect match sequence located at 7600-7609 of the HPV16 genome and is flanked by 5 bp of the natural sequence at both the 5' and 3' ends. TAMRA is at the T6 position.

NV1042 characterization was performed with ODN-20-T16, unless otherwise stated. ODN-20-T16 contains the perfect match sequence located at 3652-3664 of the HPV16 genome and is flanked by a GC/CG pair at both the 5' and 3' ends. TAMRA is at the T16 position. Characterization with a hairpin with TAMRA at T3 was attempted but technical issues made use of this hairpin with this assay unfeasible, as discussed in **sections 4.2.2.2** and **4.2.5**.

## **4.2. Results**

### **4.2.1. PA size-dependence of DNA binding affinities ( $K_d$ )**

KA1039 with ODN-6-T3 (24 nt) has the weakest  $K_d$  of the PAs tested,  $2.4 \pm 0.4$  nM (**Table 4.2**). The  $K_d$  of KA1002 with ODN-6-T3 was  $0.17 \pm 0.1$  nM. These results are consistent with previously determined  $K_d$  using SPR (Wang, S. et al., 2012). KA1002 has also been characterized via our fluorescence assay using a much bigger DNA hairpin, ODN-8-T37 (44 nt), and a  $K_d$  of  $2.3 \pm 0.3$  nM was determined (Dupureur et al., 2012). We believe that the 10-fold weaker  $K_d$  determined for KA1002 with ODN-8-T37 is due to the larger size of the DNA hairpin acting as a bulk DNA competitor given that a  $K_d$  consistent with SPR was determined using the shorter ODN-6-T3.

The  $K_d$ s of NV1028 and NV1042 for their cognate DNA hairpins were also very tight (around 0.3 nM) (**Table 4.2**). These results were as expected since previous DNase I footprinting work determined that NV1028 and NV1042 bind to larger DNA fragments with  $K_d$  similar to each other (He et al., 2014; Vasilieva et al., 2016).  $K_d$ s for NV1028 and NV1042 determined via DNase I footprinting on a



**Table 4.2** Comparison of binding affinities determined via fluorescence assay vs other methods

PA	# of Rings	Direct $K_d$ (nM)	Comp $K_d$ (nM) <sup>f</sup>	Meth	HPV16 $IC_{50}^i$ ( $\mu$ M)	HPV16 $IC_{90}^i$ ( $\mu$ M)
KA1039	6	2.4±0.4 <sup>a</sup> 7.1±0.7 <sup>b</sup>	4.3±0.6	Fluor SPR	NOT ACTIVE	
KA1002	8	0.17±0.1 <sup>a</sup> 0.3±0.1 <sup>b</sup> 2.3±0.3 <sup>c</sup>	0.56±0.1	Fluor SPR Fluor	NOT ACTIVE	
NV1028	14	0.20±0.06 <sup>a</sup> 2.2±1 <sup>d</sup> 1.9±0.7 <sup>e</sup> 4.8±0.6 <sup>h</sup>	0.03±0.01	Fluor CE CE Fluor	0.100±0.020 (4)	1.113
NV1042	20	0.34±0.16 <sup>a</sup> 0.7±0.1 <sup>c</sup>	-- <sup>g</sup>	Fluor CE <sup>d</sup>	0.036±0.0004 (3)	0.351

<sup>a</sup> 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, this work. Unless otherwise specified,  $K_d$  were determined in triplicate and averaged and the error shown is relative error.

<sup>b</sup> 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% surfactant P20, pH 7.4. (Wang, S. et al., 2012)

<sup>c</sup> 10 mM HEPES, 50 mM NaCl, pH 7.4.  $K_d$  was determined using the much longer ODN-8-T37 (Dupureur et al., 2012)

<sup>d</sup> DNase I footprinting. 200 pM DNA, 10 mM CHAPS, 10 mM Tris pH 7.5-8.0, 2% DMSO, 5 mM  $Mg^{2+}$ , 5 mM  $Ca^{2+}$ , 10 mM  $K^+$  (He et al., 2014)

<sup>e</sup> DNase I footprinting. 200 pM DNA, 10 mM CHAPS, 10 mM Tris pH 7.5-8.0, 2% DMSO, 5 mM  $Mg^{2+}$ , 5 mM  $Ca^{2+}$ , 10 mM  $K^+$  (Vasilieva et al., 2016)

<sup>f</sup> Competition data was collected by Yang Song in triplicate, relative error is shown.

<sup>g</sup> Unable to determine. See **section 4.2.2.2**

<sup>h</sup> 10 Mm HEPES, 50 Mm NaCl, pH 7.4, 200  $\mu$ M nt calf thymus DNA (Dupureur et al., 2012).

<sup>i</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011).  $IC_{50}$  and  $IC_{90}$  are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

large viral fragment tended to be weaker than those determined by fluorescence spectroscopy. Additionally, a previously published  $K_d$  for NV1028 with ODN-14-T3 determined via our fluorescence assay was 24-fold weaker than the  $K_d$  determined here. This is likely due to the presence of bulk DNA added to the sample.

As discussed in **Chapter 3**, the  $K_d$ s reported here should be treated as apparent. With the exception of KA1039, the  $K_d$ s determined via our fluorescence assay were similar to each other despite their different antiviral properties. From this we can say that antiviral activity is not correlated to DNA binding strength.

#### **4.2.2. PA-DNA binding affinities by competition**

##### **4.2.2.1. Competition DNA binding affinities for KA1039 and NV1028**

A competition affinity assay described in **Chapter 2** was performed by Yang Song to determine if there is any dye-related interference. A summary of these data can be found in **Table 4.2**. With the exception of NV1028, competition  $K_d$  were generally slightly weaker than directly determined  $K_d$ s determined at 25 °C. The weaker  $K_d$ s are attributed to the greater error associated with the nature of the competition experiment. The competition  $K_d$  determined for NV1028 at 25 °C is an outlier relative to the competition  $K_d$ s determined at other temperatures for this system. See **Chapter 3**.

##### **4.2.2.2. Difficulties with NV1042-DNA competition experiment**

For NV1042 analyses two DNA sequences with the dye close to the

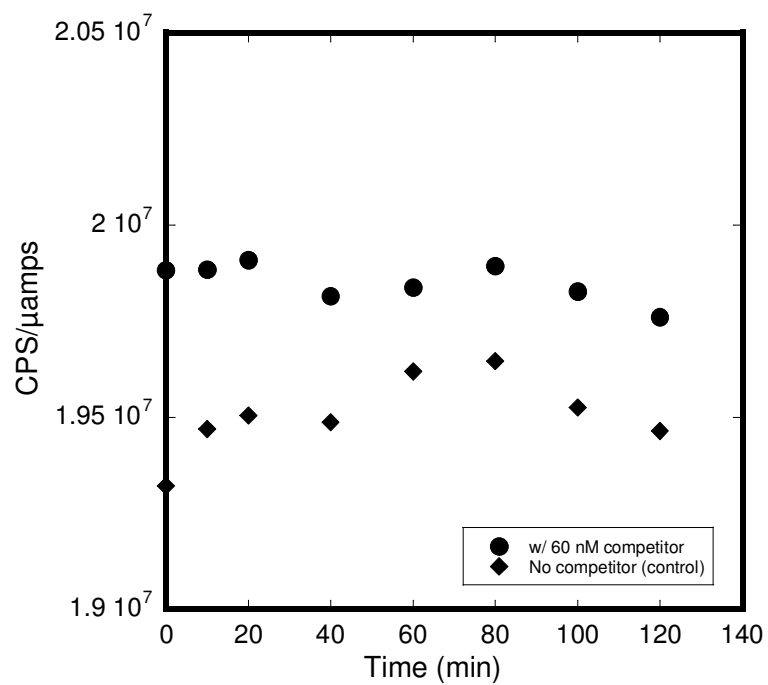
binding location of the PA N-terminus (ODN-20-T3) or the  $\gamma$ -turn (ODN-20-T16) were used. During attempts to determine the  $K_d$  of NV1042 with ODN-20-T3 (**Fig. 4.1**) and ODN-20-T16 (not shown) by competition, no recovery of intensity was observed for either oligonucleotide. Lack of recovery of intensity could be due to poor dynamic range and/or slow dissociation kinetics. Dynamic range, the observable change in dye intensity when PA binds, is directly correlated to sensitivity: a poor dynamic range means smaller incremental changes to the intensity when either PA or competitor is added. In competition experiments we typically do not see 100% signal recovery (~20% recovery observed with NV1028, discussed in **section 4.2.5**). Dye placement can have an impact on dynamic range (Dupureur et al., 2012).

While some recovery was observed for dissociation kinetics experiments using ODN-20-T16, it was not observed on a timescale appropriate for a competition titration. See **section 4.2.5**.

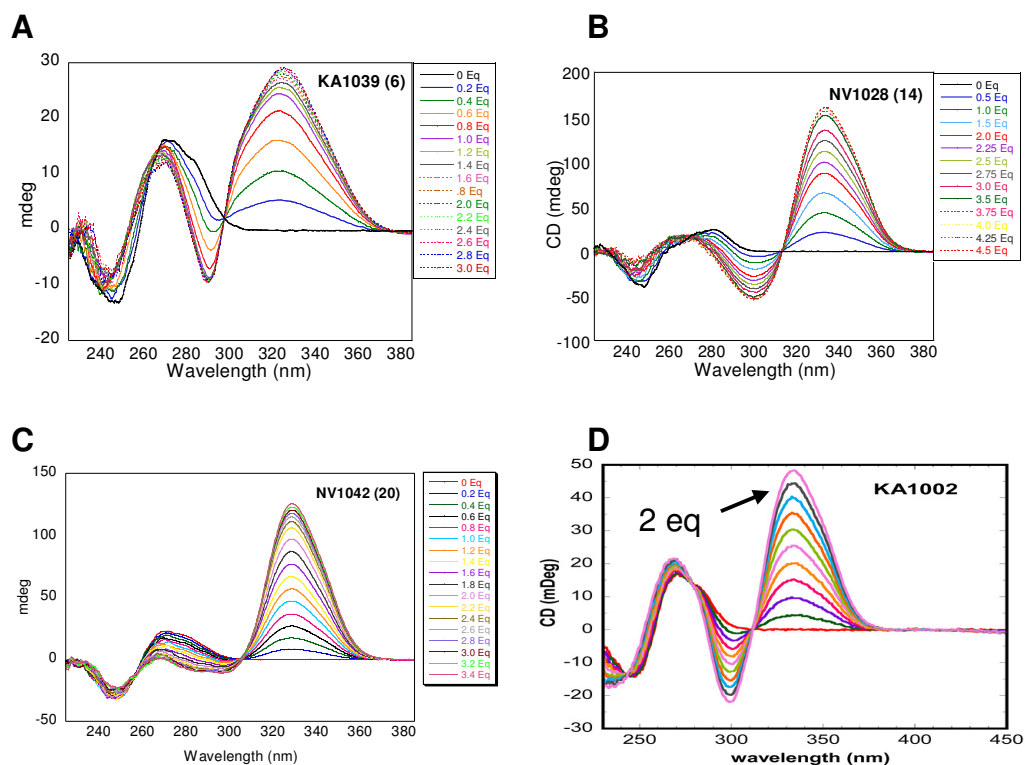
### **4.2.3. Circular dichroism spectroscopy**

#### **4.2.3.1. PA-DNA Size-dependent binding characterized via circular dichroism spectroscopy**

As discussed in **Chapter 3**, circular dichroism spectroscopy (CD) can be used to examine changes in DNA conformation in response to PA binding at  $\mu$ M concentrations (Lyng et al., 1992). CD spectra collected for all PAs discussed here are shown in **Fig 4.2**; CD spectra for NV1028 from **Chapter 3** and CD spectra for KA1002 collected by Wang and coworkers (Wang, S. et al., 2012) are



**Figure 4.1** NV1042 competition experiment with ODN-20-T3, collected by Yang Song. 60 nM DNA competitor was added to 3 nM NV1042:DNA complex in 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C.



**Figure 4.2** CD spectra collected by Kristin Bales for KA1039 with ODN-6 (A), NV1028 with ODN-14-LFTT (B), NV1042 with ODN-20 (C). PA equivalents ranged from 0-3.4. D are the CD spectra collected for KA1002 with ODN-6 adapted with permission from (Wang, S. et al., 2012). Copyright 2012 American Chemical Society. PA equivalents ranged from 0-2.0. Conditions for A-C: 5  $\mu$ M DNA, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 20  $^{\circ}$ C. Conditions for D: 5  $\mu$ M ODN-6, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 25  $^{\circ}$ C. CD spectra are averaged from 5 accumulations.

reproduced for clarity and convenience. For all PAs, PA-DNA binding was observed around 320-360 nm, which is typical for hairpin PAs binding in the minor groove of DNA. The peak for KA1039 is broader than the peaks observed with the other PAs, appearing at 300-360 nm. The broad appearance may be due to overlapping peaks—the peak observed at 320-360 nm, and second peak appearing at around 300 nm not observed in the spectra of the other PAs—which may indicate KA1039 interacts with DNA differently than the other hairpin PAs.

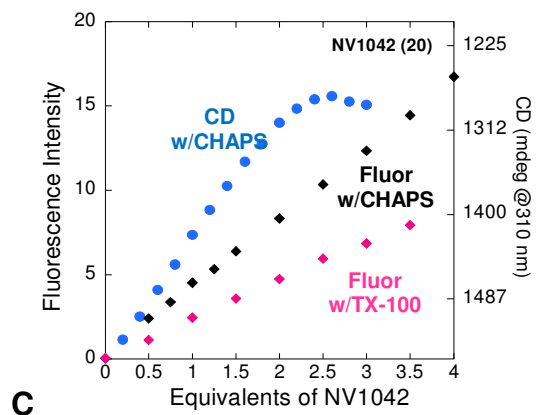
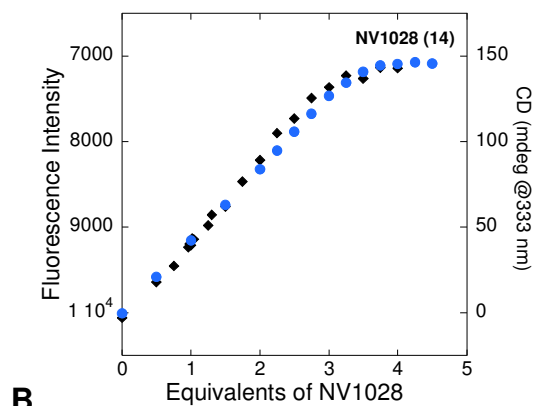
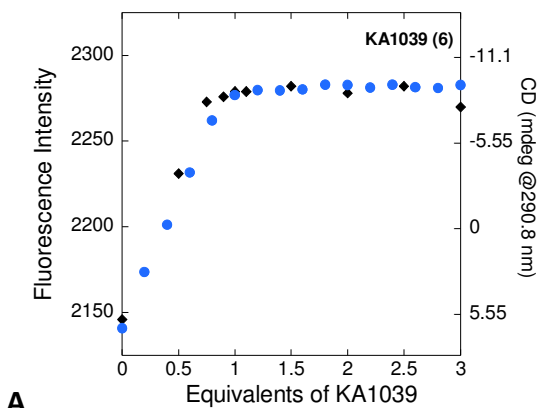
#### **4.2.3.2. PA-DNA size-dependent binding stoichiometries**

When possible, binding stoichiometries were determined by both CD ( $\mu\text{M}$ ) (**Fig. 4.2**) and fluorescence (nM) spectroscopy (**Fig. 4.3**) using the methods described in **Chapter 2**. PA-DNA binding stoichiometry of KA1039 with ODN-6 and ODN-6-T3 was determined to be 1:1 by both CD and fluorescence, respectively (**Fig. 4.3A**).

Binding stoichiometry of KA1002 was not determined via fluorescence, but it is clear from previously collected CD data that KA1002 binds to DNA at stoichiometries greater than 2:1 (Wang, S. et al., 2012), similar to what I reported for NV1028 in **Chapter 3** (**Fig. 4.3B**). NV1042 was also determined by CD spectroscopy only to bind to DNA at least 2:1 (**Fig. 4.3C**).

#### **4.2.3.3. NV1042 binding stoichiometry determination via fluorescence spectroscopy**

Several attempts were made to determine binding stoichiometry of NV1042 with both ODN-20-T3 and ODN-20-T16 by fluorescence



**Figure 4.3** Overlay of binding stoichiometry data collected via CD (●) and fluorescence spectroscopy (◆) for KA1039 (A), NV1028 (B), and NV1042 with CHAPS (◆) or TX-100 (◆) (C). Conditions for CD: 5  $\mu$ M DNA, 1X HNE, 1 mM CHAPS, pH 7.4, or 20  $^{\circ}$ C. Conditions for Fluor: 1X HNE, 1 mM CHAPS or 0.5 mM TX-100, pH 7.4, or 25  $^{\circ}$ C. CD data collected by Kristin Bales. CD: spectra were averaged from 5 accumulations, each point is the average mdeg of 2 spectra at a specific wavelength. Fluor: each point is the average of 3 intensities.

spectroscopy, but saturation was not observed at greater than 5 equivalents (**Fig. 4.3C**). Stoichiometry experiments were also attempted in Triton X-100 (TX100), a nonionic detergent, to determine if the ionic charges on CHAPS, a neutral zwitterionic detergent, may be interfering with the dye at the higher concentrations required for stoichiometry experiments, but saturation was not observed when using TX100 instead of CHAPS. Given that CD requires concentrations to be 1000-fold higher than those used for fluorescence, the issue may have been related to slow kinetics or poor binding thermodynamics in the nM range.

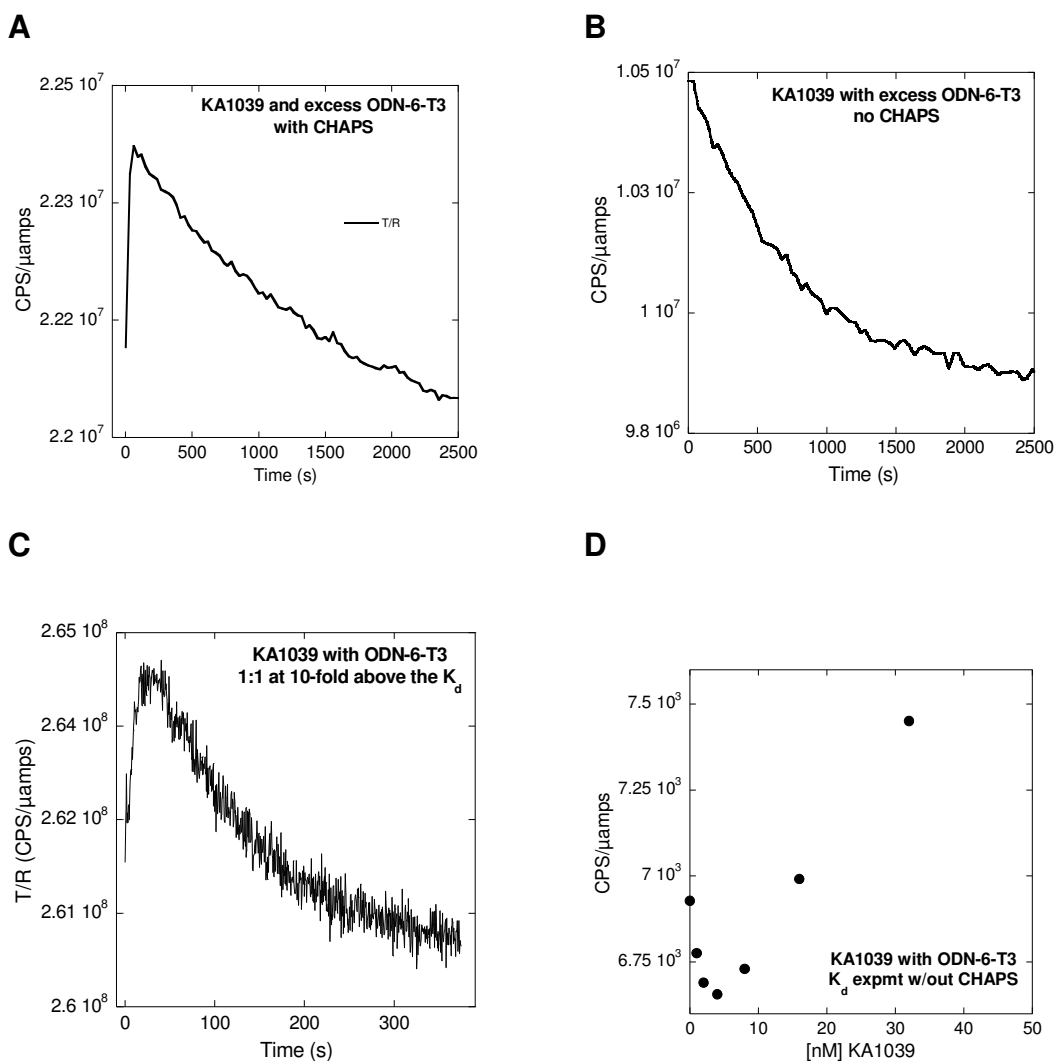
#### **4.2.4. Size-dependent PA-DNA association kinetics**

##### **4.2.4.1. KA1039 and ODN-6-T3 CHAPS-dependent dye behavior**

Unexpectedly, despite a binding stoichiometry of 1:1 for KA1039 with ODN-6-T3, the association traces resembled those of NV1028 with ODN-14-T6, where 2 distinct phases were observed (**Fig. 4.4A**) even when PA and DNA were at 1:1 (**Fig. 4.4C**). However, the observance of 2 distinct phases is not inconsistent with 1:1 binding stoichiometry given that KA1039 may bind via multiple binding modes in a concentration dependent manner.

Additionally, we determined that the dye behavior of ODN-6-T3 with KA1039 was CHAPS-dependent: with CHAPS the dye intensity increases, plateaus, then decreases upon addition of KA1039; without CHAPS the dye intensity decreases upon addition of KA1039 (**Fig. 4.4B**). Additionally, when trying to determine the  $K_d$  via titration experiments, without CHAPS, a plateau





**Figure 4.4** Comparison of KA1039 and ODN-6-T3 CHAPS-dependent dye behavior. Association curves with (A), and without CHAPS (B). Conditions: 1 nM ODN-6-T3 (w/o CHAPS) or 2.5 nM ODN-6-T3 (w/1 mM CHAPS), 5 nM KA1039, 1X HNE, pH 7.4, 25 °C. C, KA1039-DNA association curve at 30 nM PA-DNA complex. D, Attempt to determine binding affinity without CHAPS. 1 nM ODN-6-T3, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 25 °C. Points represent the average of 3 intensities taken at after the addition of each aliquot of KA1039.

region is not observed (**Fig. 4.4D**), as a sharp change in dye intensity in the opposite direction is observed mid-titration; thus, without CHAPS we are unable to determine a  $K_d$ . Consequently, CHAPS is still required for experiments since the dye behavior is different without CHAPS and therefore data is not comparable between the two methods if CHAPS is omitted.

#### **4.2.4.2. Attempts at stopped-flow characterization of the fast phase of KA1039**

Under the assumption that KA1039 fast phase  $k_{on}$  is concentration dependent, attempts to slow the association experiments to an observable timescale were made by lowering the concentrations of ODN-6-T3 and/or KA1039. However, signal-to-noise ratios became poorer as the concentration of ODN-6-T3 was lowered and I was unable to find a concentration at which I could characterize the fast phase  $k_{on}$  of KA1039 using the steady-state instrument. See Yang Song's thesis for further discussion about dye sensitivities for each PA-DNA system.

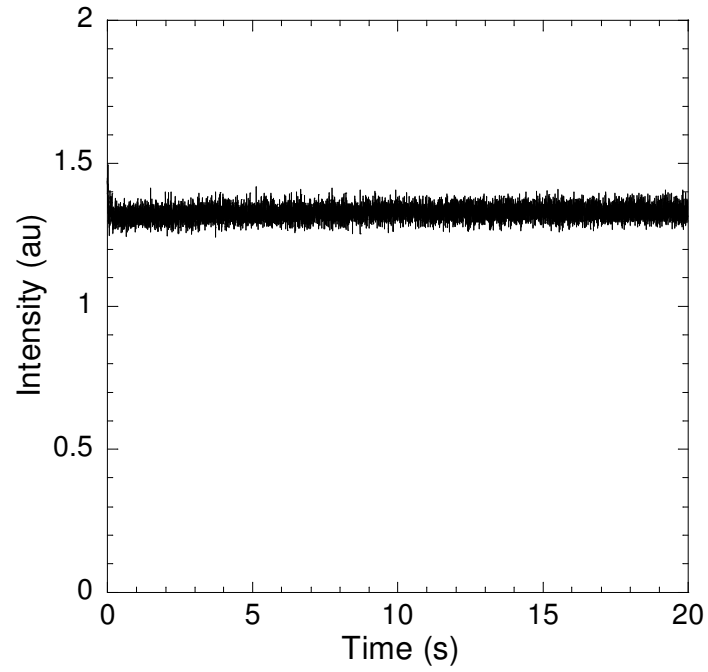
Attempts were made to characterize the fast phase DNA association event of KA1039 with ODN-6-T3 by stopped-flow kinetics using the method described in **Chapter 2**. Initial attempts were performed with KA1039 stock dissolved in DMSO. In small volumes, DMSO will mix homogeneously with water (0.5% v/v or less is used on our steady state instrument). However, the stopped-flow instrument is limited to a minimum of 5% v/v which made the use of DMSO stocks infeasible for stopped-flow experiments.

KA1039 is not known to aggregate at concentrations appropriate for isothermal calorimetry ( $\mu\text{M}$ ) (Wang, S. et al., 2012); thus, nM stocks can be made in water. Attempts to characterize the KA1039 fast phase association with KA1039 stocks made in water were still unsuccessful and no signal change was observed, most likely due to sensitivity issues. Experiments were performed at 5 °C to slow down the reaction but still no change in signal was observed via stopped-flow (**Fig. 4.5**). The fast phase  $k_{\text{on}}$  was estimated as at least  $10 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$  using the concentration range used on the steady-state instrument under the assumption that the fast phase completed within the time it takes the shutters to open and collect the initial point (2 s).

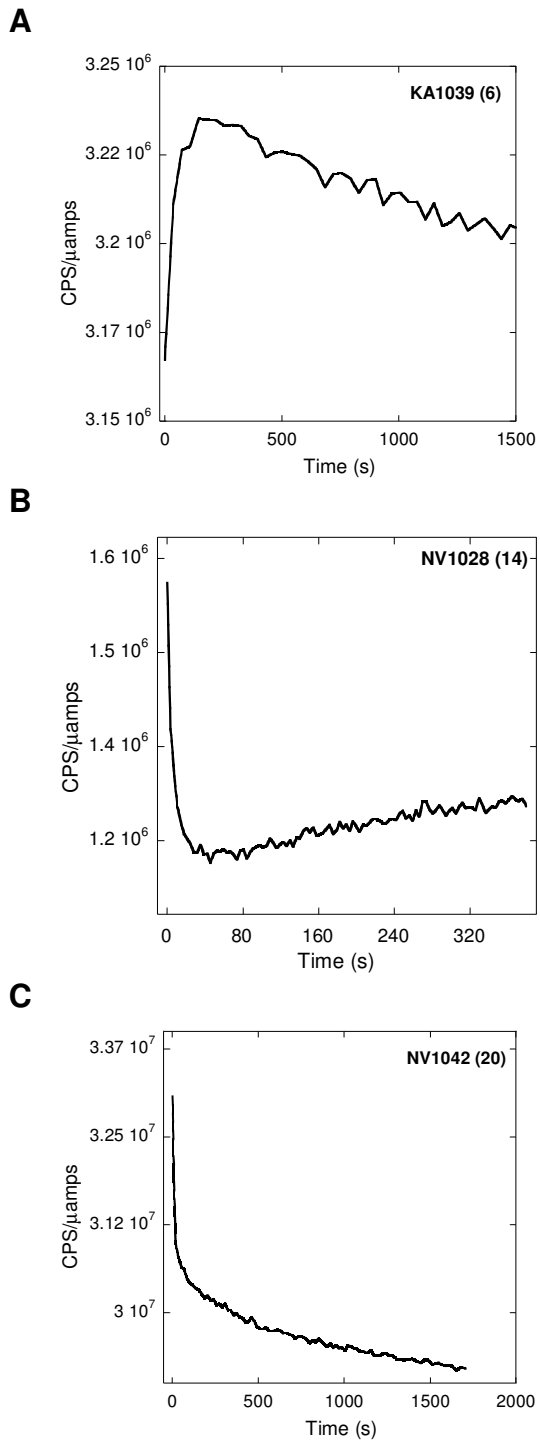
#### **4.2.4.3. Size-dependent PA-DNA association kinetic rate constants**

DNA association curves were collected for KA1039, NV1028, and NV1042 (**Fig. 4.6**). Linear concentration dependence was observed for the second phase of KA1039 and the DNA association rate constant ( $k_{\text{on}}$ ) was determined as described in **Chapters 2 and 3**. (**Fig. 4.7A, Table 4.3**).  $k_{\text{on}}$  for the slow phase DNA association event was determined to be  $0.025 \pm 0.04 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ , which is similar to the  $k_{\text{on}}$  value for the slow phase of NV1028 and discussed in **Chapter 3**.

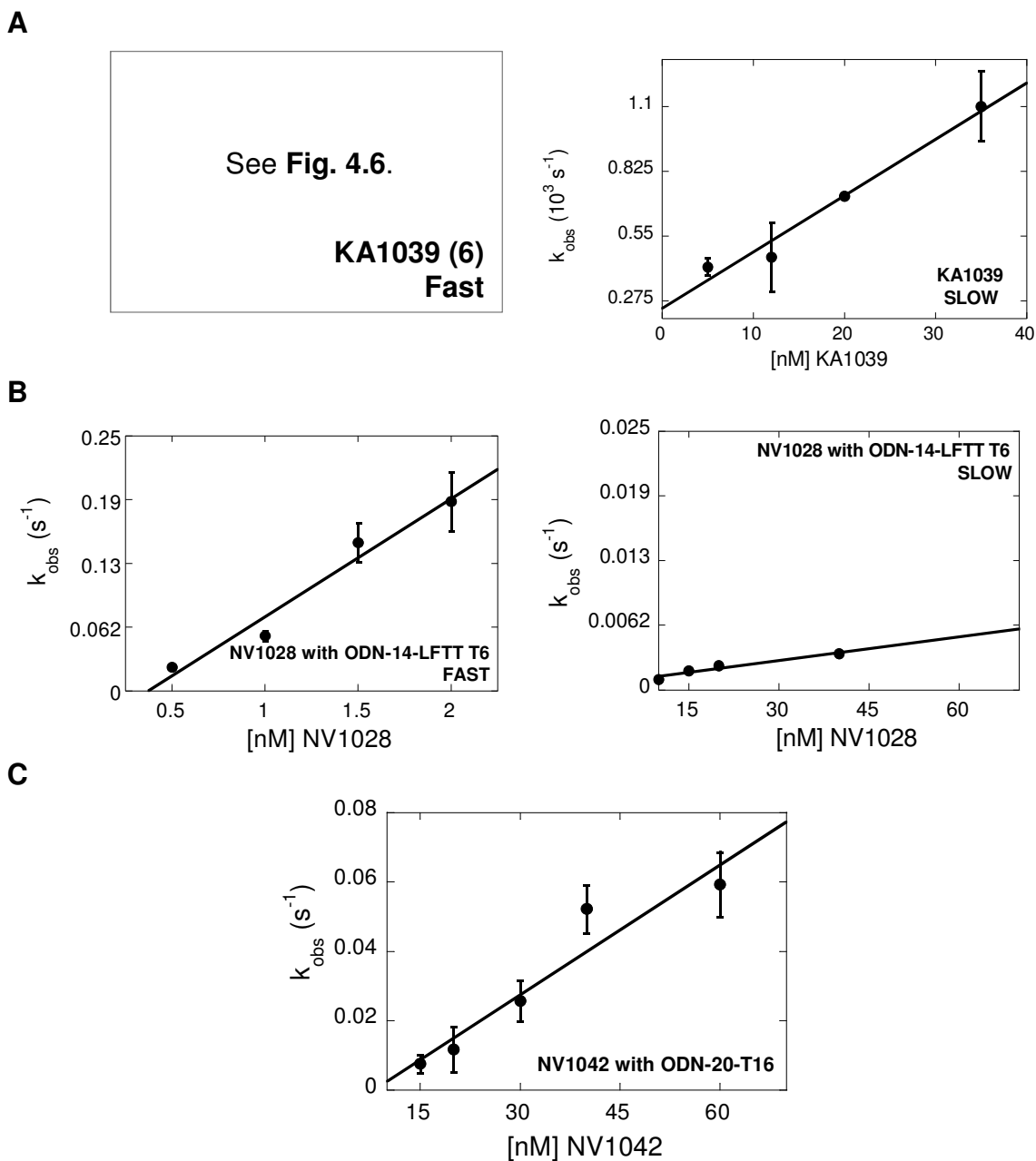
Unlike for KA1039 and NV1028, association traces for NV1042 with ODN-20-T16 did not change direction, but do appear biphasic (**Fig. 4.6C**). When fit to a monoexponential equation (**eqn. 2.3**), a linear concentration dependence was observed for NV1042  $k_{\text{obs}}$  values with respect to a single PA equivalent that was



**Figure 4.5** Example of stopped-flow data for KA1039 with ODN-6-T3 at maximum flow rate. Conditions: 1 nM ODN-6-T3, 5 nM KA1039, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 5 °C. Flow rate: 10 mL/s; dead time: 2 ms.



**Figure 4.6** Representative examples of association kinetics curves where PA is in excess of DNA for KA1039 with ODN-6-T3 (A), NV1028 with ODN-14-LFTT T6 (B), and NV1042 with ODN-20-T16 (C). Conditions: 1X HNE, 1 mM CHAPS, pH 7.4, 25 °C.



**Figure 4.7** Secondary association kinetics plots for KA1039 with ODN-6-T3 (A), NV1028 with ODN-14-LFTT T6 (B), and NV1042 with ODN-20-T16 (C). Conditions: 1X HNE, 1 mM CHAPS, pH 7.4, 25 °C. NV1028 (fast phase) and NV1042 association curves and secondary linear fits were determined in duplicate, relative error is shown. KA1039 and NV1028 (slow phase) association curves and secondary linear fits were determined once, fit error for  $k_{obs}$  is shown.

**Table 4.3** DNA association rate constants ( $k_{on}$ ) for PAs by size<sup>a</sup>

PA	# of rings	$k_{on\ fast}$ ( $10^6\ M^{-1}\ s^{-1}$ )	$k_{on\ slow}$ ( $10^6\ M^{-1}\ s^{-1}$ )	Ref	HPV16 $IC_{50}^d$ ( $\mu M$ )	HPV16 $IC_{90}^d$ ( $\mu M$ )
KA1039 <sup>e</sup>	6	>10 <sup>c</sup> 3±1.5 <sup>b</sup> 7.1±0.8 <sup>b</sup>	0.025±0.004 None reported None reported	This work (Wang, S. et al., 2012) (Wang, S. et al., 2013)	NOT ACTIVE	
KA1002	8	44±15 <sup>b</sup> 53 <sup>b</sup>		(Wang, S. et al., 2012) (Bashkin, J. K. et al., 2013)	NOT ACTIVE	
NV1028	14	120±20	0.08±0.01	This work	0.100±0.020 (4)	1.113
NV1042	20	1.4±0.2	None observed	This work	0.036±0.0004 (3)	0.351

<sup>a</sup> Fluorescence: 1X HNE, 1 mM CHAPS, pH 7., 25 °C. NV1028 (fast phase) and NV1042 association curves and secondary linear fits were determined in duplicate, relative error is shown. KA1039 and NV1028 (slow phase) association curves and secondary linear fits were determined once, fit error is shown

<sup>b</sup> SPR: 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% Surfactant P20 (pH 7.4)

<sup>c</sup> Not determined due to sensitivity issues. Estimated using the concentrations used on steady-state instrument under the assumption that the fast phase completed within the time it takes the shutter to open (2 s).

<sup>d</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011)  $IC_{50}$  and  $IC_{90}$  are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

<sup>e</sup> All characterizations of KA1039 were performed using ODNs with the same DNA sequence, ODN-6. ODN used for SPR characterizations were biotinylated at 5'; ODN used for fluorescence characterizations were labeled at T3 (**Table 3**).

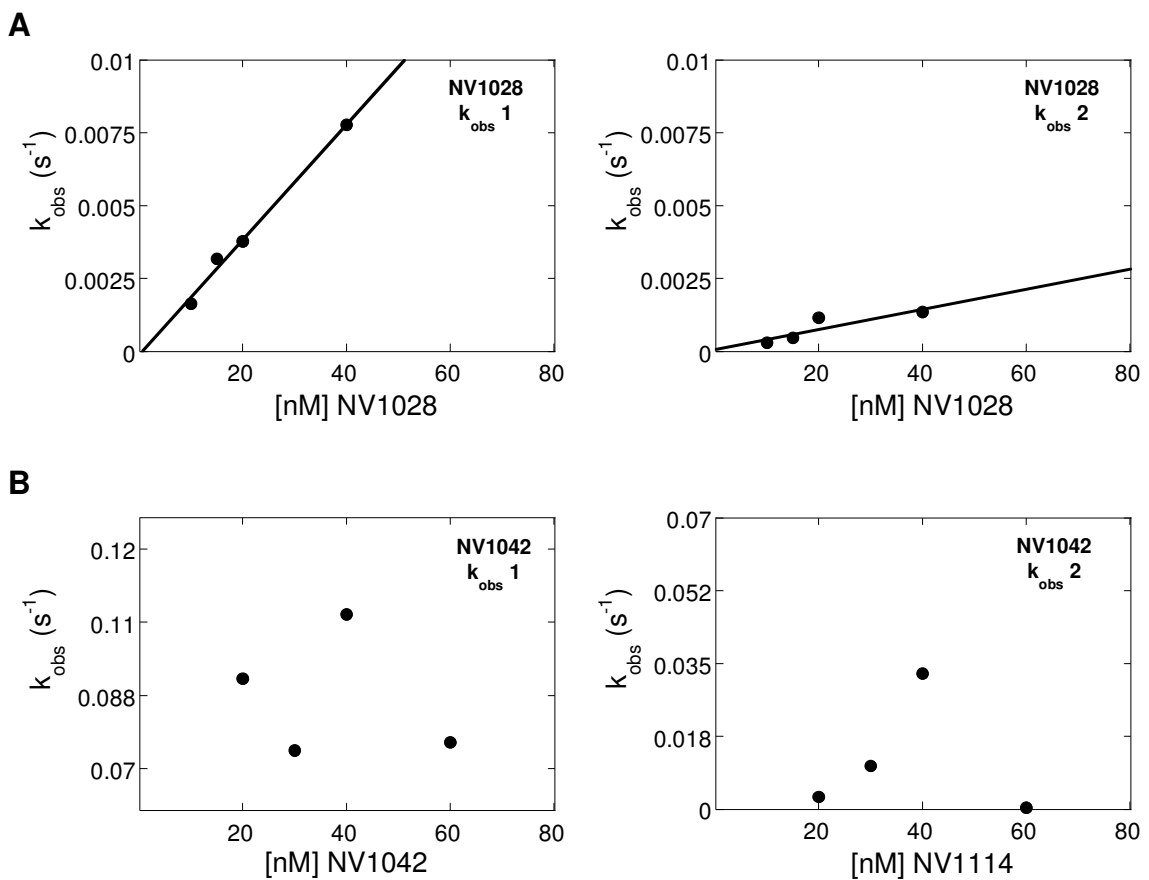
<sup>f</sup> Characterization of KA1002 were performed with 5' biotinylated ODN of different sequence. Characterizations by Wang were performed using ODN-6, characterizations by Bashkin were performed using ODN-8 (**Table 4.1**).

slower than the  $k_{on}$  values of the fast phases of KA1039 and NV1028, but faster than the slow  $k_{on}$  values of the slow phases of KA1039 and NV1028 (**Fig. 4.7C**, **Table 4.3**).

#### 4.2.4.4. NV1042-DNA association rate constants (biexponential fitting)

Given the non-monophasic appearance of the NV1042 association curves and binding stoichiometry determined to be at least 2:1 via CD spectroscopy, alternate fittings were performed with a biexponential equation (**eqn. 3.1**), like those performed for NV1028 in **Chapter 3** (reproduced in **Fig. 4.8A**). As expected, when fit to a biexponential equation (**eqn. 3.1**), the additional variables of the biexponential equation generally made fits of the curves ( $R^2$  values) better than the fits using a monoexponential equation. Unlike NV1028, 2 distinct lines were not observed for NV1042 when  $k_{obs\ 1}$  and  $k_{obs\ 2}$  were plotted with respect to NV1042 concentration: when plotted as a function of NV1042 concentration,  $k_{obs\ 1}$  does not seem to be concentration dependent; when plotted as a function of NV1042, three of the four  $k_{obs\ 2}$  values may form a line, **Fig. 4.8B**. However, it is unclear if the outlier is the  $k_{obs\ 2}$  value at 40 nM or 60 nM NV1042, as removal of either value yields a plausible line. Thus, it is unclear whether linearity was observed for either  $k_{obs\ 1}$  or  $k_{obs\ 2}$ . Therefore, despite the non-monophasic appearance, the data indicates the NV1042 association curves correlate best to a single binding event. It may be that additional binding events occur at a much slower rate, or that their influence on the dye is harder to observe. The biphasic appearance is likely due to an artifact or may be related to a dye response for an





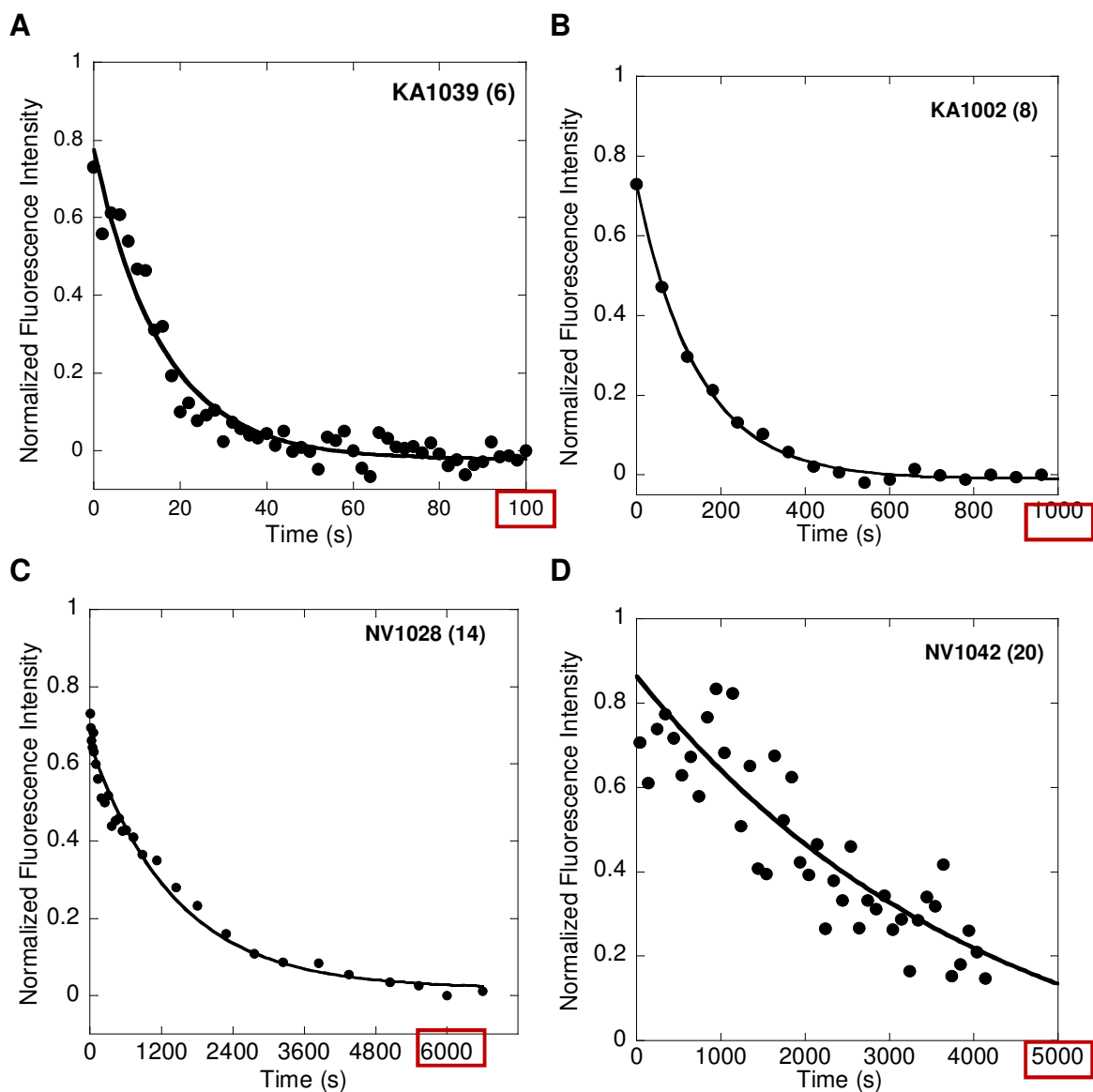
**Figure 4.8** Representative secondary plots of NV1028 and NV1042 association curves when fit to a biexponential equation. A, NV1028 with ODN-LFTT T6 slow phase; B, NV1042 with ODN-20-T16.

event we are unable to characterize at this timescale or under our experimental conditions. Simulated association traces using a monoexponential equation with a linear component are inconsistent in shape with the traces collected for NV1042 (not shown).

#### 4.2.5. PA-DNA dissociation rate constants as a function of PA size

PA-DNA dissociation behavior was observed by monitoring the change in fluorescence of a 1:1 PA-DNA complex when an excess of unlabeled DNA is added (**Fig. 4.9**). In general, dissociation experiments for the larger antivirals required a longer time frame for completion. For the inactive PAs, KA1039 and KA1002, dissociation experiments were completed in less than 10 min. NV1028 required almost 2 hours to complete. Dissociation experiments for NV1042, the most potent of the PAs characterized here, did not finish within a 2-hour timeframe.

Macroscopic apparent dissociation rate constants ( $K_{\text{off}}$ ) were determined using the method described in **Chapter 2**. Because NV1042 experiments did not finish in the experimental timeframe and because the data quality was poor, the  $K_{\text{off}}$  for NV1042 is treated as an estimate.  $K_{\text{off}}$  values ranged from  $60 \times 10^{-3} \text{ s}^{-1}$  to  $0.40 \times 10^{-3} \text{ s}^{-1}$  with residence times ( $1/K_{\text{off}}$ ) of 20 s to >2,500 s, respectively (**Table 4.4**). The  $K_{\text{off}}$  of KA1002,  $7.0 \times 10^{-3} \text{ s}^{-1}$ , determined previously (Dupureur et al., 2012), fell within the range expected for an 8-ring PA when compared to  $K_{\text{off}}$  data for PA of other sizes presented here. The  $K_{\text{off}}$  values for KA1039 and KA1002 determined using the fluorescence method were slightly faster than



**Figure 4.9** Representative dissociation rate constant plots. A, KA1039; B, KA1002 replicated from (Dupureur et al., 2012); C, NV1028; D, NV1042. All experiments were performed in 1X HNE and 1 mM CHAPS pH 7.4, 1:20 DNA-PA complex:competitor. [DNA-PA] complex was at least 10-fold the directly determined  $K_d$ .

**Table 4.4** Comparison of DNA dissociation rate constants determined via fluorescence assay vs other methods for a 1:1 PA-DNA complex

PA	# of rings	$K_{off}$ ( $10^{-3} s^{-1}$ )	Res Time (s) <sup>e</sup>	% Recov	Meth	HPV16 $IC_{50}^g$ ( $\mu M$ )	HPV16 $IC_{90}^g$ ( $\mu M$ )
KA1039	6	67.2±7.4 21±5	20 48±10	120 <sup>f</sup>	Fluor <sup>a</sup> SPR <sup>b</sup>	NOT ACTIVE	
KA1002	8	7.0 12±2 59	150 83±15 17	167 <sup>f</sup>	Fluor <sup>b</sup> SPR <sup>c</sup> SPR <sup>d</sup>	NOT ACTIVE	
NV1028	14	0.6±0.07	700	24	Fluor <sup>a</sup>	0.100±0.020 (4)	1.113
NV1042	20	<0.40	>2500	28	Fluor <sup>a</sup>	0.036±0.0004 (3)	0.351

<sup>a</sup> 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, this work. NV1028 and KA1039 dissociation rate constants were determined in duplicate with ODN-14-LFTT T6, and ODN-6-T3, respectively, relative error is shown. NV1042 dissociation rate constants were determined once with ODN-20-T16, fit error shown.

<sup>b</sup> 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4 (Dupureur et al., 2012)

<sup>c</sup> 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% surfactant P20 (pH 7.4) (Wang, S. et al., 2012)

<sup>d</sup> 10 mM HEPES, 150 mM NaCl, 3 mM EDTA, 0.05% surfactant P20 (pH 7.4) (Bashkin, J. K. et al., 2013)

<sup>e</sup> Residence time =  $1/K_{off}$

<sup>f</sup> determined from competition experiments by Yang Song

<sup>g</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011).  $IC_{50}$  and  $IC_{90}$  are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

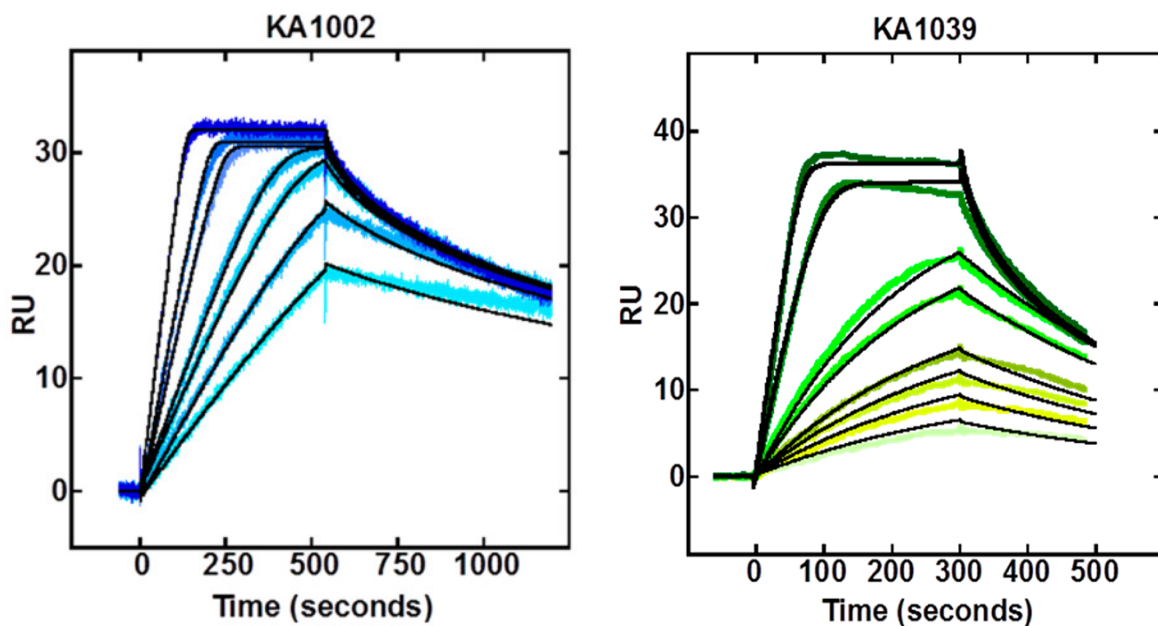
those determined by SPR by less than 3-fold which could be attributed to the difference in experimental conditions (Wang, S. et al., 2012)

Fluorescence recovery (the percentage of fluorescence intensity return due to PA dissociation from DNA) also tended to be lower for the larger antiviral PAs whereas 100% recovery was generally observed for KA1039 and KA1002 (**Table 4.4**). Given that multiple equivalents bind to the DNA even at 1:1, we believe this lack of recovery may be due to equivalents of PA not dissociating from the DNA during the experimental timeframe.

### **4.3. Discussion**

#### **4.3.1. Comparison of KA1039 DNA binding kinetics data collected by fluorescence spectroscopy and surface plasmon resonance (SPR)**

DNA binding stoichiometry is 1:1 for KA1039 and ODN-6 by both CD and fluorescence spectroscopy. However, despite this, we observe 2 distinct association phases. KA1039 DNA binding kinetics have been previously characterized by SPR but biphasic binding kinetics were not reported (Wang, S. et al., 2013; Wang, S. et al., 2012). It is unclear if this was because biphasic binding kinetics were not observed, or if the results may have been misinterpreted. Further scrutiny of KA1039 sensorgrams published reveals that the wash sequence begins around 300 s (5 mins) and prior to a plateau in SPR signal (**Fig. 4.10**). Additionally, SPR sensorgrams of KA1039 at higher concentrations have plateau regions that do not fit well when compared to the sensorgrams of other PAs, such as KA1002. It is also possible that further



**Figure 4.10** SPR sensorgrams. Left, KA1002 with ODN-6-T3 shown as an example of an SPR sensorgram where association curves have a plateau region that fits well, right, KA1039 with biotinylated version of ODN-6 where the plateau region does not fit well. Adapted with permission from (Wang, S. et al., 2012). Copyright 2012 American Chemical Society.

scrutiny of the data was not pursued because the fits obtained fit traces reasonably well without adding additional undefined fit parameters.

Low sensitivity made characterization of the fast phase association rate constant for KA1039 with ODN-6-T3 infeasible. The slow phase association rate constant was determined to be  $2.5 \times 10^4 \text{ M}^{-1} \text{ s}^{-1}$  which is 100-fold slower than the  $k_{\text{on}}$  determined by SPR ( $3 \pm 1.5 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ ). The difference in association rate constants may be due to a few possibilities. One could be the difference in experimental conditions which may affect the binding kinetics. A second possibility is that, given that only a single binding phase was reported for KA1039 for SPR experiments, the SPR determined association rate constant may correspond to the fast phase association event. However, based on my experience in collecting association rate constants with the other PAs, I estimate the lower limit to be at least  $1 \times 10^7 \text{ M}^{-1} \text{ s}^{-1}$ , which would be faster than the  $k_{\text{on}}$  of KA1039 determined via SPR.

A third possibility is that the association rate constant determined by SPR is a macroscopic  $K_{\text{on}}$  that correlates to both phases. But, with the estimated lower limit and slow phase  $k_{\text{on}}$  determined via fluorescence spectroscopy, the macroscopic  $K_{\text{on}}$  would be at least 10-fold slower than the SPR-determined association rate constant. However, if we assume the fast phase to be diffusion-controlled, like NV1028, then a  $K_{\text{on}}$  of  $1.6 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$  is computed, similar to the  $k_{\text{on}}$  determined via SPR for KA1039.

Dissociation kinetic rate constants were similar for KA1039 when determined by SPR or by fluorescence spectroscopy. I determined the  $K_{\text{off}}$  of

KA1039 from ODN-6-T3 to be  $60 \times 10^{-3} \text{ s}^{-1}$  which is close to  $k_{\text{off}}$  determined via SPR ( $21 \pm 5 \times 10^{-3} \text{ s}^{-1}$ ) (**Table 4.4**).

#### **4.3.2. Internal consistency of DNA binding kinetic rate constants and PA size**

Reconciliation of NV1028 binding kinetics and binding affinity data was accomplished in **Chapters 3** by taking the square root of the product of the two concentration dependent  $k_{\text{on}}$  values (**eqn. 3.2**) and comparing the computed values to the values determined experimentally. A similar analysis of KA1039 DNA binding kinetics rate constants and DNA binding affinities was performed. However, given that I was unable to characterize the fast phase of KA1039 I was unable to compute a macroscopic  $K_{\text{on}}$  using an experimentally determined fast phase  $k_{\text{on}}$  value. Instead,  $K_{\text{on}}$  was computed using the estimated lower limit of the fast phase DNA  $k_{\text{on}}$ . The computed  $K_{\text{off}}$  was 20-fold lower than the experimentally determined  $K_{\text{off}}$  value (**Table 4.5**).

Using the NV1042 experimental values determined via monoexponential fitting, data are internally consistent. This means that the binding events observed and characterized via the different experimental methods are likely the same and/or related events. Also, it means our method is able to be used to accurately characterize these PA-DNA binding events kinetically and at equilibrium.

It should be noted that given that the fluorescence intensity recovery upon addition of an unlabeled competitor for NV1028-DNA and NV1042-DNA complexes is typically less than 30% for both PAs, internal consistency among



**Table 4.5** Comparison of experimental and computed PA-DNA binding constants for KA1039, NV1028, and NV1042

PA	Direct $K_d$ (nM)	$k_{on\ fast}$ ( $10^6\ M^{-1}\ s^{-1}$ )	$k_{on\ slow}$ ( $10^6\ M^{-1}\ s^{-1}$ )	Calculated $K_{on}^a$ ( $10^6\ M^{-1}\ s^{-1}$ )	$K_{off}$ ( $10^{-3}\ s^{-1}$ )	Calculated $K_{off}^b$ ( $10^{-3}\ s^{-1}$ )
KA1039	2.4±0.4	>10 <sup>c</sup>	0.025±0.004	0.5	67.2±7.4	1.2
NV1028	0.23±0.02	120±19	0.077±0.01	3.0	0.6±0.07	0.69
NV1042	0.34±0.16	1.4±0.2		--	≤0.4	0.5

<sup>a</sup> computed using  $K_{on} = \sqrt{k_{on\ fast} * k_{on\ slow}}$  (Lacy et al., 2002)

<sup>b</sup> computed using  $k_{off} = K_{on} * K_d$

<sup>c</sup> estimated lower limit

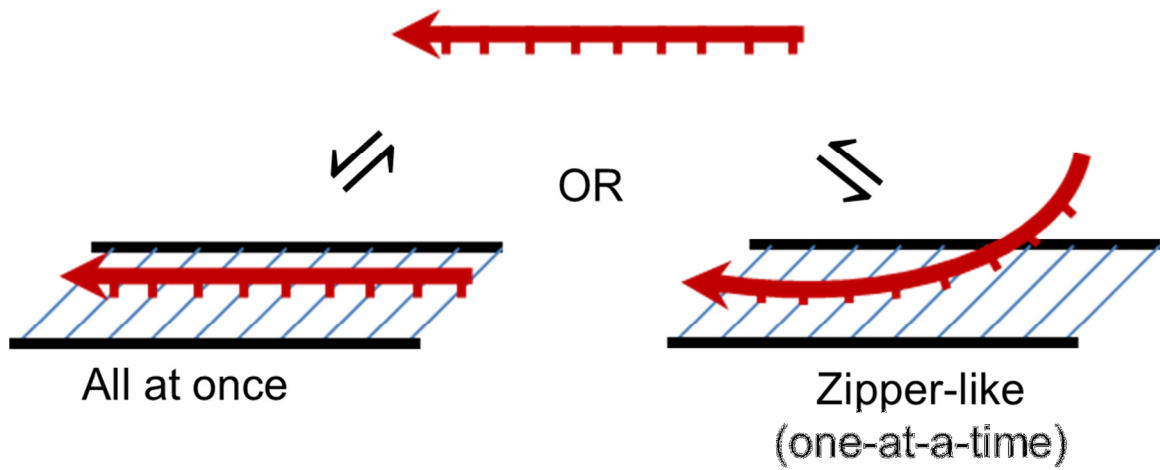
the parameters may be coincidental. As of this writing, we are unable to determine microscopic binding affinities and binding kinetic rate constants for all the binding events observed in stoichiometry experiments.

#### **4.3.3. PA-DNA fast phase association rate constant was slowest for NV1042**

Based only on the data able to be collected here, there is an inverse correlation of PA-DNA association rate constants with PA size. As there is a 100-fold difference between the fast phase  $k_{on}$  of NV1028 and the  $k_{on}$  determined via the monoexponential fitting of NV1042.

If we assume that the residues of a single hairpin PA do not bind to DNA simultaneously (all-at-once), and that nucleotide binding happens in a sequential manner (zipper-like or one-at-a-time), then a PA that makes fewer bonds would associate with DNA faster than a PA that makes more bonds (**Fig. 4.11**). If we assume the bonds are made in a zipper-like action, this correlation would be consistent given that larger hairpins are capable of making more bonds which may require more time (i.e. a slower).

If PAs bind to DNA in an all-at-once mechanism, then the slower association rate may be related to sequence recognition. While we generally believe that Dervan's rules do not adequately describe sequence specificity of large hairpin PAs, we acknowledge that some type of selection process is involved given that NV1028 and NV1042 do not bind DNA sequences indiscriminately, but rather have replicable footprinting and affinity cleavage patterns (see **Chapter 3**). The additional rings and larger size may still be more



**Figure 4.11** Comparison of all-at-once binding versus zipper-like binding. PAs are shown as red arrows.

specific—just not in the way we currently understand PA-DNA sequence specificity—and may require more time to find their recognition sequence.

There is also the possibility that the slower association rates of the larger PAs are due to a combination of a zipper-like action and site scanning. DNA-binding proteins are capable of scanning for and finding specific sequences of DNA via facilitated diffusion (sliding and hopping): 1D diffusion (sliding) is when the ligand slides along the DNA without detaching; 3D diffusion (hopping) is when the ligand detaches and moves to a different site to sample elsewhere along the DNA (Halford & Marko, 2004). Given that Dst is capable of binding to DNA via different modes in a sequence-dependent manner, larger hairpin PAs may be capable of doing the same. The larger surface area required for larger PAs to bind to DNA may require more scanning time in order to find a favorable binding site.

It should be noted that there is no overall size-dependent trend observed for PA-DNA association kinetics data collected via SPR. As shown in **Table 1.2**, PA-DNA binding kinetics studies tend to focus on smaller hairpin PAs (6-10 rings) and association rate constants can vary by orders of magnitude even among PAs of the same size. Of the larger hairpin PAs studied (12 and 14 rings) a slower DNA association rate constant was determined for the 12-ring hairpin PA than the 14-ring hairpin PA. Clearly, more studies are required to determine if PA-DNA association rate constants are, in fact, size dependent.

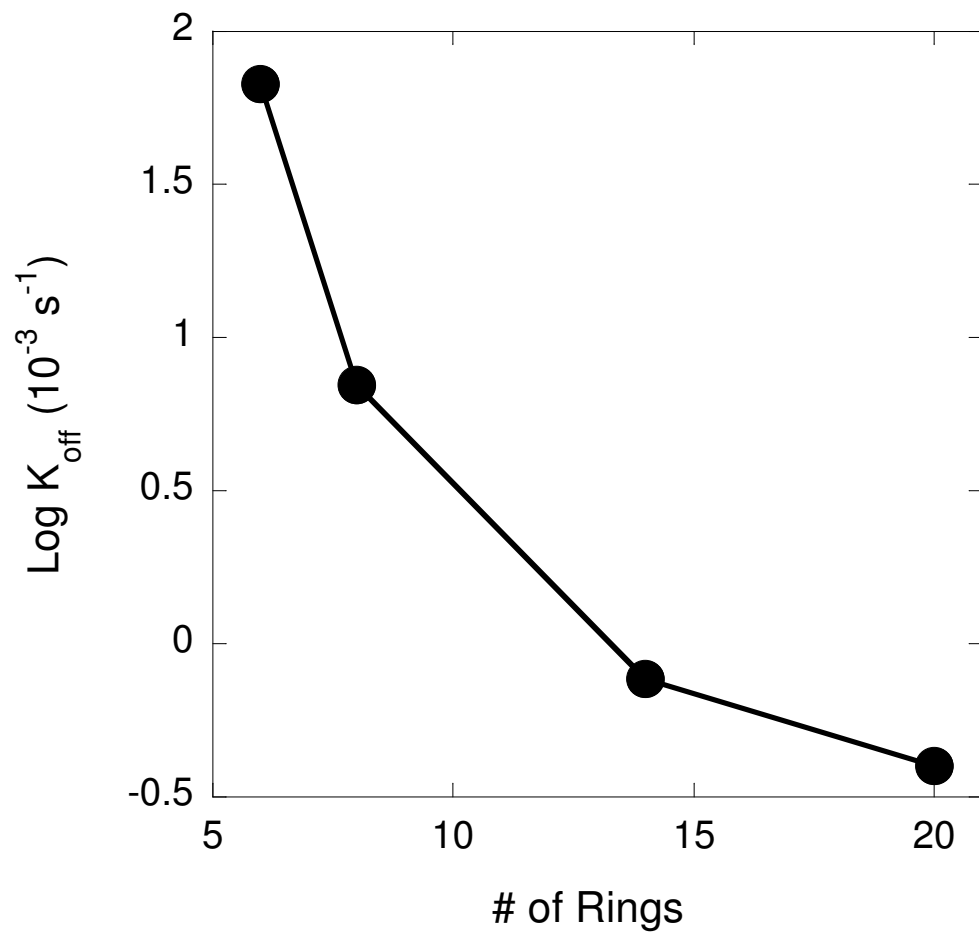
#### 4.3.4. Macroscopic DNA $K_{off}$ correlates with PA size

There is a correlation between dissociation rate constants and PA size: dissociation rates slowed as the number of rings increased (**Fig. 4.12**). The differences in  $K_{off}$  are so large that the data is better represented in log scale. This data also indicates there is a possible correlation between dissociation rate constants and antiviral activity. Further work with larger hairpin PAs that exhibit lower or no potency would be helpful in establishing to a correlation.

#### 4.3.5. Implications for the large hairpin PA antiviral mechanism

DNA binding stoichiometric ratios were determined for all PAs by fluorescence (nM) and CD ( $\mu$ M) spectroscopy. Interestingly, 1:1 binding stoichiometry was only observed for KA1039 and ODN-6-T3 despite the observation of 2 distinct phases even at 1:1 (**Fig. 4.4C**). Greater than 2:1 binding stoichiometry is observed for all larger PAs studied. Given that KA1002 does not have antiviral properties, it is unclear as to whether or not the binding stoichiometry is related to the antiviral mechanism.

As discussed in **Chapter 3**, a bridging interaction may explain the 2.5:1 binding stoichiometry observed via CD and fluorescence spectroscopy. AFM data collected by Qiao and coworkers of an 8-ring hairpin PA binding to linear calf thymus DNA shows a kinking or contracting of the DNA may be occurring upon PA binding to DNA (Qiao et al., 2015). Additionally, as discussed in **Chapter 1**, it has been shown that the antiviral mechanism of NV1028 is related to the activation of the DNA Damage Response system (DDR) in a manner not

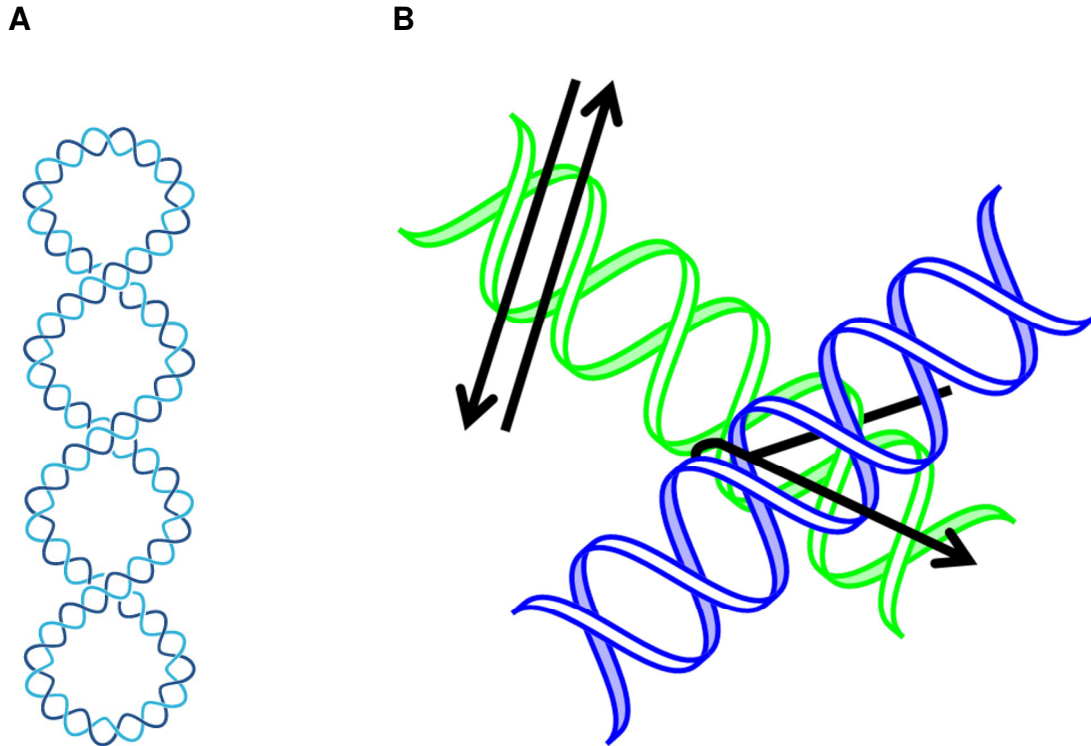


**Figure 4.12** PA size dependence of DNA dissociation rate constants.

conducive to the HPV lifecycle (Edwards et al., 2013a; Edwards et al., 2013b). HPV DNA exists as a double-stranded episome within infected cells. This viral DNA is tightly supercoiled, placing viral duplexes within close proximity to each other which would allow for multiple antiviral PAs to bind via bridging two pieces of viral DNA (**Fig. 4.13**).

Bridging of the DNA may cause structural alteration of the DNA severe enough to prevent replication and/or activate the DDR via breakage, as observed by Edwards and coworkers (Edwards et al., 2013b). The larger hairpin PAs, having a large reach, may be able to bind to two DNA molecules within close proximity to each where each of the arms of the hairpin bind to a different DNA molecule and act as a bridge; the smaller hairpins like KA1002 may bind to DNA with high stoichiometry but due to its smaller size may not have a large enough reach to adequately bridge to 2 different DNA molecules. Another possibility is that KA1002 binds to DNA similarly to NV1028 and NV1042 but dissociates too fast for it effectively activate the DNA damage response.

If this hypothesis is correct, this would show that *in vitro* biophysical studies using small DNA (versus larger DNA fragments) can provide key insights into *in vivo* drug interactions that have otherwise eluded investigators using other methods.



**Figure 4.13** Possible *in vivo* binding model for at least 2:1 NV1028-DNA binding. A, diagram of supercoiled plasmids adapted with permission from (Travers & Muskhelishvili, 2005). Copyright 2005 Springer Nature, license number 4317170561328; B, diagram of possible bridging interaction between NV1028 and supercoiled DNA. DNA helices shown in green and blue for contrast, bound PAs shown as black arrows.



## CHAPTER V

### EFFECTS OF N-TERMINAL GROUPS ON THE DNA BINDING KINETICS OF LARGE ANTIVIRAL POLYAMIDES

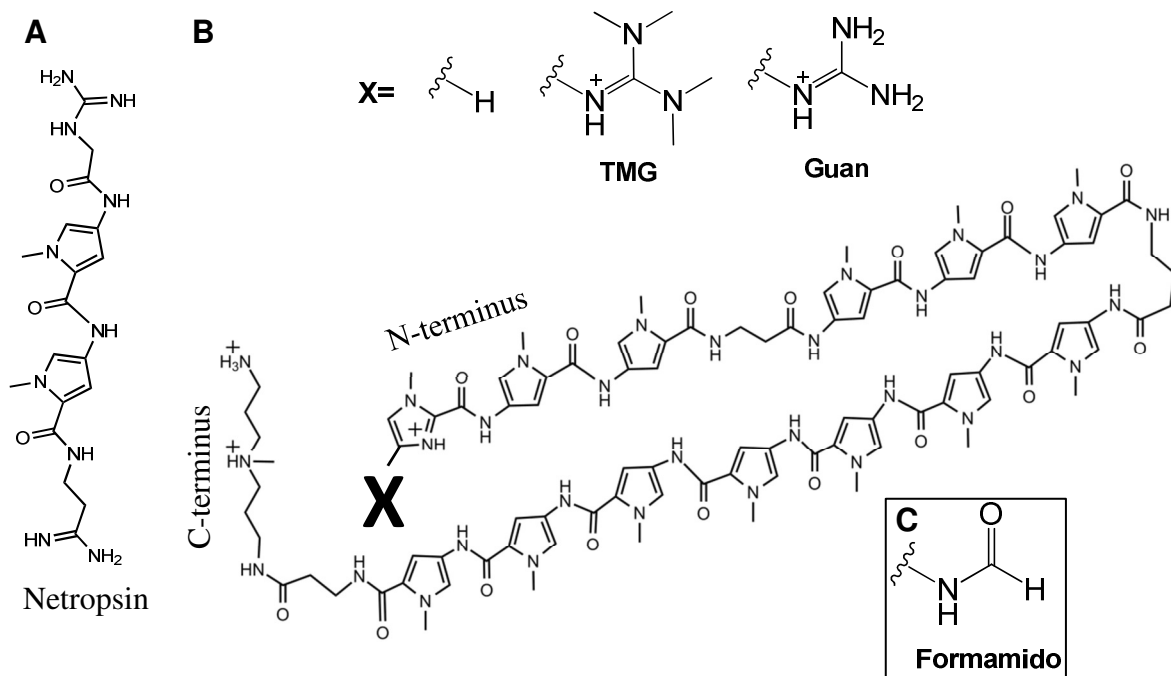
#### 5.1. Introduction and background

##### 5.1.1. Motivations

As discussed in **Chapter 1**, large hairpin polyamides (PA) NV1028 and NV1042 have been shown to have antiviral activity against oncogenic strains of HPV (Edwards et al., 2011; Koeller et al., 2014). Improvement of drug efficacy is a large part of the initial drug development process and one of the methods to accomplish this is through augmentation of the parent structure via addition of biomimetic functional groups. Guanidinylated analogs of NV1028 where the N-terminal des-amino of NV1028 is substituted with tetramethylguanidium (TMG) or guanidinium (Guan), to make NV1028-TMG and NV1028-Guan, respectively, are biomimetic to netropsin, which contains a guanidinyl functional group (**Fig. 5.1**).

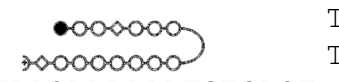
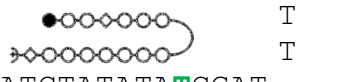
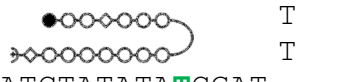
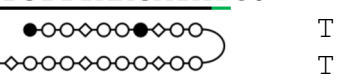
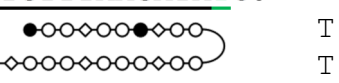
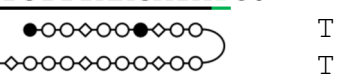
In terms of antiviral activity,  $IC_{50}$ 's of the NV1028 series were unaffected by the N-terminal substitutions (**Table 5.1**) (Castaneda, 2017). However, a 3-fold improvement of  $IC_{90}$  was observed for NV1028-Guan versus NV1028, while a 10-fold decrease in antiviral activity ( $IC_{90}$ ) was observed for NV1028-TMG versus NV1028 (ie Guan>H>TMG). It was shown that despite improved antiviral activity, DNA binding affinities were identical when studied via DNase I footprinting; thus, increased potency did not correlate with DNA binding affinity.

Guanidinylated analogs of NV1042, NV1042-TMG and NV1042-Guan,



**Figure 5.1** N-terminal functional groups studied in **Chapter 5**. A, structure of netropsin; B, structure of NV1028 shown and tetramethylguanadinium (TMG), guanadinium (Guan) used in heads series studies; C, formamido (F) used in (Liu, B. et al., 2017).

**Table 5.1** Polyamides and DNA hairpins relevant to **Chapter 5**

PA	DNA	Sequence <sup>a</sup>	HPV16 <sup>b</sup>	
			IC <sub>50</sub> (μM)	IC <sub>90</sub> (μM)
NV1028	<b>ODN-14 (LFTT T6)</b>	5' - TTCCA <b><u>TGTTTTTTA</u></b> CACTGT  3' - AAGGTAACAAAAAATGTGACT	0.100±0.020 (4)	1.113
NV1028-TMG	<b>ODN-14 SFAT T3</b>	5' - GC <b><u>TAGATATATA</u></b> GCTT  3' - CGATCTATATA <b>T</b> CGAT	0.3	>10
NV1028-Guan	<b>ODN-14 SFAT T23</b>	5' - GC <b><u>TAGATATATA</u></b> GCTT  3' - CGATCTATATA <b>T</b> CGAT	0.10	0.38
NV1042	<b>ODN-20-T3</b> <b>ODN-20-T16</b>	5' - GC <b><u>TATGTTTAAAGATA</u></b> <b>T</b> GCTT  3' - CGATACAAATTCTATACGAT	0.036±0.0004 (3)	0.351
NV1042-TMG		5' - GC <b><u>TATGTTTAAAGATA</u></b> <b>T</b> GCTT  3' - CGATACAAATTCTATACGAT	0.035	0.411
NV1042-Guan		5' - GC <b><u>TATGTTTAAAGATA</u></b> <b>T</b> GCTT  3' - CGATACAAATTCTATACGAT	0.038	0.340

<sup>a</sup> Pyrrole (○), Imidazole (●); β-alanine (◇), γ-aminobutyric acid shown as a loop. Expected binding sequences are **bold and underlined**. Colors denote dye position: primary dye positions are labeled in **red**, alternate dye positions are shown in **green**. **Black** ODNs are unlabeled and used as competitors or for CD spectroscopy.<sup>b</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011)

have also been prepared. Unlike NV1028, no improvement in either  $IC_{50}$ 's or  $IC_{90}$ 's was observed for these analogs compared to NV1042 (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013). It is unclear why changes in antiviral activity is observed for NV1028, but not for NV1042. Perhaps the functional groups affect the DNA binding kinetics of NV1028, but not those of NV1042. Work presented here seeks to determine the influence these N-terminal groups have on NV1028 and NV1042 DNA binding kinetics.

### **5.1.2. DNA hairpins used to study DNA binding kinetics with respect to N-terminal functional groups**

All DNA hairpins relevant to this chapter are shown in **Table 5.1**. A detailed discussion about DNA hairpin design can be found in **Chapter 2**. The binding sequences for NV1028 and its TMG and Guan analogs, also referred to as the NV1028 series, are identical; thus, the same DNA hairpins were used to characterize NV1028, NV1028-TMG, and NV1028-Guan. The data collected for the NV1028 series were principally collected with ODN-14-LFTT T6. ODN-14-SFAT T23 was used to characterize the fast phase association kinetics of NV1028-Guan only.

Since the recognition sequences for NV1042 and its TMG and Guan analogs (the NV1042 series) were assumed to be identical, the same DNA hairpins were used. ODN-20-T16 was used for the NV1042 series (unless otherwise specified).

## 5.2. Results

### 5.2.1. DNA binding affinities for NV1028 and NV1042 head series

Binding affinities ( $K_d$ ) were determined for both series using the method and fitting protocols described in **Chapter 2**.  $K_d$ s were very similar for all PAs within the NV1028 series ( $\sim 0.2$  nM) (**Table 5.2**).  $K_d$ s for NV1042 and NV1042-TMG were similar. However, the  $K_d$  of NV1042-Guan with ODN-20-T16 was 50-fold weaker than NV1042 and NV1042-TMG. It is not only distinct from the other NV1042 series competition  $K_d$ s, but also from the pattern observed with NV1028 series. Given that Guan is smaller than TMG, it is unlikely that steric hindrance is the cause. The difference in  $K_d$  might indicate a unique binding interaction between NV1042-Guan and the DNA, or dye interference.

### 5.2.2. DNA binding affinities determined via competition for NV1028 and NV1042 head series

In order to rule out the possibility of dye interference, DNA binding affinities of PAs to their DNA hairpins were determined indirectly via competition experiments using the method described in **Chapter 2**.

Competition  $K_d$ s obtained for both NV1028 and NV1042 series are shown in **Table 5.2**. The competition  $K_d$  for NV1028-Guan is very similar to the directly determined  $K_d$ . Competition experiments for NV1028-Guan and ODN-14-LFTT T6 are pending as of this writing.

Competition  $K_d$ s could not be determined for NV1042 or NV1042-TMG with ODN-20-T16 due to a small dynamic range and slow dissociation kinetics as discussed in **Chapter 4**. The weak  $K_d$  of NV1042-Guan made a competition

**Table 5.2** Binding affinities for NV1028 and NV1042 series<sup>a</sup>

PA	Direct K <sub>d</sub> (nM)	Comp K <sub>d</sub> (nM)	Alternate K <sub>d</sub> (nM)	HPV16 IC <sub>50</sub> <sup>f</sup> (μM)	HPV16 IC <sub>90</sub> <sup>f</sup> (μM)
<b>NV1028</b>	0.23±0.02	0.03±0.01 <sup>b</sup>		0.100±0.020 (4)	1.113
<b>NV1028-TMG</b>	0.23±0.06	-- <sup>e</sup>		0.304	>10
<b>NV1028-Guan</b>	0.18±0.01 <sup>c</sup>	0.17 <sup>b,c</sup>		0.103	0.378
<b>NV1042</b>	0.34±0.16	--	0.13±0.02 <sup>d</sup>	0.036±0.0004 (3)	0.351
<b>NV1042-TMG</b>	0.74±0.34	--	0.63±0.23 <sup>d</sup>	0.035	0.411
<b>NV1042-Guan</b>	51.7±2.5	--	101±13.7 <sup>d</sup>	0.038	0.340

<sup>a</sup>K<sub>d</sub> collected via fluorescence method, conditions: 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. Unless specified otherwise, data was collected in triplicate with relative error shown.

<sup>b</sup>Collected by Yang Song, see discussion in **Chapter 3**

<sup>c</sup>As of this writing, only 1 replicate. Fit error shown.

<sup>d</sup>Competition infeasible with ODN-20-T16, K<sub>d</sub> determined with alternate DNA ODN-20-T3

<sup>e</sup> competition experiment not performed as of this writing

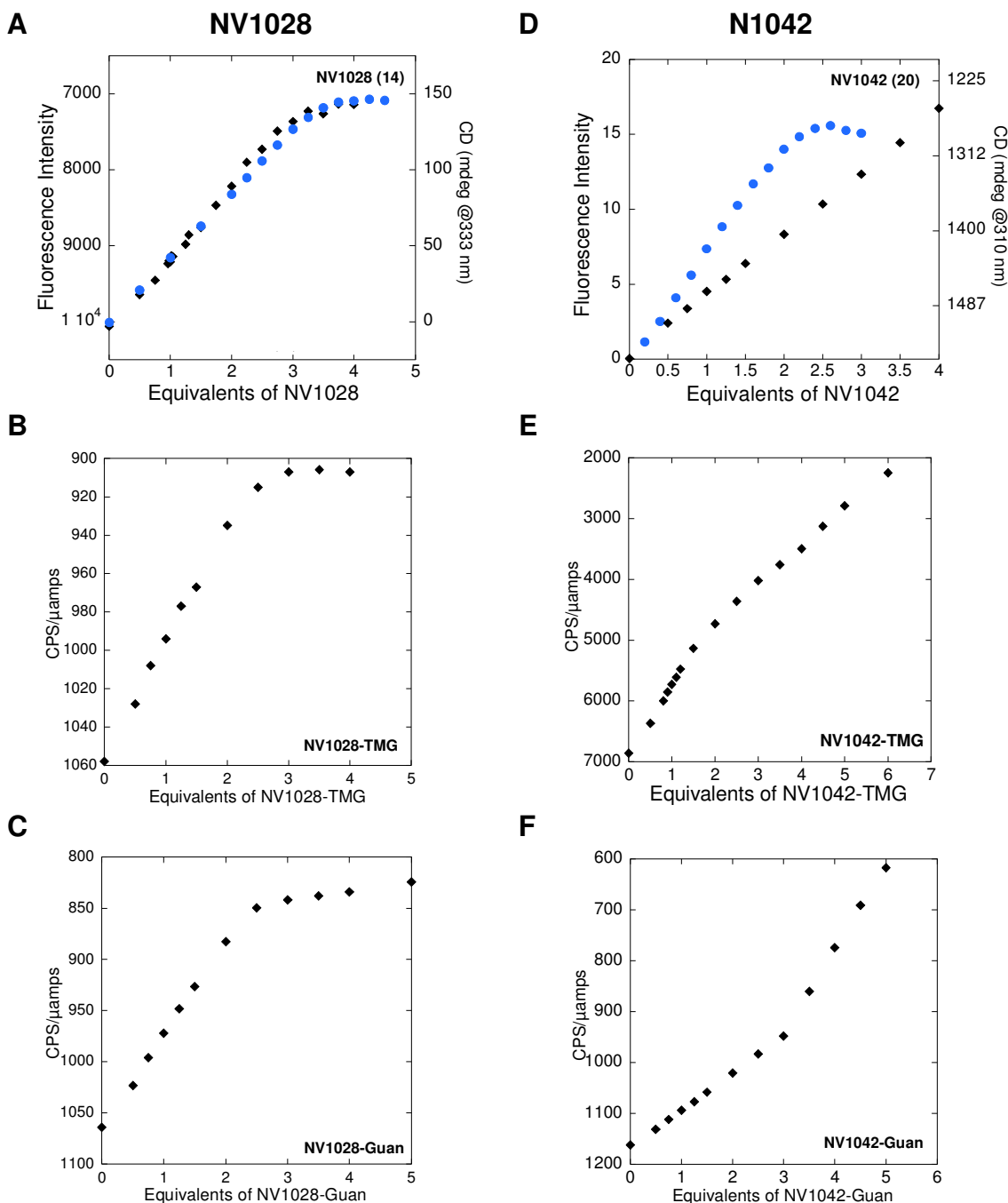
<sup>f</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011). IC<sub>50</sub> and IC<sub>90</sub> are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

experiment impractical. Instead,  $K_d$ s were determined directly using ODN-20-T3 which is identical in sequence to ODN-20-T16 but with the dye moved from T16 to T3.  $K_d$ s of NV1042 and NV1042-TMG with ODN-20-T3 were similar to those determined with ODN-20-T16; thus, is it unlikely the dye has an observable effect on NV1042 and NV1042-TMG DNA binding. The  $K_d$  of NV1042-Guan with ODN-20-T3 was also similar to the  $K_d$  of NV1042-Guan with ODN-20-T16 which suggests that the dye placement does not significantly influence NV1042-Guan binding to DNA.

### 5.2.3. PA-DNA binding stoichiometry for NV1028 and NV1042 series

In **Chapters 3** and **4**, NV1028 and NV1042 were shown to bind to DNA with a stoichiometric ratio of at least 2:1 (**Fig. 5.2**). Due to a limited supply of the guanidylated analogs and the high concentration required for CD spectroscopy, only fluorescence spectroscopy was used to determine PA-DNA binding stoichiometries.

PA-DNA binding stoichiometry for NV1028-TMG and NV1028-Guan were determined to have binding stoichiometric ratios of at least 2:1 (**Fig 5.2A-C**). For the NV1042 series no plateau in fluorescence intensity change was observed in excess of 4 equivalents (**Fig. 5.2D-F**). A change in slope, or break, is observed for NV1042-TMG at around 1 equivalent, and for NV1042-Guan at around 3 equivalents which indicates an interaction between the PA and the DNA. However, since no plateau in fluorescence signal and further interactions between the DNA and PA are observed, these breaks should not be considered



**Figure 5.2** Stoichiometry data for NV1028 series with ODN-14-LFTT T6 (left) and NV1042 series with ODN-20-T16 (right). ●, data collected via CD spectroscopy; ◆, data collected via fluorescence spectroscopy. A and D are reproduced from **Chapters 3** and **4**. Conditions: 5 μM DNA (CD), 5 nM DNA (Fluor) 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. CD: spectra were averaged from 5 accumulations, each point is the average mdeg of 2 spectra at a specific wavelength. Fluor: each point is the average of 3 intensities.



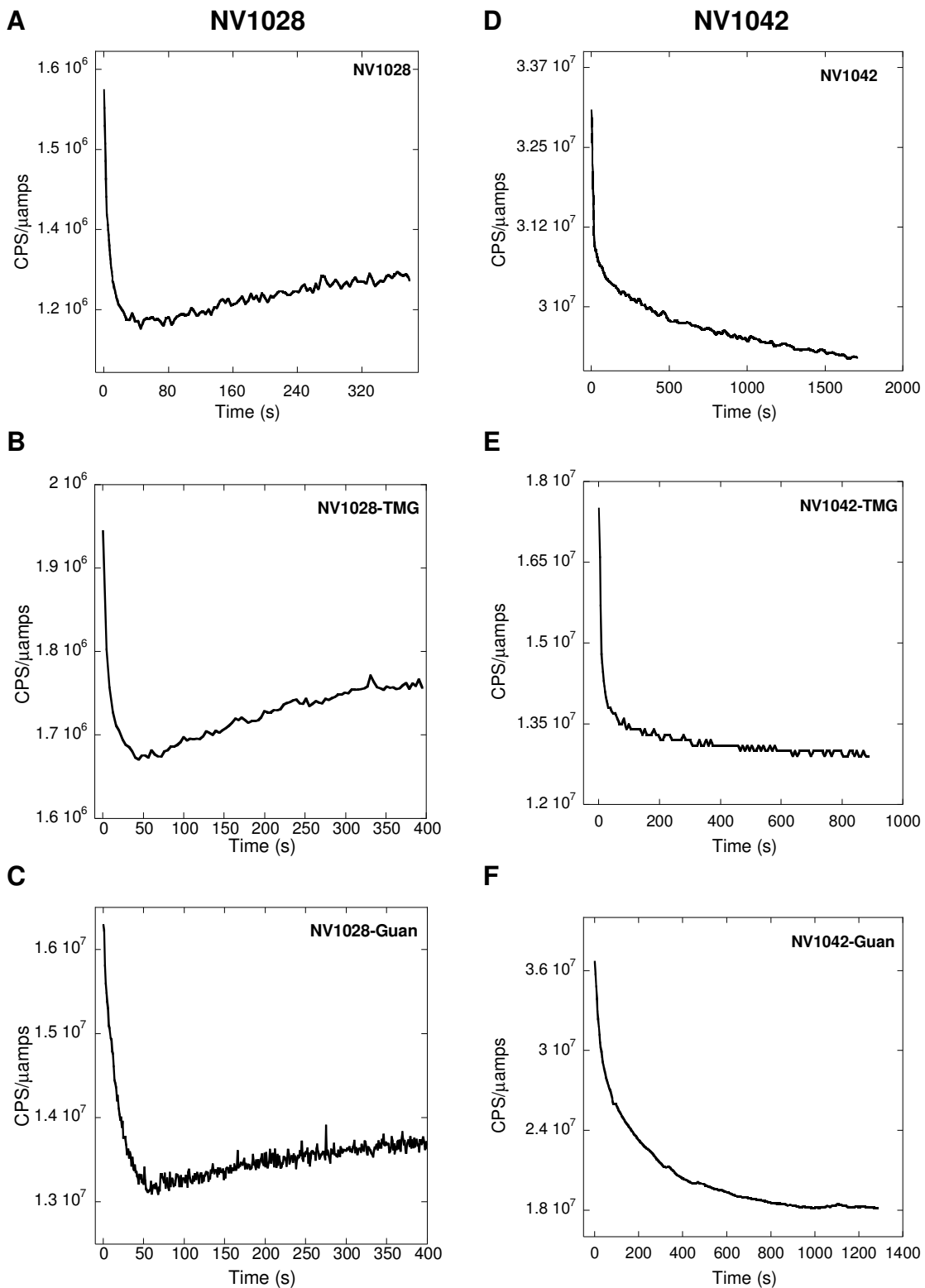
saturation points. Full saturation of the DNA by the NV1042 series PAs may include very weak and/or slow interactions, which are more difficult to observe at nM concentrations. Currently, the only determination of binding stoichiometry for a PA from the NV1042 series has been via CD by Kristin Bales for NV1042, which was determined to bind to DNA at least 2:1 (**Fig. 5.2D**). See **Chapter 4** for further discussion. A change in slope, but not saturation of the DNA by NV1042-TMG and NV1042-Guan, is observed (**Fig. 5.2EF**). Due to limited quantities of NV1042-TMG and NV1042-Guan, these experiments were performed only once: replicates are pending as of this writing.

#### **5.2.4. PA-DNA association kinetics for NV1028 and NV1042 head series**

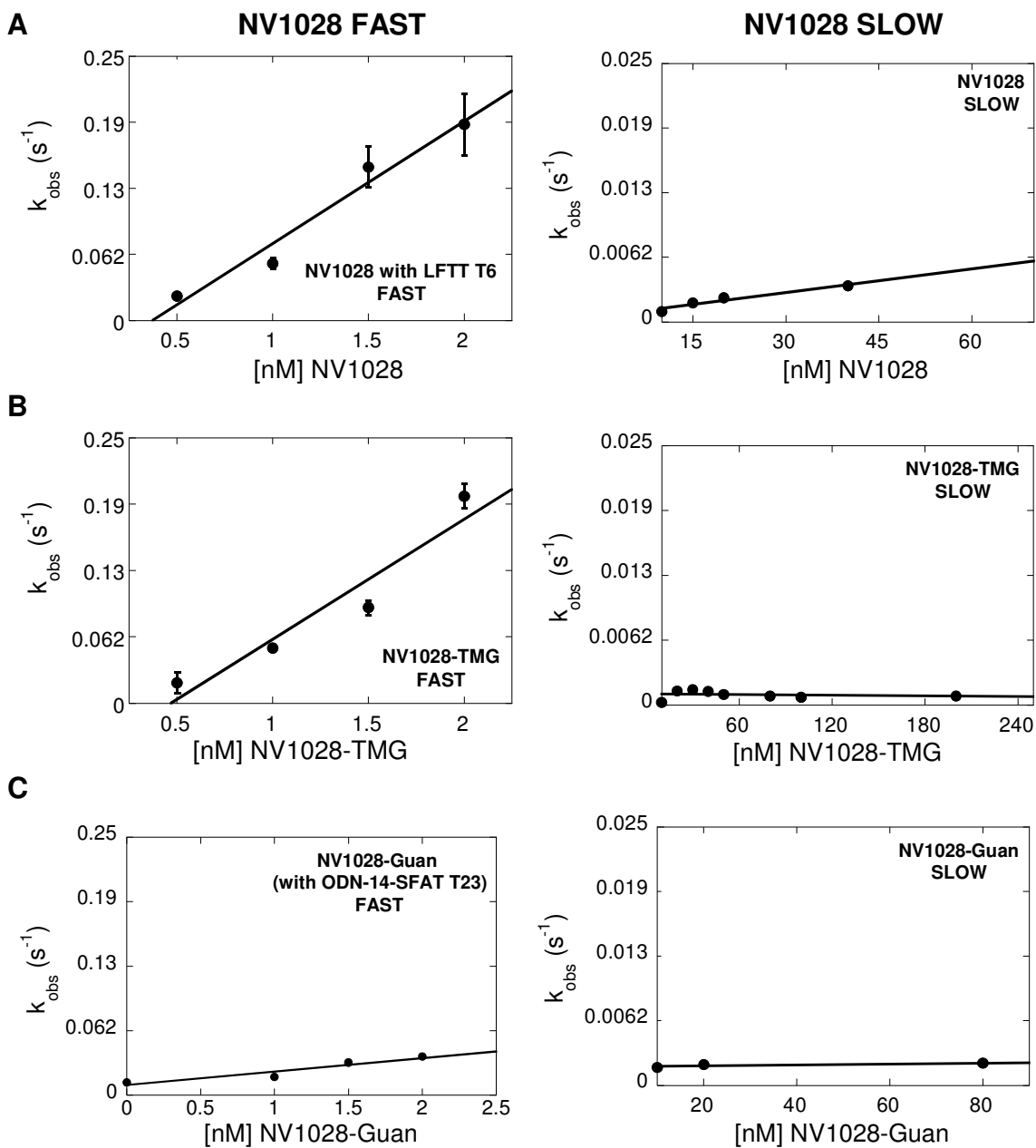
##### **5.2.4.1 NV1028 series-DNA association kinetics**

In **Chapter 3**, NV1028-DNA association kinetics were shown to be biphasic with an initial fast phase where fluorescence intensity decreased, followed by an exponentially slower phase where fluorescence intensity recovered slightly. These phases were present for NV1028 regardless of A/T pattern or arrangement within the binding sequence. NV1028-TMG and NV1028-Guan also exhibited biphasic association kinetics (**Fig. 5.3A-C**). This suggests that NV1028 interacts with the DNA similarly regardless of the N-terminal group.

The 2 phases of NV1028 were determined to be linearly dependent on PA concentration meaning each phase was related to an intermolecular event and not a rearrangement (**Fig. 5.4, Table 5.3**). The 2 observable phases were fit independently, as previously described in **Chapter 3**. For NV1028 and



**Figure 5.3** Association kinetics traces for NV1028 series with ODN-14-LFTT T6 (left) and NV1042 series with ODN-20-T16 (right). Conditions: 2.5 nM DNA, excess PA, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C.



**Figure 5.4** NV1028 series secondary association plots. NV1028 and NV128-TMG with ODN-14-LFTT T6, NV1028-Guan (NV1115) fast phase shown with ODN-14-T23, NV1028-Guan slow phase shown with ODN-14-LFTT T6. NV1028 association curves and secondary fits were determined in duplicate, relative error is shown. Association curves and the secondary linear fits for the TMG and Guan analogs were determined once, fit error is shown.

**Table 5.3** Association rate constants for NV1028 and NV1042 series<sup>a</sup>

PA	Direct $K_d$ (nM)	$k_{on}^{fast}$ ( $\times 10^6 M^{-1} s^{-1}$ )	$k_{on}^{slow}$ ( $\times 10^6 M^{-1} s^{-1}$ )	HPV16 $IC_{50}^c$ ( $\mu M$ )	HPV16 $IC_{90}^c$ ( $\mu M$ )
NV1028	0.23±0.02	120±20	0.08±0.01 <sup>c</sup>	0.100±0.02 (4)	1.113
NV1028-TMG	0.23±0.06	130±25	[PA] ind	0.304	>10
NV1028-Guan	0.18±0.01 <sup>c</sup>	13±3 <sup>b</sup>	[PA] ind	0.103	0.378
NV1042	0.34±0.16	1.4±0.25		0.036±0.0004 (3)	0.351
NV1042-TMG	0.74±0.34	11±2		0.035	0.411
NV1042-Guan	51.7±2.5	0.2±0.1		0.038	0.340

<sup>a</sup> determined via fluorescence assay, conditions: 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, pH 7.4, 1 mM CHAPS, 25 °C. NV1028 and NV1042 association curves and secondary fits were determined in duplicate, relative error is shown. Association curves and the secondary linear fits for the TMG and Guan analogs were determined once, fit error is shown.

<sup>b</sup> NV1028 and NV1028-TMG fast phase collected with ODN-14-LFTT T6, NV1028-Guan fast phase collected with ODN-14-SFAT T23

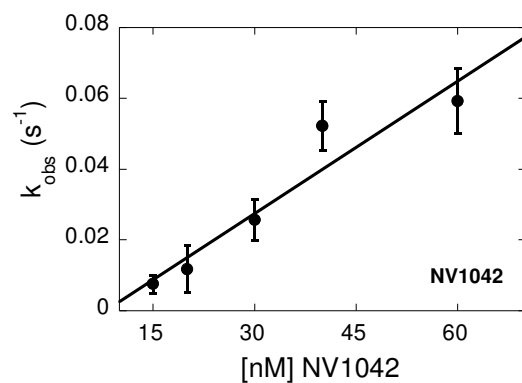
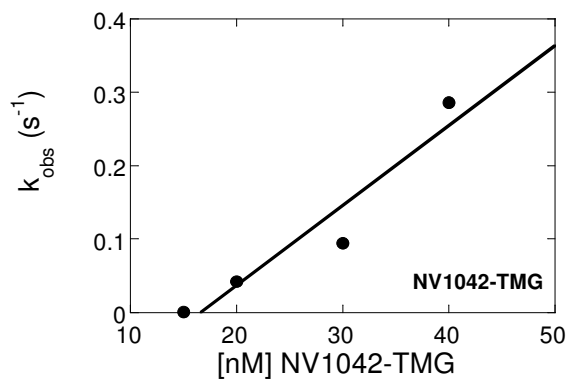
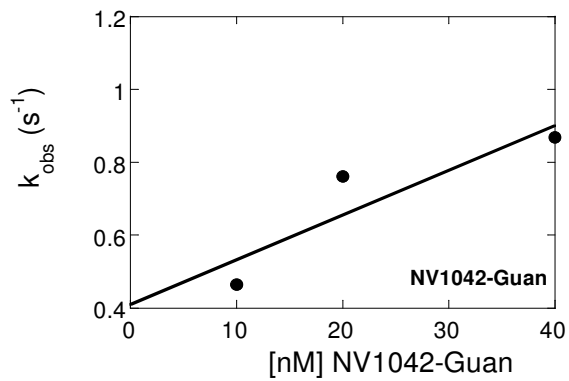
<sup>c</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011).  $IC_{50}$  and  $IC_{90}$  are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

NV1028-TMG, the fast phase association rate constant ( $k_{\text{on fast}}$ ) is at the diffusion-control limit. In initial experiments, NV1028-Guan  $k_{\text{on fast}}$  was determined with ODN-14-SFAT T23 to be 10-fold slower than the  $k_{\text{on fast}}$  of NV1028 and NV1028-TMG determined with ODN-14-LFTT T6. Characterization of  $k_{\text{on fast}}$  of NV1028-Guan with ODN-14-LFTT T6 is pending as of this writing.

For NV1028, the slow phase association rate constant ( $k_{\text{on slow}}$ ) was  $10^3$  slower than  $k_{\text{on fast}}$ . Unexpectedly, while the slow phase of NV1028 was determined to be concentration dependent, the slow phases of NV1028-TMG and NV1028-Guan appear to be concentration independent on the scale of 2 hours. This means that the secondary phase observed for NV1028-TMG and NV1028-Guan could either be a rearrangement or is very slow on this timescale.

#### **5.2.4.2. NV1042 series-DNA association curves**

In **Chapter 4**, NV1042 DNA association curves also appeared to be biphasic but the fluorescence intensity change was uniform in direction (**Fig 5.3D-F**). NV1042-TMG association curves behaved similarly, but NV1042-Guan appears clearly more monophasic. When fit to a monophasic equation, the association rate constants ( $k_{\text{on}}$ ) for the NV1042 series were much more varied ranging from  $0.2\text{-}11 \times 10^6 \text{ M}^{-1} \text{ s}^{-1}$ . Secondary plots for the NV1042 series are shown in **Fig. 5.5** and summarized in **Table 5.3**. NV1042-Guan DNA association is slowest; this is reasonable given the much weaker binding affinity observed with ODN-20-T16. NV1042-TMG  $k_{\text{on}}$  was approximately 8-fold faster than that observed with NV1042. Our findings are consistent with what Liu and coworkers

**A****B****C**

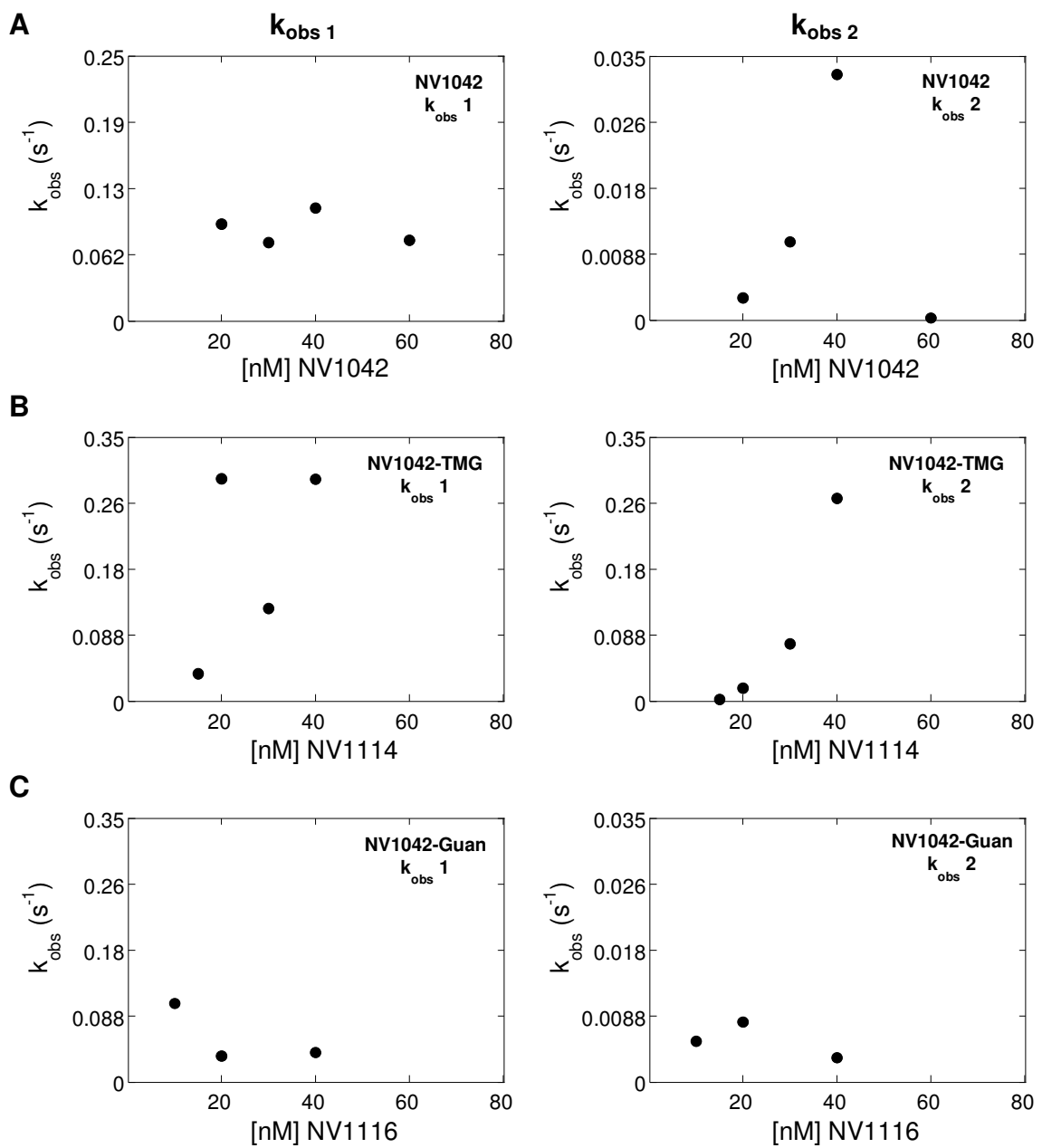
**Figure 5.5** NV1042 series with ODN-20-T16 secondary association plots. NV1042 association curves and secondary fits were determined in duplicate, relative error is shown. Association curves and the secondary linear fits for the TMG and Guan analogs were determined once, fit error is shown.

observed when comparing association rate constants for TMG versus formamido N-terminal substituted 8-ring hairpin PAs: a 2-5-fold increase in  $k_{on}$  was observed for the TMG analogs when compared to  $k_{on}$  of the parent structure (Liu, B. et al., 2017) (**Fig. 5.1**). In general, NV1042 series DNA association was slower than that of the NV1028 series, with the Guan analog associating with DNA the slowest of both series.

Given the biphasic appearance of the NV1042 and NV1042-TMG DNA association curves, curves were also fit using a biexponential equation to see if the biphasic appearance could be attributed to multiple binding events. And if so, could we separate the binding events and determine their microscopic  $k_{on}$ . When fit to a biexponential equation (**eqn. 3.1**), the additional variables of the biphasic equation generally made fits of the curves (R values) better than the fits using a monophasic equation. As discussed in **Chapter 4**, no clear linear concentration dependence was observed with respect to PA concentration for NV1042.

For both NV1042-TMG and NV1042-Guan,  $k_{obs\ 1}$  values seem to be concentration independent (**Fig. 5.6 left**). However, like NV1042, when  $k_{obs\ 2}$  is plotted with respect to PA concentration dependence, NV1042-TMG and NV1042-Guan seem to have a linear concentration dependence (**Fig. 5.6 right**). Unfortunately, since association curves were only collected once, it is unclear whether these trends are genuine or coincidental.

Thus, despite the non-monophasic appearance of NV1042 and NV1042-TMG curves, the data indicates these association curves correlate to only a single binding event on this timescale. It may be that additional binding events



**Figure 5.6** Representative secondary plots NV1042 series when association curves were fit to a biexponential equation (eqn. 3.1).

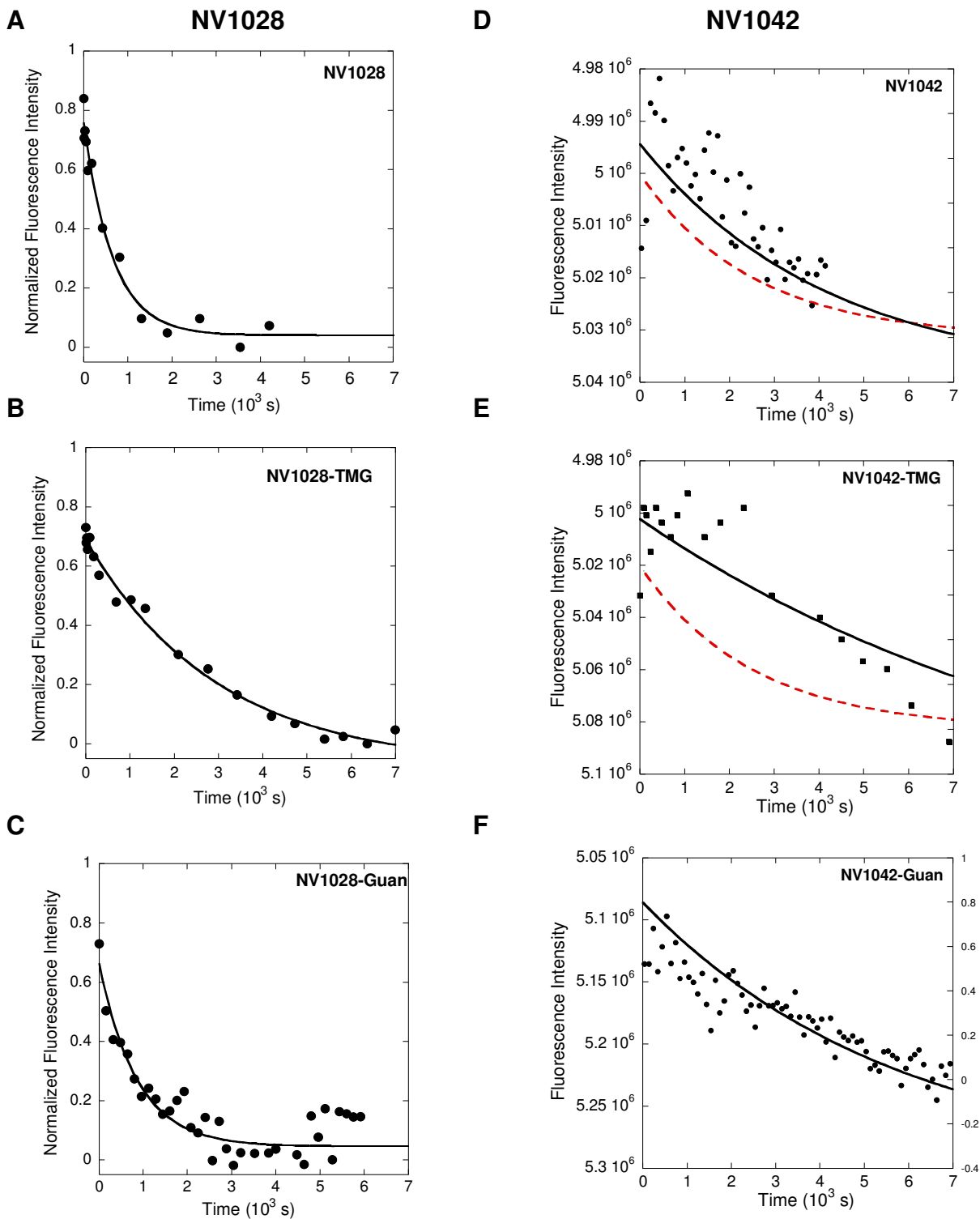


occur at a much slower rate, or that their influence on the dye is harder to observe.

### 5.2.5. PA-DNA dissociation rate constants for NV1028 and NV1042 head series

Dissociation rate constants for 1:1 PA-DNA complexes were determined using the method and fitting protocol described in **Chapter 2**. In general, the NV1042 series dissociated from DNA at a slower rate than the NV1028 series. Additionally, less than 50% signal recovery (return in emission intensity due to PA dissociation from DNA) is observed for both series (**Fig. 5.7, Table 5.4**). NV1028 dissociation half-life ( $t_{1/2}$ ) was ~15 min, while  $t_{1/2}$  of NV1042 was ~30 min. This observation is consistent with what we found in **Chapter 4** where it was shown that as PA size increased, dissociation rates slowed. If we assume residence time to be the most important kinetic parameter to antiviral behavior, it was expected that NV1028-Guan, being the most potent of the series, would dissociate the slowest from its target DNA, and NV1028-TMG, being the least potent of the series, would dissociate the fastest. The opposite trend was observed for the NV1028 series (**Fig. 5.8A**). However, the differences in  $K_{off}$  were small only 2-4-fold different from each other, so it is unclear whether these small differences in  $K_{off}$  could affect the antiviral potencies in a significant way.

NV1042 series  $k_{off}$  were difficult to obtain due to a small dynamic range and very slow dissociation kinetics, as discussed in **Chapter 4**. Although the data was fit using the fitting protocol in **Chapter 2**, the fits are poor (low  $R^2$ ) due to low signal-to-noise ratios. Additionally, because a true plateau was not reached



**Figure 5.7** Representative dissociation curves for NV1028 series with ODN-14-LFTT T6 (left) and NV1042 series with ODN-20-T16 (right). Conditions: 1:1 PA-DNA complex, 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. Red dotted line is a trend line with estimated  $K_{\text{off}}$  of  $0.4 \times 10^{-3} \text{ s}^{-1}$ .

**Table 5.4** Dissociation rate constants and percent fluorescence recovery for NV1028 and NV1042 series<sup>a</sup>

PA	$K_{off}$ ( $\times 10^{-3} \text{ s}^{-1}$ )	% recov	HPV16 $IC_{50}^c$ ( $\mu\text{M}$ )	HPV16 $IC_{90}^c$ ( $\mu\text{M}$ )
<b>NV1028</b>	0.6±0.07	24	0.100±0.02 (4)	1.113
<b>NV1028-TMG</b>	0.33±0.04	45	0.304	>10
<b>NV1028-Guan</b>	1.2±0.10	70	0.103	0.378
<b>NV1042</b>	≤0.4 <sup>b</sup>	28	0.036±0.0004 (3)	0.351
<b>NV1042-TMG</b>	≤0.4 <sup>b</sup>	12	0.035	0.411
<b>NV1042-Guan</b>	0.17±0.01	7	0.038	0.340

<sup>a</sup>Collected via fluorescence method, conditions: 10 mM HEPES, 50 mM NaCl, 1 mM EDTA, 1 mM CHAPS, pH 7.4, 25 °C. NV1028 dissociation curves were determined in duplicate, relative error is shown. Dissociation curves for NV1042 and all TMG- and Guan- analogs were determined once, fit error is shown.

<sup>b</sup>estimate reported given data quality. Estimations were made based on the similarity in line trends of NV1042 and NV1042-TMG and the maximum  $K_{off}$  value determined by Kaleidagraph for the two PA. Estimated  $K_{off}$  value is shown as red dotted line in **Fig 5.7**.

<sup>c</sup> (Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J., 2013; Castaneda et al., 2017; Edwards et al., 2011).  $IC_{50}$  and  $IC_{90}$  are defined as the concentration of PA at viral concentration is decreased by 50% or 90%, respectively, in vitro. Numbers in parenthesis denote sample size.

within the experimental timeframe, data could not be properly normalized. Thus, the  $K_{\text{off}}$ s determined for the NV1042 series were treated as estimates. Unlike the NV1028 series, there is no difference in potency between the different NV1042 analogs and, as expected,  $K_{\text{off}}$  (and, thus, residence times) were similar to each other (**Fig. 5.8B**). Representative dissociation kinetics traces are shown unnormalized in **Fig. 5. 7**. The fit for the “normalized” data is overlaid as a solid black line, while a trend line with the estimated  $K_{\text{off}}$  of  $0.4 \times 10^{-3} \text{ s}^{-1}$  is shown as a dashed red line.

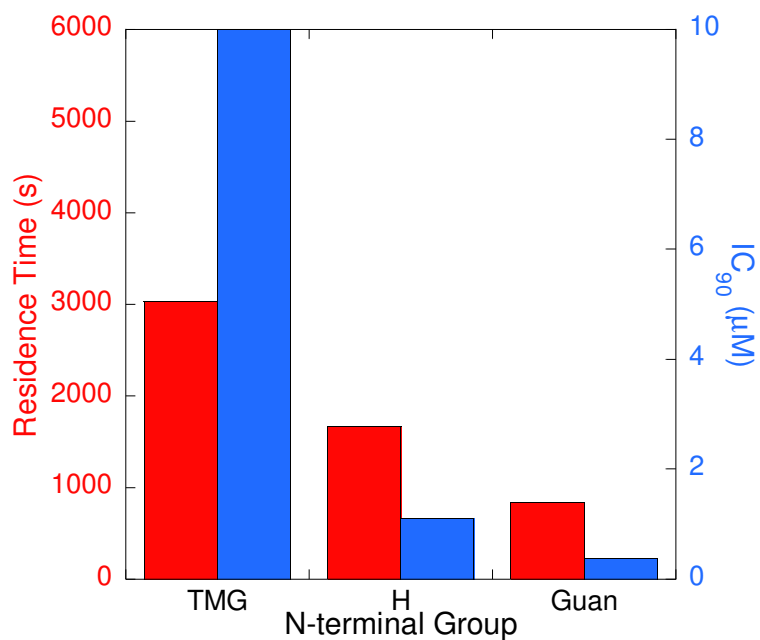
Given that all NV1028 series and NV1042 bind at least 2:1 and we suspect NV1042-TMG and NV1042-Guan binding stoichiometries are similar to that of NV1042, the dissociation rate constants presented here should be treated as macroscopic dissociation rate constants ( $K_{\text{off}}$ ). Treating these dissociation rate constants as macroscopic is also supported by the lack of signal recovery observed which suggests that not all the PA is dissociated within the experimental timeframe despite no further change in signal intensity.

### **5.3. Discussion**

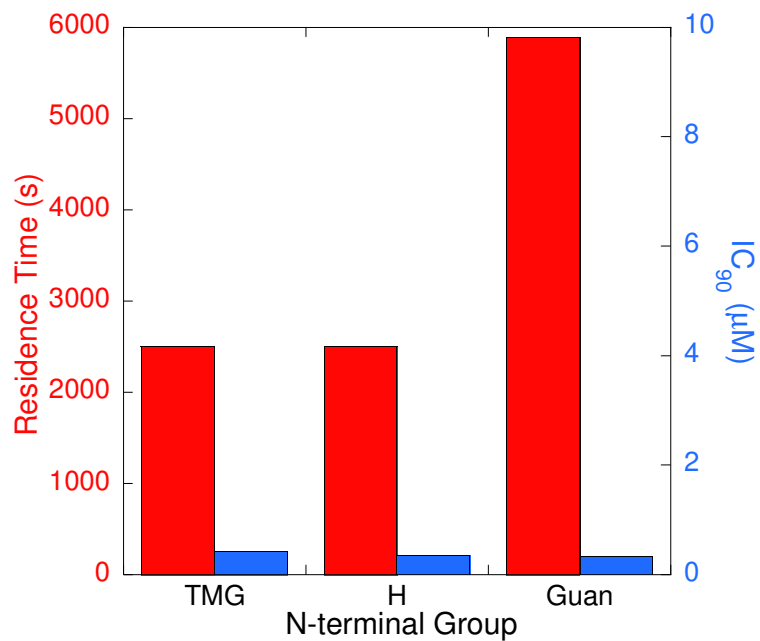
#### **5.3.1. DNA binding affinities within an N-terminal series**

Despite their different antiviral activities, NV1028 series exhibited similar DNA binding affinities ( $K_{\text{d}}$ ). This result is consistent with what was observed by Castaneda and coworkers using DNase I footprinting and shows that the differences between the methods do not affect the DNA binding affinities (Castaneda et al., 2017). It also further emphasizes how  $K_{\text{d}}$  does not correlate

**A**



**B**



**Figure 5.8** Residence time and potency with respect to N-terminal functional groups for the NV1028 series, A, and NV1042 series, B.

with PA antiviral activity.

Antiviral activity is identical for the PAs in the NV1042 series. However, only the  $K_d$ s of NV1042 and NV1042-TMG were similar to each other. Additionally, despite the improved antiviral activity of NV1042 and NV1042-TMG compared to the NV1028 series, their  $K_d$ s were similar to the NV1028 series. These results are consistent with what has been observed previously with DNase I footprinting (He et al., 2014; Vasilieva et al., 2016). The  $K_d$  of NV1042-Guan was at least 100-fold weaker than all the other PAs tested. The weaker  $K_d$  is likely due to slower DNA association ( $k_{on}$ ) by NV1042-Guan since  $K_{off}$  is similar to NV1042 and NV1042-TMG.

### **5.3.2. Reconciliation of $K_d$ and kinetic rate constants with respect to N-terminal series**

As discussed in **Chapters 3** and **4**,  $K_d$  is a function of the kinetic rate constants,  $k_{on}$  and  $K_{off}$ , by **eqn. 3.3**. To determine if the experimentally determined parameters were consistent with one another, data reconciliation was performed on the NV1028 and NV1042 series. Reconciliation of the NV1028 series binding kinetics and binding affinity was accomplished in **Chapters 3** and **4** by taking the square root of the product of the 2 concentration-dependent  $k_{on}$  (**eqn. 3.2**) (see **Chapter 3** for discussion). Using the  $K_{on}$ , and experimentally determined  $K_d$ , the  $K_{off}$  was computed and compared to the experimentally determined  $K_{off}$ . If the computed  $K_{off}$  and experimentally determined  $K_{off}$  were similar, then experimental parameters were deemed internally consistent.

However, given that the slow phases of both NV1028-TMG and NV1028-Guan were concentration independent, experimentally determined biophysical parameters were not internally consistent.

NV1042 biophysical parameters are internally consistent (**Table 5.5**) (see **Chapter 4** for discussion). NV1042-TMG and NV1042-Guan experimental parameters do not reconcile as the calculated  $K_{\text{off}}$  is greater than the experimental  $K_{\text{off}}$  by at least an order of magnitude.

The lack of internal consistency for the guanidinylated analogs of both series is likely due to my ability to experimentally determine only apparent values and not microscopic values. Furthermore, given that fluorescence intensity recovery is typically less than 30% for PAs in both series, internal consistency of NV1028 and NV1042 may only be coincidental since this would seem to indicate that not all PA is dissociating.

### **5.3.3. PA-DNA dissociation rate constants with respect to N-terminal functional groups**

The  $K_{\text{off}}$  for PAs of the NV1028 series ranged from 0.3-1.2  $\text{s}^{-1}$  where  $\text{TMG} < \text{H} < \text{Guan}$ , but both TMG and Guan  $K_{\text{off}}$  values could be considered similar to H given that they were only 2-fold different. If we assume that better drugs have slower dissociation rate constants, then it is counterintuitive that NV1028-TMG would dissociate from DNA the slowest and that NV1028-Guan would dissociate from DNA the fastest, given their respective potencies (**Fig. 5.8A**). NV1042 series  $K_{\text{off}}$  are less precise in their determination, but DNA dissociation curves appear objectively slower than those of the NV1028 series. Under the

**Table 5.5** Comparison of experimental and computed PA-DNA binding constants for NV1028 and NV1042 series

PA	Direct $K_d$ (nM)	$k_{on\ fast}$ ( $10^6\ M^{-1}\ s^{-1}$ )	$k_{on\ slow}$ ( $10^6\ M^{-1}\ s^{-1}$ )	Calculated $K_{on}^a$ ( $10^6\ M^{-1}\ s^{-1}$ )	$K_{off}$ ( $10^{-3}\ s^{-1}$ )	Calculated $K_{off}^b$ ( $10^{-3}\ s^{-1}$ )
NV1028	0.23±0.02	120±19	0.077±0.01	3.0	0.6±0.07	0.69
NV1028-TMG	0.23±0.06	130±25	[PA] ind	--	0.33±0.04	--
NV1028-Guan	0.18±0.01	37	[PA] ind	--	1.2±0.10	--
NV1042	0.34±0.16	1.4±0.2		--	≤0.4	0.5
NV1042-TMG	0.74±0.34	11±2		--	≤0.4	8
NV1042-Guan	51.7±2.5	0.2±0.1		--	0.17±0.01	10

<sup>a</sup> computed using  $K_{on} = \sqrt{k_{on\ fast} * k_{on\ slow}}$  (Lacy et al., 2002)

<sup>b</sup> computed using  $K_{off} = K_{on} * K_d$



above assumption, this result is expected given the improved potency of the NV1042 series versus the NV1028 series. This correlation has already been discussed with respect to size in **Chapter 4** and is unrelated to the N-terminal groups.

Recovery of fluorescence intensity upon addition of competitor DNA for both series is well below 100% which suggests that if multiple equivalents are binding to DNA, then not all equivalents are dissociating from DNA upon addition of competitor. In general, lower percent recovery of fluorescence is observed for the NV1042 series than the NV1028 series. Perhaps potency is related to the equivalents that do not dissociate from DNA within our experimental timeframe.

#### **5.3.4. PA-DNA association rate constants with respect to N-terminal functional groups**

For all PAs in the NV1028 series, the initial fast phase is concentration dependent; thus, it corresponds to an intermolecular event, such as DNA binding. Association curves for NV1028 indicate the second observable association phase also corresponds to a binding event. However, data for NV1028-TMG and NV1028-Guan indicate the slow association phases observed for these guanidinylated PAs are related to unimolecular events: perhaps a rearrangement of the DNA, or sliding of the PA on the DNA.

With respect to N-terminal functional groups, the patterns in association rate constants differ for the 2 series (comparing NV1028-Guan  $k_{on\ fast}$  with the monophasic  $k_{on}$  determined for NV1042). For the NV1028 series, H=TMG>Guan; for the NV1042 series, TMG>H>Guan. It is noteworthy that both NV1028-Guan

and NV1042-Guan have the slowest DNA association rate constants within their respective series. However, the slower DNA binding is likely unrelated to steric hindrance. NV1042-TMG associates with DNA the fastest within its series despite the bulkiness of TMG; both NV1028 and NV1028-TMG  $k_{on\ fast}$  are both diffusion-limited. In a recently published study comparing the DNA binding affinities and kinetics of formamido substituted versus TMG-substituted 8-ring hairpin PAs, faster association rate constants were observed for TMG analogs with a  $\beta$  in the top strand of the hairpins than their formamido counterparts. See **Fig. 5.1** for formamido structure (Liu, B. et al., 2017). If NV1042-TMG and NV1042-Guan interact with DNA similarly they might both have a second phase that is not observable using our fluorescence assay, or within the timescale of our association experiments. However, given their 50-fold difference in binding affinities, it seems unlikely that they would interact with DNA similarly.

Perhaps the bulkiness of TMG affects the conformation of the PA and/or limits the overall number of favorable enthalpic interactions a PA can make with DNA. Castaneda and coworkers reported subtle differences in DNase I footprints and affinity cleavage sites of the different NV1028 analogs and speculated that TMG may bind more non-specifically than the other analogs. As discussed in **Chapter 4**, if we assume that the residues of hairpin PAs do not bind to individual base pairs simultaneously, and that nucleotide binding happens in a zipper-type manner (**Fig. 4.11**), then a PA that makes fewer bonds would associate with DNA faster than a PA that makes more bonds. NV1028-TMG and NV1028-Guan

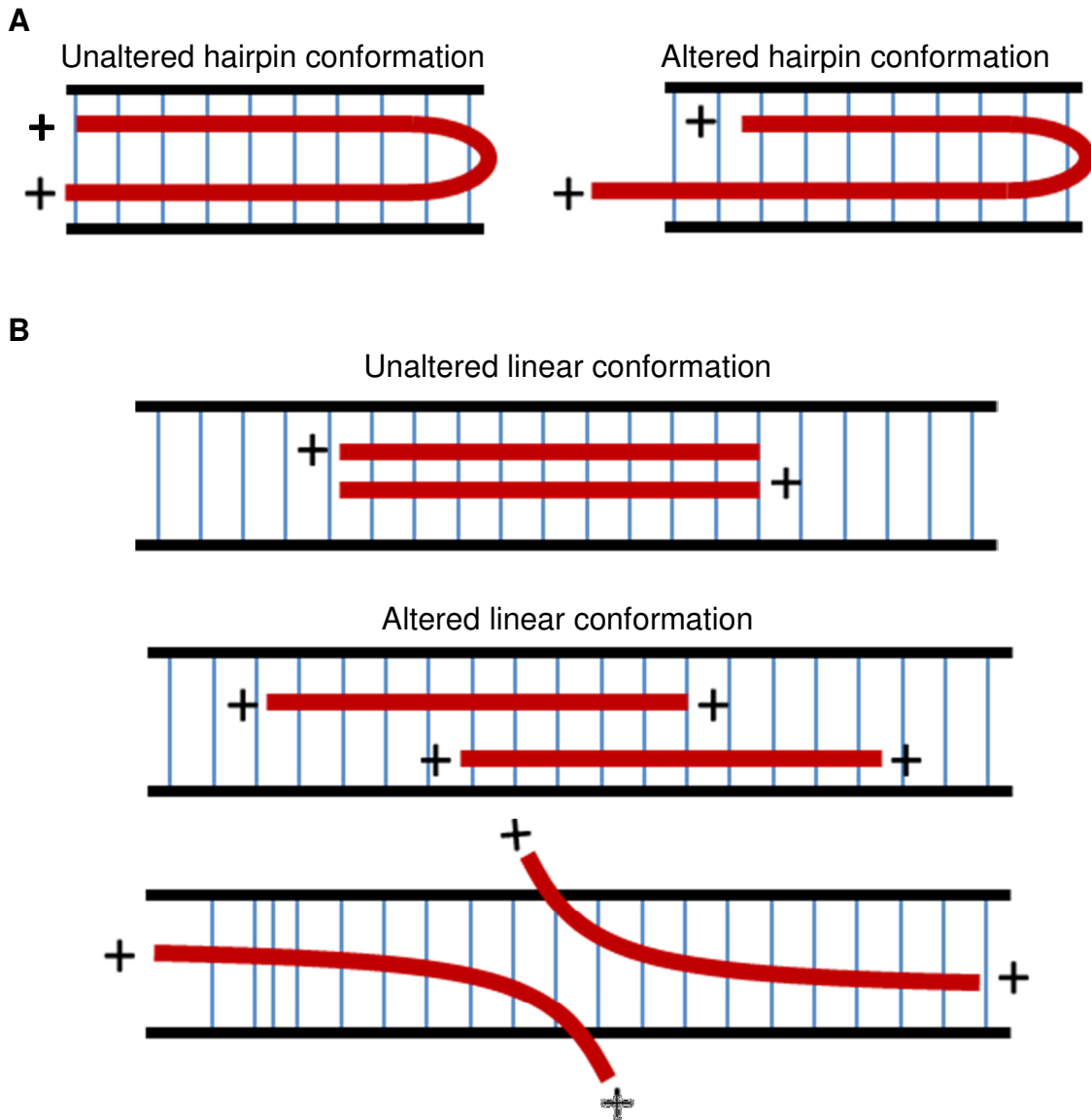
may also bind quickly to a non-specific DNA sequence then slide along the DNA to a more favorable position.

The placement of the charged guanidinyll groups at the N-terminus of the PA places a positive charge closer to the dicationic C-terminal Ta than the non-guanidinyllated parent structure (**Fig. 5.1**). Given the flexibility of both the tail and the overall hairpin structure, perhaps electrostatic repulsion forces the PA into an altered PA confirmation that is still capable of interacting with DNA, but in a different orientation and/or at a different DNA sequence than the non-guanidinyllated PA. Some examples of an altered structure would be an open an asymmetrical hairpin (**Fig. 5.9B**) or an altered linear PA-DNA complex structure (**Fig. 5.9C**).

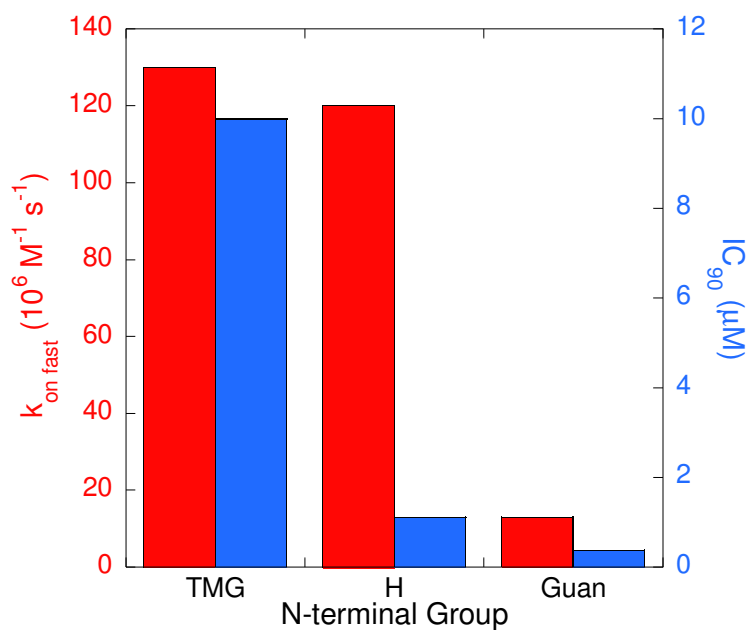
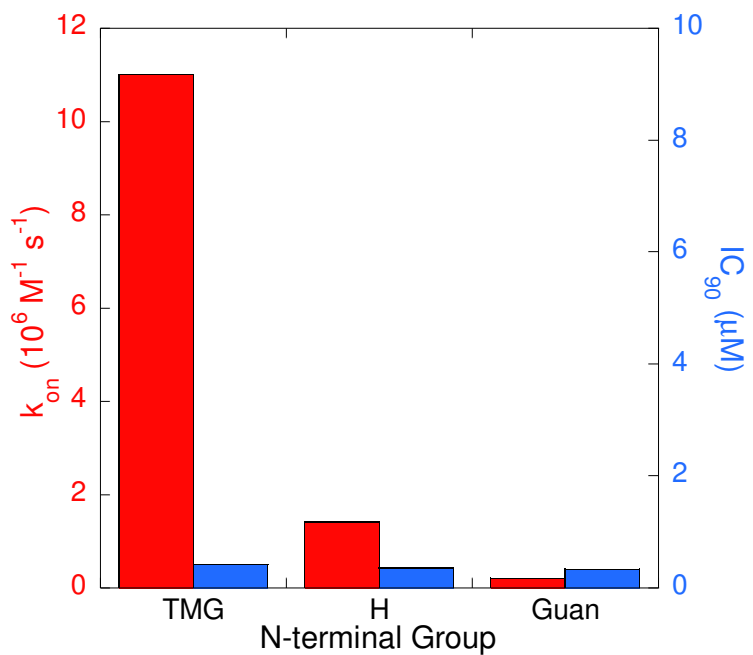
With respect to antiviral activity, there does not appear to be a consistent correlation between association rate constants of the fast phase of NV1028, nor for the monoexponential association rate constants of NV1042. For the NV1028 series, there is an inverse correlation between association rate constants and potency (**Fig. 5.10A**). However, for the NV1042 series, the association rate constants of the three PAs are significantly different from each other despite there being no difference in their antiviral activities (**Fig. 5.10B**).

### **5.3.5. N-terminal functional groups and their implications on the viral mechanism**

All PAs in the NV1028 series bound to DNA at least 2:1 despite the different N-terminal groups. However, given the differences in the concentration dependence of the DNA association it is possible that multiple equivalents bind to



**Figure 5.9** Proposed structures for repulsion-induced altered conformations of polyamides with charged N-terminal groups. A, unaltered hairpin conformation (left) versus altered hairpin conformation (right). B, unaltered linear hairpin conformation (left) versus altered linear hairpin conformation (right).

**A****B**

**Figure 5.10** Association rate constants and potency with respect to N-terminal groups for the NV1028 series, A, and NV1042 series, B

the DNA, but with slight differences in conformation or position. There is some evidence to support the claim that the guanidinylated analogs bind somewhat differently to DNA with respect to their overall geometries (Castaneda et al., 2017; Liu, B. et al., 2017).

Mechanistically, these analogs may still bind in a manner consistent with their non-guanidinylated parent structures given their similar  $IC_{50}$ 's. Additionally, given the similar  $IC_{50}$ 's, it seems likely that all NV1028 series PAs are capable of permeating the cell membrane similarly; thus, it seems less likely that these charged groups affect the uptake of these PAs into the cell. However, the significantly different  $IC_{90}$ 's of the NV1028 series indicates that, while uptake and binding affinities may not be affected, the guanidinylated groups do impart some type of antiviral benefit. Perhaps NV1028-TMG and NV1028-Guan bind to DNA in a similar manner but impact the overall DNA shape and structure more significantly. The repulsive nature of the charged groups is believed to help prevent PA aggregation and increase PA solubility in solution; when bound to the DNA, these charged groups would still continue to repulse one another and may affect the overall coverage of the DNA by the PA with a PA-DNA complex structure that is slightly different from the NV1028-DNA complex structure. A slight alteration to how the PA binds to DNA may have a bigger impact on whether and or how the DDR is elicited. Additionally, this subtle difference in PA-DNA interactions would still result in an affinity cleavage and DNase I footprinting map that would be slightly different, but still very similar.

Binding stoichiometry was unable to be determined for NV1042-Guan and NV1042-TMG so it unclear whether these PAs bind to DNA similarly to NV1042. However, some inferences can still be made with regards to the viral mechanism. Unlike the NV1028 series, the addition of the guanidinium functional groups to NV1042 has no impact on the potency. Perhaps the overall size of NV1042 nullifies the effects of (or is able to compensate for the effects of) the charged groups, making it possible for NV1042-TMG and NV1042-Guan to still bind to DNA in a manner closer to that of NV1042. As discussed previously, it is believed that the larger PAs are able to compensate for the enthalpic penalty of mismatches due to their overall number of hydrogen bond interactions. The larger PAs, especially NV1042, may be both large and flexible enough, that it is capable of altering its conformation while still maintaining a minimal number of bonds to maintain a favorable binding orientation.

#### **5.4. Concluding remarks**

To summarize, this thesis examined the DNA binding kinetics of 7 hairpin PAs. The findings support our hypothesis that better antiviral potencies tend to correlate to longer dissociation rate constants. Multiple binding events were observed for multiple PAs characterized and the results presented here should help to further elucidate the PA antiviral mechanism.

## References

- . (!!! INVALID CITATION !!! ).
- Baliga, R., Baird, E. E., Herman, D. M., Melander, C., Dervan, P. B., & Crothers, D. M. (2001). Kinetic Consequences of Covalent Linkage of DNA Binding Polyamides. *Biochem*, *40*(1), 3-8. doi:10.1021/bi0022339
- Baliga, R., & Crothers, D. M. (2000a). The Kinetic Basis for Sequence Discrimination by Distamycin A. *J Am Chem Soc*, *122*(47), 11751-11752. doi:10.1021/ja002386j
- Baliga, R., & Crothers, D. M. (2000b). On the kinetics of distamycin binding to its target sites on duplex DNA. *Proc Natl Acad Sci U S A*, *97*(14), 7814-7818.
- Baraldi, P. G., Bovero, A., Fruttarolo, F., Preti, D., Tabrizi, M. A., Pavani, M. G., & Romagnoli, R. (2004). DNA minor groove binders as potential antitumor and antimicrobial agents. *Med Res Rev*, *24*(4), 475-528. doi:10.1002/med.20000
- Barrett, M. P., Gemmell, C. G., & Suckling, C. J. (2013). Minor groove binders as anti-infective agents. *Pharmacol & Therapeut*, *139*(1), 12-23. doi:10.1016/j.pharmthera.2013.03.002
- Bashkin, J. K., Aston, K., Ramos, J. P., Koeller, K. J., Nanjunda, R., He, G., . . . Wilson, W. D. (2013). Promoter scanning of the Human COX-2 gene with 8-ring polyamides: unexpected weakening of polyamide-DNA binding and selectivity by replacing an internal N-Me-pyrrole with  $\beta$ -alanine. *Biochimie*, *95*(2), 271-279. doi:10.1016/j.biochi.2012.09.023
- Bashkin, J. K., Edwards, T. G., Fisher, C., Harris, Jr., G. D., Koeller, K. J. (2013). United States Patent No. 9,133,228. U. S. P. Office.
- Bradshaw, J. M., McFarland, J. M., Paavilainen, V. O., Bisconte, A., Tam, D., Phan, V. T., . . . Taunton, J. (2015). Prolonged and tunable residence time using reversible covalent kinase inhibitors. *Nat Chem Biol*, *11*(7), 525-531. doi:10.1038/nchembio.1817
- Buchmueller, K. L., Taherbhai, Z., Howard, C. M., Bailey, S. L., Nguyen, B., O'Hare, C., . . . Lee, M. (2005). Design of a Hairpin Polyamide, ZT65B, for Targeting the Inverted CCAAT Box (ICB) Site in the Multidrug Resistant (MDR1) Gene. *Chembiochem*, *6*(12), 2305-2311. doi:10.1002/cbic.200500179
- Burckhardt, G., Förtsch, I., Simon, H., Birch-Hirschfeld, E., Kittler, L., Schütz, H., . . . Zimmer, C. (2000). DNA Sequence Recognition of a Cross-Linked Polyamide: CD Studies, Footprinting and Effects on the Activity of DNA Gyrase. *J Biomol Struct Dyn*, *17*(sup1), 355-363. doi:10.1080/07391102.2000.10506641
- Burkhoff, A. M., & Tullius, T. D. (1988). Structural details of an adenine tract that does not cause DNA to bend. *Nature*, *331*, 455. doi:10.1038/331455a0
- Castaneda, C. H. (2017). *Biophysical Studies of Hairpin Polyamides with Broad-Spectrum Activity Against High-Risk Human Papillomaviruses*. (Doctor of Philosophy), University of Missouri - St. Louis.
- Castaneda, C. H., Scuderi, M. J., Edwards, T. G., Harris Jr, G. D., Dupureur, C. M., Koeller, K. J., . . . Bashkin, J. K. (2017). Improved antiviral activity of a



- polyamide against high-risk human papillomavirus via N-terminal guanidinium substitution. *MedChemComm*. doi:10.1039/C6MD00371K
- CDC. (1999). *Prevention of Genital HPV Infection and Sequelae: Report of an External Consultants' Meeting*. Retrieved from Atlanta: [http://www.cdc.gov/nchstp/dstd/Reports\\_Publications/99HPVReport.htm](http://www.cdc.gov/nchstp/dstd/Reports_Publications/99HPVReport.htm)
- Changeux, J. P. (2012). Allosterity and the Monod-Wyman-Changeux model after 50 years. *Annu Rev Biophys*, 41, 103-133. doi:10.1146/annurev-biophys-050511-102222
- Chen, F.-M., & Sha, F. (1998). Circular Dichroic and Kinetic Differentiation of DNA Binding Modes of Distamycin. *Biochem*, 37(32), 11143-11151. doi:10.1021/bi980950l
- Chenoweth, D. M., & Dervan, P. B. (2009). Allosteric modulation of DNA by small molecules. *Proc Natl Acad Sci U S A*, 106(32), 13175. doi:10.1073/pnas.0906532106
- Cho, J., Parks, M. E., & Dervan, P. B. (1995). Cyclic polyamides for recognition in the minor groove of DNA. *Proc Natl Acad Sci U S A*, 92(22), 10389-10392.
- Conlan, L. H., & Dupureur, C. M. (2002). Multiple Metal Ions Drive DNA Association by PvuII Endonuclease. *Biochem*, 41(50), 14848-14855. doi:10.1021/bi026403o
- Copeland, R. A. (2016). The drug-target residence time model: a 10-year retrospective. *Nat Rev Drug Discov*, 15(2), 87-95. doi:10.1038/nrd.2015.18
- Copeland, R. A., Pompliano, D. L., & Meek, T. D. (2006). Drug-target residence time and its implications for lead optimization. *Nat Rev Drug Discov*, 5(9), 730-739.
- Crothers, D. M., & Shakked, Z. (1999). Oxford Handbook of Nucleic Acid Structure. *Neidle, S., Ed*, 455-468.
- Crow, J. M. (2012). HPV: The global burden. *Nature*, 488(7413), S2-S3.
- Cusack, K. P., Wang, Y., Hoemann, M. Z., Marjanovic, J., Heym, R. G., & Vasudevan, A. (2015). Design strategies to address kinetics of drug binding and residence time. *Bioorg Med Chem*, 25(10), 2019-2027. doi:10.1016/j.bmcl.2015.02.027
- Cutts, F. T., Franceschi, S., Goldie, S., Castellsague, X., de Sanjose, S., Garnett, G., . . . Markowitz, L. (2007). Human papillomavirus and HPV vaccines: a review. *Bull World Health Organ*, 85(9).
- de Witte, W. E. A., Danhof, M., van der Graaf, P. H., & de Lange, E. C. M. (2016). In vivo Target Residence Time and Kinetic Selectivity: The Association Rate Constant as Determinant. *Trends in Pharmacol Sci*, 37(10), 831-842. doi:10.1016/j.tips.2016.06.008
- Dervan, P. B., & Edelson, B. S. (2003). Recognition of the DNA minor groove by pyrrole-imidazole polyamides. *Curr Opin Struct Biol*, 13(3), 284-299.
- Dickinson, L. A., Trauger, J. W., Baird, E. E., Ghazal, P., Dervan, P. B., & Gottesfeld, J. M. (1999). Anti-repression of RNA Polymerase II Transcription by Pyrrole-Imidazole Polyamides. *Biochem*, 38(33), 10801-10807. doi:10.1021/bi9912847

- Dupureur, C. M., Bashkin, J. K., Aston, K., Koeller, K. J., Gaston, K. R., & He, G. (2012). Fluorescence assay of polyamide-DNA interactions. *Anal Biochem*, *423*(1), 178-183. doi:10.1016/j.ab.2012.01.017
- Edwards, T. G., Helmus, M. J., Koeller, K., Bashkin, J. K., & Fisher, C. (2013a). Human papillomavirus episome stability is reduced by aphidicolin and controlled by DNA damage response pathways. *J Virol*, *87*(7), 3979-3989. doi:10.1128/jvi.03473-12
- Edwards, T. G., Koeller, K. J., Slomczynska, U., Fok, K., Helmus, M., Bashkin, J. K., & Fisher, C. (2011). HPV episome levels are potently decreased by pyrrole-imidazole polyamides. *Antiviral Res*, *91*(2), 177-186. doi:10.1016/j.antiviral.2011.05.014
- Edwards, T. G., Vidmar, T. J., Koeller, K., Bashkin, J. K., & Fisher, C. (2013b). DNA Damage Repair Genes Controlling Human Papillomavirus (HPV) Episome Levels under Conditions of Stability and Extreme Instability. *PLoS One*, *8*(10), e75406. doi:10.1371/journal.pone.0075406
- Ellenberger, T. E., Brandl, C. J., Struhl, K., & Harrison, S. C. (1992). The GCN4 basic region leucine zipper binds DNA as a dimer of uninterrupted alpha helices: crystal structure of the protein-DNA complex. *Cell*, *71*(7), 1223-1237.
- Fagan, P., & Wemmer, D. E. (1992). Cooperative binding of distamycin-A to DNA in the 2:1 mode. *J Am Chem Soc*, *114*(3), 1080-1081. doi:10.1021/ja00029a042
- Farkas, M. E., Li, B. C., Dose, C., & Dervan, P. B. (2009). DNA sequence selectivity of hairpin polyamide turn units. *Bioorg Med Chem Lett*, *19*(14), 3919-3923. doi:10.1016/j.bmcl.2009.03.072
- Finlay, A. C., Hochstein, F. A., Sobin, B. A., & Murphy, F. X. (1951). Netropsin, a New Antibiotic Produced by a Streptomyces. *J Am Chem Soc*, *73*(1), 341-343. doi:10.1021/ja01145a113
- Flores, L. V., Staples, A. M., Mackay, H., Howard, C. M., Uthe, P. B., Sexton, J. S., 3rd, . . . Lee, M. (2006). Synthesis and evaluation of an intercalator-polyamide hairpin designed to target the inverted CCAAT box 2 in the topoisomerase IIalpha promoter. *Chembiochem*, *7*(11), 1722-1729. doi:10.1002/cbic.200600155
- Gambari, R., Feriotto, G., Rutigliano, C., Bianchi, N., & Mischiati, C. (2000). Biospecific interaction analysis (BIA) of low-molecular weight DNA-binding drugs. *J Pharmacol Exp Ther*, *294*(1), 370-377.
- Garner-Hamrick, P. A., & Fisher, C. (2002). HPV Episomal Copy Number Closely Correlates with Cell Size in Keratinocyte Monolayer Cultures. *Virology*, *301*(2), 334-341. doi:10.1006/viro.2002.1558
- Halford, S. E., & Marko, J. F. (2004). How do site-specific DNA-binding proteins find their targets? *Nucleic Acids Res*, *32*(10), 3040-3052. doi:10.1093/nar/gkh624
- Hamilton, P. L., & Arya, D. P. (2012). Natural product DNA major groove binders. *Nat Prod Rep*, *29*(2), 134-143. doi:10.1039/C1NP00054C
- Han, Y. W., Kashiwazaki, G., Morinaga, H., Matsumoto, T., Hashiya, K., Bando, T., . . . Sugiyama, H. (2013). Effect of single pyrrole replacement with  $\beta$ -

- alanine on DNA binding affinity and sequence specificity of hairpin pyrrole/imidazole polyamides targeting 5'-GCGC-3'. *Bioorg Med Chem*, 21(17), 5436-5441. doi:10.1016/j.bmc.2013.06.005
- Han, Y. W., Matsumoto, T., Yokota, H., Kashiwazaki, G., Morinaga, H., Hashiya, K., . . . Sugiyama, H. (2012). Binding of hairpin pyrrole and imidazole polyamides to DNA: relationship between torsion angle and association rate constants. *Nucleic Acids Res*, 40(22), 11510-11517. doi:10.1093/nar/gks897
- Han, Y. W., Tsunaka, Y., Yokota, H., Matsumoto, T., Kashiwazaki, G., Morinaga, H., . . . Harada, Y. (2014). Construction and characterization of Cy3- or Cy5-conjugated hairpin pyrrole-imidazole polyamides binding to DNA in the nucleosome. *Biomater Sci*, 2(3), 297-307. doi:10.1039/C3BM60202H
- Harper, J. W., & Elledge, S. J. (2007). The DNA Damage Response: Ten Years After. *Mol Cell*, 28(5), 739-745. doi:10.1016/j.molcel.2007.11.015
- He, G., Tolic, A., Bashkin, J. K., & Poon, G. M. K. (2015). Heterogeneous dynamics in DNA site discrimination by the structurally homologous DNA-binding domains of ETS-family transcription factors. *Nucleic Acids Res*, 43(8), 4322-4331. doi:10.1093/nar/gkv267
- He, G., Vasilieva, E., Harris, G. D., Jr., Koeller, K. J., Bashkin, J. K., & Dupureur, C. M. (2014). Binding studies of a large antiviral polyamide to a natural HPV sequence. *Biochimie*, 102, 83-91. doi:10.1016/j.biochi.2014.02.011
- Henry, J. A., Le, N. M., Nguyen, B., Howard, C. M., Bailey, S. L., Horick, S. M., . . . Lee, M. (2004). Targeting the inverted CCAAT box 2 in the topoisomerase II $\alpha$  promoter by JH-37, an imidazole-pyrrole polyamide hairpin: design, synthesis, molecular biology, and biophysical studies. *Biochem*, 43(38), 12249-12257. doi:10.1021/bi048785z
- Herman, D. M., Baird, E. E., & Dervan, P. B. (1998). Stereochemical Control of the DNA Binding Affinity, Sequence Specificity, and Orientation Preference of Chiral Hairpin Polyamides in the Minor Groove. *J Am Chem Soc*, 120(7), 1382-1391. doi:10.1021/ja9737228
- Hurley, L. H. (2002). DNA and its associated processes as targets for cancer therapy. *Nat Rev Cancer*, 2(3), 188-200. doi:10.1038/nrc749
- Kashiwazaki, G., Chandran, A., Asamitsu, S., Kawase, T., Kawamoto, Y., Sawatani, Y., . . . Sugiyama, H. (2016). Comparative Analysis of DNA-Binding Selectivity of Hairpin and Cyclic Pyrrole-Imidazole Polyamides Based on Next-Generation Sequencing. *Chembiochem*, 17(18), 1752-1758. doi:10.1002/cbic.201600282
- Khalaf, A. I., Ebrahimabadi, A. H., Drummond, A. J., Anthony, N. G., Mackay, S. P., Suckling, C. J., & Waigh, R. D. (2004). Synthesis and antimicrobial activity of some netropsin analogues. *Org Biomol Chem*, 2(21), 3119-3127. doi:10.1039/B408386P
- Kielkopf, C. L., Baird, E. E., Dervan, P. B., & Rees, D. C. (1998a). Structural basis for GC recognition in the DNA minor groove. *Nat Struct Mol Biol*, 5(2), 104-109. doi:10.1038/nsb0298-104
- Kielkopf, C. L., White, S., Szewczyk, J. W., Turner, J. M., Baird, E. E., Dervan, P. B., & Rees, D. C. (1998b). A Structural Basis for Recognition of A·T and

- T·A Base Pairs in the Minor Groove of B-DNA. *Science*, 282(5386), 111-115. doi:10.1126/science.282.5386.111
- Kim, Y., Grable, J. C., Love, R., Greene, P. J., & Rosenberg, J. M. (1990). Refinement of Eco RI Endonuclease Crystal Structure: A Revised Protein Chain Tracing. *Sci*, 249(4974), 1307-1309.
- Koeller, K. J., Harris, G. D., Aston, K., He, G., Castaneda, C. H., Thornton, M. A., . . . Bashkin, J. K. (2014). DNA Binding Polyamides and the Importance of DNA Recognition in their use as Gene-Specific and Antiviral Agents. *Med Chem*, 4, 338-344. doi:10.4172/2161-0444.1000162
- Koshland, D. E., Jr., Nemethy, G., & Filmer, D. (1966). Comparison of experimental binding data and theoretical models in proteins containing subunits. *Biochem*, 5(1), 365-385.
- Lace, M. J., Anson, J. R., Klussmann, J. P., Wang, D. H., Smith, E. M., Haugen, T. H., & Turek, L. P. (2011). Human papillomavirus type 16 (HPV-16) genomes integrated in head and neck cancers and in HPV-16-immortalized human keratinocyte clones express chimeric virus-cell mRNAs similar to those found in cervical cancers. *J Virol*, 85(4), 1645-1654. doi:10.1128/jvi.02093-10
- Lacy, E. R., Le, N. M., Price, C. A., Lee, M., & Wilson, W. D. (2002). Influence of a terminal formamido group on the sequence recognition of DNA by polyamides. *J Am Chem Soc*, 124(10), 2153-2163. doi:10.1021/ja016154b
- Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990). The Energetic Basis of Specificity in the Eco RI Endonuclease-DNA Interaction. *Science*, 250(4982), 776-786. doi:10.1126/science.2237428
- Li, B. C., Montgomery, D. C., Puckett, J. W., & Dervan, P. B. (2013). Synthesis of Cyclic Py-Im Polyamide Libraries. *J Org Chem*, 78(1), 124-133. doi:10.1021/jo302053v
- Liu, B., Wang, S., Aston, K., Koeller, K. J., Kermani, S. F. H., Castaneda, C. H., . . . Wilson, W. D. (2017). Beta-Alanine and N-terminal cationic substituents affect polyamide-DNA binding. *Org Biomol Chem*, 15(46), 9880-9888. doi:10.1039/C7OB02513K
- Liu, Y., & Wilson, W. D. (2009). Quantitative Analysis of Small Molecule–Nucleic Acid Interactions With a Biosensor Surface and Surface Plasmon Resonance Detection. *Methods Mol Biol*, 613, 1-23. doi:10.1007/978-1-60327-418-0\_1
- Lown, J. W. (1988). Lexitropsins: rational design of DNA sequence reading agents as novel anti-cancer agents and potential cellular probes. *Anti-cancer drug design*, 3(1), 25-40.
- Lown, J. W. (1992). Lexitropsins in antiviral drug development. *Antiviral Res*, 17(3), 179-196. doi:[https://doi.org/10.1016/0166-3542\(92\)90040-C](https://doi.org/10.1016/0166-3542(92)90040-C)
- Lyng, R., Rodger, A., & Nordén, B. (1992). The CD of ligand-DNA systems. 2. Poly(dA-dT) B-DNA. *Biopolymers*, 32(9), 1201-1214. doi:10.1002/bip.360320910
- Marques, M. A., Doss, R. M., Urbach, A. R., & Dervan, P. B. (2002). Toward an Understanding of the Chemical Etiology for DNA Minor-Groove

- Recognition by Polyamides. *Helvetica Chimica Acta*, 85(12), 4485-4517. doi:10.1002/hlca.200290024
- Marverti, G., Guaitoli, G., Ligabue, A., Frassinetti, C., Monti, M. G., Lombardi, P., & Costi, M. P. (2012). Distamycin A and derivatives as synergic drugs in cisplatin-sensitive and -resistant ovarian cancer cells. *Amino Acids*, 42(2), 641-653. doi:10.1007/s00726-011-1039-3
- Monod, J., Wyman, J., & Changeux, J. P. (1965). On the nature of allosteric transitions: A plausible model. *J Mol Biol*, 12(1), 88-118. doi:10.1016/S0022-2836(65)80285-6
- Moody, C. A., & Laimins, L. A. (2009). Human Papillomaviruses Activate the ATM DNA Damage Pathway for Viral Genome Amplification upon Differentiation. *PLOS Pathog*, 5(10), e1000605. doi:10.1371/journal.ppat.1000605
- Morinaga, H., Bando, T., Takagaki, T., Yamamoto, M., Hashiya, K., & Sugiyama, H. (2011). Cysteine Cyclic Pyrrole–Imidazole Polyamide for Sequence-Specific Recognition in the DNA Minor Groove. *J Am Chem Soc*, 133(46), 18924-18930. doi:10.1021/ja207440p
- Mrksich, M., & Dervan, P. B. (1994). Design of a Covalent Peptide Heterodimer for Sequence-Specific Recognition in the Minor Groove of Double-Helical DNA. *J Am Chem Soc*, 116(8), 3663-3664. doi:10.1021/ja00087a088
- NCI. (2018). HPV and Cancer. Retrieved from <https://www.cancer.gov/about-cancer/causes-prevention/risk/infectious-agents/hpv-fact-sheet - r7>
- Neidle, S. (2001). DNA minor-groove recognition by small molecules. *Nat Prod Rep*, 18(3), 291-309. doi:10.1039/A705982E
- Nelson, S. M., Ferguson, L. R., & Denny, W. A. (2007). Non-covalent ligand/DNA interactions: Minor groove binding agents. *Mut Res*, 623(1–2), 24-40. doi:10.1016/j.mrfmmm.2007.03.012
- O'Hare, C. C., Uthe, P. B., Mackay, H., Blackmon, K., Jones, J., Brown, T., . . . Hartley, J. A. (2007). Sequence Recognition in the Minor Groove of DNA by Covalently Linked Formamido Imidazole–Pyrrole–Imidazole Polyamides: Effect of H-Pin Linkage and Linker Length on Selectivity and Affinity. *Biochem*, 46(42), 11661-11670. doi:10.1021/bi701053a
- Peixoto, P., Liu, Y., Depauw, S., Hildebrand, M. P., Boykin, D. W., Bailly, C., . . . David-Cordonnier, M. H. (2008). Direct inhibition of the DNA-binding activity of POU transcription factors Pit-1 and Brn-3 by selective binding of a phenyl-furan-benzimidazole dication. *Nucleic Acids Res*, 36(10), 3341-3353. doi:10.1093/nar/gkn208
- Pelton, J. G., & Wemmer, D. E. (1990). Binding modes of distamycin A with d(CGCAAATTTGCG)<sub>2</sub> determined by two-dimensional NMR. *J Am Chem Soc*, 112(4), 1393-1399. doi:10.1021/ja00160a016
- Pilch, D. S., Poklar, N., Gelfand, C. A., Law, S. M., Breslauer, K. J., Baird, E. E., & Dervan, P. B. (1996). Binding of a hairpin polyamide in the minor groove of DNA: sequence-specific enthalpic discrimination. *Proc Natl Acad Sci U S A*, 93(16), 8306-8311.

- Qiao, H., Ma, C., Zhang, X., Jing, X., Li, C., & Zhao, Y. (2015). Insight into DNA Minor Groove Unspecific Binding of Pyrrole Polyamide. *Bioconjug Chem*, 26(10), 2054-2061. doi:10.1021/acs.bioconjchem.5b00309
- Reddy, B. S. P., Sharma, S. K., & Lown, J. W. (2000). Recent Developments In Sequence Selective Minor Groove DNA Effectors. *Curr Med Chem*, 8(5), 475-508. doi:10.2174/0929867003373292
- Reddy, P. M., Toporowski, J. W., Kahane, A. L., & Bruice, T. C. (2005). Recognition of a 10 base pair sequence of DNA and stereochemical control of the binding affinity of chiral hairpin polyamide–Hoechst 33258 conjugates. *Bioorg Med Chem Lett*, 15(24), 5531-5536. doi:<https://doi.org/10.1016/j.bmcl.2005.08.076>
- Rucker, V. C., Melander, C., & Dervan, P. B. (2003). Influence of  $\beta$ -Alanine on Hairpin Polyamide Orientation in the DNA Minor Groove. *Helv Chim Acta*, 86(6), 1839-1851. doi:doi:10.1002/hlca.200390148
- Sakakibara, N., Mitra, R., & McBride, A. A. (2011). The Papillomavirus E1 Helicase Activates a Cellular DNA Damage Response in Viral Replication Foci. *J Virol*, 85(17), 8981-8995. doi:10.1128/jvi.00541-11
- Structure and Function of DNA. (2018, Jan 14, 2018). *Biology LibreTexts*. Retrieved from [https://bio.libretexts.org/TextMaps/Map%3A\\_Microbiology\\_\(OpenStax\)/10%3A\\_Biochemistry\\_of\\_the\\_Genome/10.2%3A\\_Structure\\_and\\_Function\\_of\\_DNA](https://bio.libretexts.org/TextMaps/Map%3A_Microbiology_(OpenStax)/10%3A_Biochemistry_of_the_Genome/10.2%3A_Structure_and_Function_of_DNA)
- Surovaya, A. N., Bazhulina, N. P., Lepehina, S. Y., Andronova, V. L., Galegov, G. A., Moiseeva, E. D., . . . Gursky, G. V. (2016). Interaction of a dimeric distamycin analog with poly(dA)poly(dT), poly[d(A–T)]poly[d(A–T)], and duplex O23 at the origin of replication of the herpes simplex virus. *Biophys*, 61(2), 227-232. doi:10.1134/s0006350916020214
- Swalley, S. E., Baird, E. E., & Dervan, P. B. (1999). Effects of  $\gamma$ -Turn and  $\beta$ -Tail Amino Acids on Sequence-Specific Recognition of DNA by Hairpin Polyamides. *J Am Chem Soc*, 121(6), 1113-1120. doi:10.1021/ja9830905
- Swinney, D. C. (2004). Biochemical mechanisms of drug action: what does it take for success? *Nat Rev Drug Discov*, 3(9), 801-808.
- Taylor Rhys, D., Asamitsu, S., Takenaka, T., Yamamoto, M., Hashiya, K., Kawamoto, Y., . . . Sugiyama, H. (2013). Sequence-Specific DNA Alkylation Targeting for Kras Codon 13 Mutation by Pyrrole–Imidazole Polyamide seco-CBI Conjugates. *Chem Eur J*, 20(5), 1310-1317. doi:10.1002/chem.201303295
- Travers, A., & Muskhelishvili, G. (2005). DNA supercoiling — a global transcriptional regulator for enterobacterial growth? *Nat Rev Microbiol*, 3, 157. doi:10.1038/nrmicro1088
- Tummino, P. J., & Copeland, R. A. (2008). Residence Time of Receptor–Ligand Complexes and Its Effect on Biological Function. *Biochem*, 47(20), 5481-5492. doi:10.1021/bi8002023
- Turner, J. M., Swalley, S. E., Baird, E. E., & Dervan, P. B. (1998). Aliphatic/Aromatic Amino Acid Pairings for Polyamide Recognition in the Minor Groove of DNA. *J Am Chem Soc*, 120(25), 6219-6226. doi:10.1021/ja980147e

- Urbach, A. R. (2011). DNA complexes: Durable binders. *Nat Chem*, 3(11), 836-837. doi:10.1038/nchem.1173
- Vasilieva, E. (2014). *Binding Properties of Large Antiviral Polyamides (PA1 and PA25) and Natural Product Derivatives as Inhibitors of Serine Hydrolases*. (Doctor of Philosophy), University of Missouri - St. Louis.
- Vasilieva, E., Niederschulte, J., Song, Y., Harris, G. D., Jr., Koeller, K. J., Liao, P., . . . Dupureur, C. M. (2016). Interactions of two large antiviral polyamides with the long control region of HPV16. *Biochimie*, 127, 103-114. doi:10.1016/j.biochi.2016.04.022
- Vilums, M., Zweemer, A. J. M., Yu, Z., de Vries, H., Hillger, J. M., Wapenaar, H., . . . Ijzerman, A. P. (2013). Structure–Kinetic Relationships—An Overlooked Parameter in Hit-to-Lead Optimization: A Case of Cyclopentylamines as Chemokine Receptor 2 Antagonists. *J Med Chem*, 56(19), 7706-7714. doi:10.1021/jm4011737
- Voet, D., Voet, J. G., & Pratt, C. W. (2013). *Fundamentals of biochemistry : life at the molecular level*. Hoboken, NJ: Wiley.
- Wang, C. C., Ellervik, U., & Dervan, P. B. (2001). Expanding the recognition of the minor groove of DNA by incorporation of beta-alanine in hairpin polyamides. *Bioorg Med Chem*, 9(3), 653-657. doi:10.1016/s0968-0896(00)00282-0
- Wang, S., Aston, K., Koeller, K. J., Harris, G. D., Jr., Rath, N. P., Bashkin, J. K., & Wilson, W. D. (2014). Modulation of DNA-polyamide interaction by beta-alanine substitutions: a study of positional effects on binding affinity, kinetics and thermodynamics. *Org Biomol Chem*, 12(38), 7523-7536. doi:10.1039/c4ob01456a
- Wang, S., Kumar, A., Aston, K., Nguyen, B., Bashkin, J. K., Boykin, D. W., & Wilson, W. D. (2013). Different thermodynamic signatures for DNA minor groove binding with changes in salt concentration and temperature. *Chem Commun*, 49(76), 8543-8545. doi:10.1039/C3CC44569K
- Wang, S., Munde, M., Wang, S., & Wilson, W. D. (2011). Minor Groove to Major Groove, an Unusual DNA Sequence-Dependent Change in Bend Directionality by a Distamycin Dimer. *Biochem*, 50(35), 7674-7683. doi:10.1021/bi201010g
- Wang, S., Nanjunda, R., Aston, K., Bashkin, J. K., & Wilson, W. D. (2012). Correlation of Local Effects of DNA Sequence and Position of  $\beta$ -Alanine Inserts with Polyamide–DNA Complex Binding Affinities and Kinetics. *Biochem*, 51(49), 9796-9806. doi:10.1021/bi301327v
- Watson, J. D., & Crick, F. H. C. (1953). Molecular Structure of Nucleic Acids: A Structure for Deoxyribose Nucleic Acid. *Nature*, 171(4356), 737-738.
- Wemmer, D. E., Geierstanger, B. H., Fagan, P., Dwyer, T. J., Jacobsen, J. P., Pelton, J. G., . . . Kollman, P. (1994). *Structural Biology: The state of the art* (R. H. Sarma & M. H. Sarma Eds.). New York: Adenine Press.
- White, C. M., Satz, A. L., Bruice, T. C., & Beerman, T. A. (2001). Inhibition of transcription factor-DNA complexes and gene expression by a microgonotropen. *Proc Natl Acad Sci U S A*, 98(19), 10590-10595. doi:10.1073/pnas.191374698

- White, S., Baird, E. E., & Dervan, P. B. (1997a). On the pairing rules for recognition in the minor groove of DNA by pyrrole-imidazole polyamides. *Chem Biol*, 4(8), 569-578.
- White, S., Baird, E. E., & Dervan, P. B. (1997b). Orientation Preferences of Pyrrole–Imidazole Polyamides in the Minor Groove of DNA. *J Am Chem Soc*, 119(38), 8756-8765. doi:10.1021/ja971569b
- Williams, V. M., Filippova, M., Soto, U., & Duerksen-Hughes, P. J. (2011). HPV-DNA integration and carcinogenesis: putative roles for inflammation and oxidative stress. *Future virology*, 6(1), 45-57. doi:10.2217/fvl.10.73
- Woods, C. R., Faucher, N., Eschgfaller, B., Bair, K. W., & Boger, D. L. (2002). Synthesis and DNA binding properties of saturated distamycin analogues. *Bioorg & Med Chem Letters*, 12(18), 2647-2650. doi:10.1016/S0960-894X(02)00467-5
- Wu, H.-M., & Crothers, D. M. (1984). The locus of sequence-directed and protein-induced DNA bending. *Nature*, 308, 509. doi:10.1038/308509a0
- Wurtz, N. R., & Dervan, P. B. (2000). Sequence specific alkylation of DNA by hairpin pyrrole–imidazole polyamide conjugates. *Chem Biol*, 7(3), 153-161. doi:10.1016/S1074-5521(00)00085-5
- Zhang, W., Jiang, S. K., Wu, Y. L., Guo, C. X., Zhang, H. F., Sugiyama, H., & Chen, X. L. (2012). Discrimination between T/A and A/T base pairs of pyrrole-imidazole polyamides substituted with chiral beta-hydroxy-gamma-aminobutyric acid/beta-alanine pairs. *Chembiochem*, 13(1), 47-50. doi:10.1002/cbic.201100675
- Zhang, W., Minoshima, M., & Sugiyama, H. (2006). Base pair recognition of the stereochemically alpha-substituted gamma-turn of pyrrole/imidazole hairpin polyamides. *J Am Chem Soc*, 128(46), 14905-14912. doi:10.1021/ja064369l