Name: Desmond Ka Wing Lau
Degree: Master of Science
Title of Thesis: The Replication Protein A function of HIN-200 family proteins is mediated by complementation of PAAD and HIN200 domains

Examining Committee:
Chair: Dr. Nicholas Harden
Assistant Professor, Department of Molecular Biology and Biochemistry

Dr. Frederic Pio
Senior Supervisor
Assistant Professor, Department of Molecular Biology and Biochemistry

Dr. Mark Paetzel
Supervisor
Associate Professor, Department of Molecular Biology and Biochemistry

Dr. Edgar Young
Supervisor
Assistant Professor, Department of Molecular Biology and Biochemistry

Dr. Lisa Craig
Internal Examiner
Assistant Professor, Department of Molecular Biology and Biochemistry

Date Defended/Approved: Dec 9th, 2008
Declaration of Partial Copyright Licence

The author, whose copyright is declared on the title page of this work, has granted to Simon Fraser University the right to lend this thesis, project or extended essay to users of the Simon Fraser University Library, and to make partial or single copies only for such users or in response to a request from the library of any other university, or other educational institution, on its own behalf or for one of its users.

The author has further granted permission to Simon Fraser University to keep or make a digital copy for use in its circulating collection (currently available to the public at the “Institutional Repository” link of the SFU Library website <www.lib.sfu.ca> at: <http://ir.lib.sfu.ca/handle/1892/112>) and, without changing the content, to translate the thesis/project or extended essays, if technically possible, to any medium or format for the purpose of preservation of the digital work.

The author has further agreed that permission for multiple copying of this work for scholarly purposes may be granted by either the author or the Dean of Graduate Studies.

It is understood that copying or publication of this work for financial gain shall not be allowed without the author’s written permission.

Permission for public performance, or limited permission for private scholarly use, of any multimedia materials forming part of this work, may have been granted by the author. This information may be found on the separately catalogued multimedia material and in the signed Partial Copyright Licence.

While licensing SFU to permit the above uses, the author retains copyright in the thesis, project or extended essays, including the right to change the work for subsequent purposes, including editing and publishing the work in whole or in part, and licensing other parties, as the author may desire.

The original Partial Copyright Licence attesting to these terms, and signed by this author, may be found in the original bound copy of this work, retained in the Simon Fraser University Archive.

Simon Fraser University Library
Burnaby, BC, Canada

Revised: Fall 2007
ABSTRACT

Our discovery of the death domain PAAD shows not only its involvement in cell death but in other biological processes. For example, we demonstrated a DNA repair function similar to Replication Protein A for the PAAD family member IFI16 (Interferon-Inducible protein 16) which also contains two HIN-200 domains by functional complementation between both individual domains. However, it is not known if the discovered ssDNA binding function is general to the HIN200 family. Here I studied MNDA (Myeloid Nuclear Differentiation Antigen), another HIN200 family member that recognizes dsDNA using similar biochemical and biophysical approaches. Our results show that the MNDA has similar but distinct DNA binding properties compared to IFI16. To further map the protein-nucleic acids interface involved in these interactions, I designed mutagenesis and chemical shift mapping experiments and show that specific amino-acids side chains of the PAAD domain are critical to ssDNA and dsDNA recognition as well as nucleic-acids stabilization.

Keywords: PAAD domain, HIN-200 domain, Replication Protein A, dsDNA, ssDNA, IFI16, MNDA
DEDICATION

To my dearest family
ACKNOWLEDGEMENTS

I would like to thank Dr. Frederic Pio for giving me the opportunity to work for him as a graduate student and thank you for your patience and allowance in teaching me how to work in the lab and the fundamentals of biochemistry. It is my honour to work with the past and current lab members, Kush Dalal, Karen Yan and Benjamin Hon. They showed enormous support and guidance throughout my master degree. I would also like to thank my supervisory committee, Dr. Mark Paetzal and Dr. Edgar Young, for their valuable comments and directions on my projects. Thanks to Dr. Sinclair for giving me endless opportunity to learn how to teach. I would like to thank my family for their everlasting support and encouragement.

Finally yet importantly, I need to thank all the faculties who have taught me in my graduate career and all my friends for their help and support namely: Dr. Brandhorst, Dr. Cornell, Dr. Craig, Dr. Thewalt, K Fan, A Yuen, K Chan, S Wu, D Ng, J Lum, S Nam, M Lee and to anyone else I have missed.
# TABLE OF CONTENTS

Approval ............................................................................................................................. ii
Abstract ............................................................................................................................... iii
Dedication ............................................................................................................................ iv
Acknowledgements ............................................................................................................ v
Table of Contents ............................................................................................................... vi
List of Figures ..................................................................................................................... ix
List of Tables ..................................................................................................................... xi
Glossary .............................................................................................................................. xii

1: Introduction ..................................................................................................................... 1
  1.1. Overview of the HIN-200 family members ......................................................... 1
  1.2. Domain organization of the HIN-200 protein ................................................. 3
  1.3. Interferon-γ inducible protein 16 (IFI16) ......................................................... 6
      1.3.1 IFI16 and cell cycle arrest ............................................................................. 6
      1.3.2 IFI16 binding partners and involvement in disease ............................... 10
      1.3.3 IFI16 is an RPA like protein ................................................................. 12
      1.3.4 The PAAD domain of IFI16 is a novel ssDNA binding domain........... 14
  1.4. Myeloid Nuclear Differentiation Antigen (MNDA) ........................................... 16
      1.4.1 MNDA expression and interaction profile ............................................. 16
  1.5. Overview of objectives ....................................................................................... 20

2: Materials and Methods .................................................................................................. 22
  2.1. Cloning and expression ...................................................................................... 22
  2.2. Mass spectroscopy (ESI and MALDI-TOF) ....................................................... 23
  2.3. Circular dichroism (CD) ................................................................................... 24
  2.4. Intrinsic tyrosine fluorescence ......................................................................... 25
  2.5. Comparative modelling ................................................................................... 25
  2.6. Electrophoretic mobility shift assay ................................................................. 25
  2.7. UV cross linking ............................................................................................... 29
  2.8. Chemical cross linking ...................................................................................... 29
  2.9. Fluorescence quenching assay ........................................................................... 29
  2.10. Double stranded DNA melting ....................................................................... 30
  2.11. Fluorescence resonance energy transfer ......................................................... 32
  2.12. Site-directed mutagenesis .............................................................................. 33
  2.13. Nuclear magnetic resonance .......................................................................... 33
2.15. Mutant expression .......................................................... 37
2.16. Far UV and thermodynamic parameters assessment of IFI16-PAAD mutants using circular dichroism .................................................. 37
2.17. Chemical stability assessment of IFI16-PAAD using circular dichroism .......................................................... 37

3: Results ............................................................................ 38

3.1 OB fold properties are conserved within the PAAD domain family .......................................................... 38

3.1.1 Understanding the spectroscopic property and stability of MNDA-PAAD .......................................................... 38
3.1.1.1 Rationale ........................................................................ 38
3.1.1.2 Expression and purification of MNDA-PAAD .......................................................... 38
3.1.1.3 Secondary structure determination using Circular Dichroism .......................................................... 39
3.1.1.4 Thermodynamic parameters of the PAAD domain .......................................................... 45
3.1.1.5 Intrinsic fluorescence – tertiary structure unfolding of MNDA-PAAD .......................................................... 49
3.1.1.6 Polydispersity determination of MNDA-PAAD using dynamic light scattering .......................................................... 52
3.1.1.7 Comparative modelling reveals disordered region in IFI16-PAAD .......................................................... 54

3.1.2 The OB fold binding properties of MNDA .......................................................... 56
3.1.2.1 Rationale ........................................................................ 56
3.1.2.2 MNDA-PAAD binds to nucleic acids .......................................................... 57
3.1.2.3 The PAAD domain forms a stable complex with ssDNA as determined by UV cross linking, mass spectroscopy and western blot .......................................................... 63
3.1.2.4 MNDA-PAAD binding constant determination using fluorescence quenching assay .......................................................... 76
3.1.2.5 MNDA-PAAD does not compact or extend ssDNA by Fluorescence Resonance Energy Transfer (FRET) .......................................................... 83
3.1.2.6 PAAD domains destabilize dsDNA by hyperchromicity assay .......................................................... 88
3.1.2.7 Oligomerization of the PAAD domain determine ssDNA binding property .......................................................... 90
3.1.2.8 MNDA-H120 also has RPA like function .......................................................... 93

3.2 Chemical shift mapping of MNDA-PAAD with ssDNA .......................................................... 107

3.3 Correlation of IFI16-PAAD’s structure and stability to dsDNA destabilization property .......................................................... 113
3.3.1 Rationale ........................................................................ 113
3.3.2 Identification of IFI16-PAAD mutations with higher stability and secondary structure content by GFP directed evolution .......................................................... 113
3.3.3 IFI16-PAAD mutants with increased stability and secondary structure confer enhanced dsDNA destabilization .......................................................... 116
4: Discussion ......................................................................................................................... 122
4.1 Different structural properties and stability within the PAAD domain family .................. 122
4.2 Comparison of nucleic acid binding properties between IFI16 and MNDA ...................... 124
4.3 IFI16-PAAD mutants with different thermodynamics signature obtained by directed evolution can modulate nucleic acids destabilization ................................................. 130
5: Conclusion and Future Direction ....................................................................................... 133
Appendices ......................................................................................................................... 137
Appendix A: EMSA curve fitting parameters of MNDA-PAAD ............................................ 137
Appendix B: Tyrosine fluorescence quenching curve fitting parameters of MNDA-PAAD .................. 142
Appendix C: Tryptophan fluorescence quenching curve fitting parameters of MNDA-HIN200 ................................................................. 152
Appendix D: Circular dichroism spectra of IFI16-PAAD mutants with increased secondary structure ............................................................... 158
Appendix E: Circular dichroism spectra of IFI16-PAAD mutants with decreased secondary structure ............................................................... 159
Appendix F: Thermal denaturation of IFI16-PAAD mutants with increased thermostability ................................................................. 160
Appendix G: Thermal denaturation of IFI16-PAAD mutants with decreased thermostability ................................................................. 161
Appendix H: Chemical denaturation of IFI16-PAAD mutants .............................................. 162
Appendix I: IFI16-PAAD disorder region prediction .......................................................... 163
Appendix J: MNDA-PAAD 2D HSQC titration spectra .......................................................... 166
Reference List ...................................................................................................................... 169
LIST OF FIGURES

Figure 1.1. Domain organization of the human HIN-200 family members IFI16, MNDA, AIM2 and the newly identified IFIX. ........................................... 4

Figure 1.2. Schematic representation of the involvement of IFI16 in cell cycle arrest ................................................................................. 9

Figure 1.3. Modelling of PAAD domain of IFI16 suggest that MNDA has a six helix bundle motif but IFI16 has a disordered region .......... 19

Figure 3.1. Superdex 75, SEC of MNDA-PAAD. .......................................................... 42

Figure 3.2. Cation exchange chromatography of IFI16-PAAD. .............................. 43

Figure 3.3. CD spectra of the PAAD domains .................................................. 44

Figure 3.4. Thermal denaturation of the PAAD domain ........................................ 47

Figure 3.5. MNDA-PAAD did not regain secondary structure after thermal denaturation .......................................................... 48

Figure 3.6. Tyrosine spectroscopic property of MNDA-PAAD. ............................ 51

Figure 3.7. Dynamic light scattering of MNDA-PAAD ........................................ 53

Figure 3.8. Superimposition of IFI16-PAAD and MNDA-PAAD showing the six helix bundle and disordered helix region. ....................... 55

Figure 3.9. MNDA-PAAD binds to different single stranded DNA by EMSA ................................................................................. 61

Figure 3.10. Binding curves and one site binding $K_D$ calculations from EMSA experiments of MNDA-PAAD .................................................. 62

Figure 3.11. UV cross linking of MNDA-PAAD with ssDNA .................................. 64

Figure 3.12. Gel calibration curve for molecular weight determination of UV cross linking species of MNDA-PAAD with ssDNA. ................. 65

Figure 3.13. Mass Spectroscopy of the PAAD domain with ssDNA .................. 70

Figure 3.14. Western blot of the PAAD domain with ssDNA ............................ 72

Figure 3.15. Electrostatic potential of MNDA-PAAD .......................................... 74

Figure 3.16. Length dependence of MNDA-PAAD with ssDNA by UV cross linking ................................................................................. 75

Figure 3.17. Dissociation constant of MNDA-PAAD with nucleic acids determined by tyrosine fluorescence quenching ......................... 82
Figure 3.18. Oligonucleotides design to study compaction or extension of ssDNA by MNDA using FRET ................................................................. 85

Figure 3.19. Fluorescence resonance energy transfer suggests that MNDA-PAAD does not compact or extend single stranded DNA. ...................................................................................... 86

Figure 3.20. Melting depression of dsDNA by MNDA-PAAD ....................... 89

Figure 3.21. Chemical cross linking of MNDA-PAAD to polythymine 70 .......... 92

Figure 3.22. SDS-PAGE analysis of MNDA-HIN200 ..................................... 97

Figure 3.23. MNDA-HIN200 recognizes ssDNA by EMSA ............................ 99

Figure 3.24. UV cross linking of MNDA-HIN200 with ssDNA ...................... 100

Figure 3.25. Dissociation constant determination of MNDA-HIN200 by Tryptophan fluorescence quenching .................................................. 103

Figure 3.26. Chemical cross linking of MNDA-HIN200 to T70 .................... 105

Figure 3.27. MNDA-HIN200 compact and extend ssDNA by FRET ............... 106

Figure 3.28. MNDA-PAAD recognizes ssDNA in the fast/medium exchange rate by chemical shift mapping ......................................................... 110

Figure 3.29. ssDNA destabilizes MNDA-PAAD structure by thermal denaturation ....................................................................................... 111

Figure 3.30. Relative chemical shift change in the chemical shift mapping of MNDA-PAAD ........................................................................ 112

Figure 3.31. Modelling of IFI16-PAAD mutant with increased secondary structure and stability ............................................................... 115

Figure 3.32. Different thermodynamic signatures of IFI16-PAAD mutants modulate dsDNA destabilization .................................................. 118
## LIST OF TABLES

Table 1-1. The HIN200 domain motifs. ................................................................. 5
Table 2-1. Oligonucleotides used in this study ...................................................... 28
Table 3-1. Experimental molecular weight determination of MNDA-PAAD-nucleic acid complexes after SDS-PAGE analysis and UV cross linking with ssDNA ................................................................. 66
Table 3-2. Nucleic acid binding parameters of MNDA-PAAD by tyrosine fluorescence quenching. ................................................................. 79
Table 3-3. Nucleic acid binding parameters of MNDA-HIN200 by tryptophan fluorescence quenching. ................................................................. 104
Table 3-4. Thermodynamic parameters of MNDA-PAAD in the absence and presence of ssDNA. ................................................................. 111
Table 3-5. Secondary structure and thermodynamic parameters of IFI16-PAAD mutants after directed evolution. ................................................................. 119
Table 3-6. Thermodynamics and chemical stability parameters of four IFI16-PAAD mutants. ................................................................. 120
Table 3-7. Summary of RPA nucleic acid binding properties for the PAAD and HIN200 domain of MNDA and IFI16 ................................. 121
GLOSSARY

AIM2  Absent in Melanoma protein 2
Apaf1  Apoptotic Proteases Activation Factor 1
ΔASA  Change in Accessible Surface Area upon denaturation
ASC  Apoptosis associated Speck like protein with Caspase Recruitment Domain
ATM  Ataxia Telangiectasia Mutated protein kinase
BASC  BRCA1 Associated Genome Surveillance Complex
BLM  Bloom protein helicase
B_{max}  Maximum Binding constant
bp  base pair
BRCA1  Breast Cancer Associated protein 1
CARD  Caspase Associated Recruitment Domain
CD  Circular Dichroism
CDK  Cyclin Dependent Kinase
DBD  DNA Binding Domain
DD  Death Domain
DED  Death Effector Domain
DLS  Dynamic Light Scattering
EMSA  Electrophoretic Mobility Shift Assay
ESI  Electro spray Ionization
FRET  Fluorescence Resonance Energy Transfer
GFP  Green Fluorescence Protein
HIN-200  Hematopoietic Interferon inducible Nuclear with 200 amino acids repeats
HL-60  Human Promyelocytic Leukemia cells 60
HSQC  Heteronuclear Single Quantum Correlation
IFI16  Interferon Inducible protein 16
IFN  Interferon
IFI-X  Interferon-gamma Inducible protein X
JAK1  Janus Kinase 1
Kd  Dissociation constant
LPS  Lipopolysaccharide
MALDI-TOF  Matrix Assisted Laser Desorption/Ionization – Time Of Flight Mass
MNDAG  Myeloid Nuclear Differentiation Antigen
MS  Spectroscopy
NALP1  NACHT, LRR and PYD-containing protein 1
NF-KB  Nuclear Factor of immunoglobulin k locus in B cells
NLS  Nuclear Localization Signal
NMR  Nuclear Magnetic Resonance
Ni-NTA  Nickel – Nitrilotriacetic acid
OB fold  Oligonucleotide/Oligosaccharide Binding fold
<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>P21</td>
<td>Tumor Protein 21</td>
</tr>
<tr>
<td>P53</td>
<td>Tumor Protein 53</td>
</tr>
<tr>
<td>PAAD</td>
<td>Pyrin, Aim, Asc, Death domain like</td>
</tr>
<tr>
<td>PAGE</td>
<td>Polyacrylamide Gel Electrophoresis</td>
</tr>
<tr>
<td>PMSF</td>
<td>Phenylmethanesulphonylfluoride</td>
</tr>
<tr>
<td>pRB</td>
<td>Retinoblastoma protein</td>
</tr>
<tr>
<td>RFC</td>
<td>DNA Replication Factor C</td>
</tr>
<tr>
<td>RPA</td>
<td>Replication Protein A</td>
</tr>
<tr>
<td>SDS</td>
<td>Sodium Dodecyl Sulphate</td>
</tr>
<tr>
<td>SEC</td>
<td>Size exclusion chromatography</td>
</tr>
<tr>
<td>SSB</td>
<td>Single Stranded DNA Binding protein</td>
</tr>
<tr>
<td>STAT</td>
<td>Signal Transducers and Activator of Transcription protein</td>
</tr>
<tr>
<td>TAE</td>
<td>Tris Acetate EDTA buffer</td>
</tr>
<tr>
<td>TBE</td>
<td>Tris Borate EDTA buffer</td>
</tr>
<tr>
<td>TFE</td>
<td>2,2,2 Trifluoroethanol</td>
</tr>
<tr>
<td>TNFα</td>
<td>Tumor necrosis factor α</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet light</td>
</tr>
</tbody>
</table>
1: INTRODUCTION

1.1. Overview of the HIN-200 family members

Since the discovery of the first murine gene of the HIN-200 family, \textit{Ifi202a}, in 1982 (Samanta, Dougherty, Brawner, Schmidt, & Lengyel, 1982), many new members have been characterized. The nomenclature of HIN-200 was initially proposed according to its hematopoietic expression, interferon induction, nuclear localization, and a characteristic domain of 200 amino acids repeats. Currently there are five murine and four human HIN-200 family members that all share common structural motifs and are interferon inducible. Murine HIN-200 proteins include p202a, p202b, p203, p204 and p205, while human HIN-200 members include IFI16 (Interferon-Inducible protein 16), MNDA (Myeloid Nuclear Differentiation Antigen), AIM2 (Absent In Melanoma 2) and IFIX (Interferon-Inducible protein X). The HIN-200 family members have similar structural domains, a PAAD/DAPIN/Pyrin domain at their N-termini with the exception of p202a and p202b, followed by either one or two copies of the HIN200 domain except IFIXy isoform that lacks any HIN200 domain. The inductive ability of HIN-200 family members by interferon suggests that they are involved in antiviral and immunomodulatory activities because interferon is a type of cytokine, which is produced by the cell in response to viral infection or abnormal cell growth such as tumor (Choubey, Deka, & Ho, 2008). There is substantial evidence in the literature that link HIN-200 proteins to diseases. For example, IFI16 expression is absent in
many tumor samples such as prostate carcinoma (Xin, Curry, Johnstone, Nickoloff, & Choubey, 2003) and malignant breast epithelial cells (Fujiuchi et al., 2004) suggesting that IFI16 may be a tumor suppressor (Raffaella et al., 2004; Xin et al., 2003). More studies on the HIN-200 family members revealed that they share similar hematopoietic expression, nuclear localization and a characteristic domain of 200 amino acids (HIN200 domain) whose structure and functions are unknown. All three human HIN-200 members are localized on human chromosome 1q21-23 and their close proximity in syntenic loci suggests that the family members arose by gene duplication (R. C. Briggs et al., 1994; DeYoung et al., 1997; Trapani et al., 1992). Some exceptions to the characteristics of HIN-200 proteins discussed above include the gene Ifi202a that is constitutively expressed in mice’s spleen lacking interferon receptors indicating that it is interferon independent (H. Wang et al., 2002). Also Ifi205 gene expression is induced by agents other than interferon such as lipopolysaccharide (LPS) and tumor necrosis factor α (TNFα) (Hamilton, Bredon, Ohmori, & Tannenbaum, 1989). Moreover, IFI16 is found in epithelial cells of the skin and ducts of breast tissue suggesting HIN-200 expression in non-hematopoietic system (Gariglio et al., 2002; Wei et al., 2003). HIN-200 proteins were thought to be localized exclusively in the nucleus as mediated by a recognized nuclear localization signal (NLS) (L. J. Briggs et al., 2001), but MNDA and AIM2 both lack the NLS. Even with the lack of consistency in HIN-200 protein classification, there is increasing evidence pointing toward the involvement of HIN-200 protein in regulating cell proliferation and differentiation.
1.2. Domain organization of the HIN-200 protein

HIN-200 proteins contain one or two repeats of a C-terminal 200 amino acid domain (HIN200 domain) as well as an 85-100 amino acid PAAD domain at their N-termini. The PAAD domain is a subfamily of the Death domain super family that modulates inflammation and apoptosis (Figure 1.1). There are three different types of HIN200 domain, α, β, and γ, distinguished by their amino acid sequence and specific motifs (Table 1-1). Some HIN-200 proteins have their domains separated by a spacer region with unknown function but vary in overall sequence and size. The HIN-200 proteins have several highly conserved peptide motifs such as the MFHATVAT motif, located in the HIN200 domain and proposed to be involved in protein-protein interaction (Koul, Obeyesekere, Gutterman, Mills, & Choubey, 1998). LXCXE is another highly conserved motif located on the HIN200 domain for the binding of phosphorylated retinoblastoma protein, a tumor suppressor protein that helps regulates cell cycle (Choubey & Lengyel, 1995). More details about this motif will be discussed in the subsequent section (Section 1.3.1).
Figure 1.1. Domain organization of the human HIN-200 family members IFI16, MNDA, AIM2 and the newly identified IFIX.
<table>
<thead>
<tr>
<th>Residue</th>
<th>α domain</th>
<th>β domain</th>
<th>γ domain</th>
</tr>
</thead>
<tbody>
<tr>
<td>14-20</td>
<td>EYESPEX</td>
<td>TYDXXEX</td>
<td>E(C/F)ETQEG</td>
</tr>
<tr>
<td>32-39</td>
<td>M(F/L)HATVA(T/S)</td>
<td>M(F/L)HATVAT</td>
<td>(I/M)F HATVAT</td>
</tr>
<tr>
<td>33-28</td>
<td>X(S/T)QYFH</td>
<td>ETEFFR</td>
<td>ETDFFF</td>
</tr>
<tr>
<td>63-71</td>
<td>(E/K)XXGILE(I/V)N</td>
<td>GCNGFLEIY</td>
<td>(R/W)HSXFXEV(T/N)</td>
</tr>
<tr>
<td>98-103</td>
<td>XTPKIX</td>
<td>ATPKIS</td>
<td>ETPKIS</td>
</tr>
<tr>
<td>152-158</td>
<td>IXC(E/K)(E/K)GD</td>
<td>(I/V)XCEPG(D/X)</td>
<td>(T/M)KC(E/K)EGD</td>
</tr>
<tr>
<td>162-168</td>
<td>LFCF(H/R)L(R/K)</td>
<td>L(V/F)CFEL(T/S)</td>
<td>LTEF(E/T)(V/L)S</td>
</tr>
<tr>
<td>176-187</td>
<td>LV(C/S)GXHFSIKX</td>
<td>LRSVRHSYMQV</td>
<td>LKSGX(C/H)SXXKV</td>
</tr>
</tbody>
</table>

Table 1-1. The HIN200 domain motifs.
Different motifs used to differentiate HIN-200 domain. Amino acids (K,R), (I,L), (D,E), (Y,F), (V,A) are considered equivalent. Bolded amino acids are common to all subclasses and underlined amino acids can be used to discriminate between them (Ludlow, Johnstone, & Clarke, 2005).
1.3. Interferon-γ inducible protein 16 (IFI16)

1.3.1 IFI16 and cell cycle arrest

Cell cycle arrest is a well-known host defense mechanism against viral infection and since HIN-200 proteins are interferon induced, it provides a link for the HIN-200 proteins to cell cycle arrest. There are two types of interferon identified that modulate biological activities during viral infection, Type I (IFN-α and IFN-β) and Type II (IFN-γ). A well known mechanism describing how interferons mediate transcriptional activation during host defense was determined by Stark (Stark, 2007). Briefly, interferon is a cytokine produced by infected cells during the anti-viral response; the interferons are secreted into the cytoplasm of the cell to promote the anti-viral response of the neighboring cells. Interferon binds to the corresponding cell surface receptor and activates tyrosine kinase (JAK1). Once JAK1 is phosphorylated, it activates signal transducer and activator of transcription protein (STAT) to initiate the transcription of their target genes such as IFI16. A well-studied HIN-200 family member, IFI16, is found to be involved in gene regulation and cell cycle arrest (Fujiuchi et al., 2004; Raffaella et al., 2004; Xin et al., 2003). Typically, IFI16 arrests cell growth in the G1 to S phase transition with the E2F-mediated transcription pathway involving regulatory proteins such as p53, pRb, c-Myc and cyclin A (Ludlow et al., 2005). In the normal human cell, cell cycle progression in the E2F pathway is led by activated cyclin-dependent kinase 2 (cdk), which phosphorylates pRb and turns pRb into its hyper-phosphorylated state (ppRb). The hyper-phosphorylated pRb (ppRb) is no longer able to bind to and inhibits transcription factor E2F; therefore, E2F become active and cell cycle progresses. Certain stimuli such as DNA
damage often causes cell cycle arrest via p53 dependent pathway, when p53 and p21 are over-expressed by response to a stimulus. The tumor suppressor transcription factor p53 activates other regulatory proteins such as p21, a cyclin-dependent kinase inhibitor protein. When p21 is active, it inhibits the phosphorylation activity of cdk, resulting in pRb reverting to its hypophosphorylated state. In this state, pRb is able to bind to transcription factor E2F and represses it, and cells lacking E2F cannot progress from G1 to S phase during cell cycle (Figure 1.2). IFI16 has been shown to interact with most of the proteins mentioned above in the E2F pathway, namely Rb, E2F, p21 and p53. Xin et al. in 2003 found that IFI16 is able to specifically bind to Rb and E2F1 by GST pull-down assay (Xin et al., 2003). It has been established that HIN-200 proteins possess at least one conserved LXCXE motif, which is a putative Rb binding motif. However IFI16 contains a closely related IXCXE motif but Mahnaghi and co-workers suggested that the closely related motifs are involved in binding to Rb (Magnaghi-Jaulin et al., 1998; Trapani et al., 1992). In an attempt to find a link between the function of IFI16 and its mouse homolog p202, Johnstone et al. (Johnstone, Wei, Greenway, & Trapani, 2000) revealed that IFI16 also binds transcription factor p53 by immuno-precipitation and enhances the DNA binding activity of p53. Interestingly, besides the ability to form protein-protein interaction, IFI16 also has a putative DNA binding site and its DNA binding ability was demonstrated by Dawson and Trapani (Dawson & Trapani, 1995). The DNA binding ability of IFI16 was further characterized by Dalal et al. and Yan et al. In their studies, they tested individual domains of IFI16 for binding and concluded that the HIN200 domain of IFI16 DNA binding property behaves like RPA.
(replication protein A), another single stranded DNA binding protein containing the OB fold (oligonucleotides/oligosaccharide binding) (Yan et al., 2008).
Figure 1.2. Schematic representation of the involvement of IFI16 in cell cycle arrest (Ludlow et al., 2005).
1.3.2 IFI16 binding partners and involvement in disease

There are many studies showing the importance of IFI16 in human cancer. In 2003, Fujiuchi showed that normal human epithelial cells expressed detectable amounts of IFI16 protein, whereas in breast cancer patients the level of IFI16 mRNA and protein were relatively decreased (Fujiuchi et al., 2004). The loss of IFI16 in breast epithelial cells was thought to be the cause of breast cancer. Congruently in the same year, Xin showed the human prostate cancer cells did not express or have reduced levels of IFI16 compared to the normal prostate epithelial cells (Xin et al., 2003). In the same study, the authors showed that IFI16 is linked to cellular senescence by over expressing functional IFI16 in prostate cancer cells, revealing that cancer cells colony formation was suppressed, and the production of senescence-associated β-galactosidase, a biochemical marker for cellular senescence, was activated. Moreover, IFI16 was found to be involved in DNA repair as IFI16 interacts with BRCA1 in BASC (BRCA1 Associated Surveillance protein Complex) (Aglipay et al., 2003). BRCA1 locus encodes a large, multifunctional nuclear phosphoprotein that has an N-terminal RING finger and C-terminal BRCT domain (Bork et al., 1997; Koonin, Altschul, & Bork, 1996; Miki et al., 1994). The BRCA1 and associated proteins are called BASC and by mass spectrometry, Wang discovered BASC contains tumor suppressor and DNA repair proteins, namely MSH2, MSH6, MLH1, ATM, BLM, and DNA replication factor C (RFC), all of which are able to interact with DNA (Wang et al., 2000). Interestingly, every member of the BASC could also recognize abnormal DNA structure and damaged DNA, suggesting this complex is a sensor for DNA damage. Aglipay et al. revealed that the PAAD domain of IFI16 is responsible
for the binding to BRCA1 (Aglipay et al., 2003). Since every member in BASC is able to interact with nucleic acids, it is anticipated that IFI16 also possess DNA binding property. Not surprisingly, the N-terminus of IFI16 is able to bind to double stranded DNA but the binding property requires further characterization (Choubey & Gutterman, 1996; Dawson & Trapani, 1995; Johnstone, Kerry, & Trapani, 1998; Luu & Flores, 1997).
1.3.3 IFI16 is an RPA like protein

Full length IFI16 has the ability to bind to nucleic acids but each individual domain remains to be investigated. Comparative modelling of IFI16 predicted by our laboratory and others (Albrecht, Choubey, & Lengauer, 2005) showed that the two HIN200 domains (α and β) contain two OB folds and this was later confirmed by x-ray crystallography (PDB code: 2OQ0, 3B6Y). Although the HIN200 domain has two OB folds, the sequence identity between the HIN200 domain and the OB fold of RPA is less than 10% (Theobald, Mitton-Fry, & Wuttke, 2003). If the sequence identity between two proteins is less than the critical threshold of 25%, the function of these proteins may be very different. However, since the typical OB fold contains many proteins that perform single stranded binding function in spite of having little sequence identity; it is possible that HIN-200 proteins also have single stranded binding function. Moreover, the structural similarity between the HIN200 domain and RPA allows us to consider if the HIN200 domain exhibits similar OB fold function. These properties remain to be determined for this family.

RPA contains an OB fold and plays a critical role in preventing DNA being damaged and ensuring the repair process is executed properly (Wold & Kelly, 1988). As a ssDNA binding protein, it keeps the damaged DNA in its single stranded form and prevents it from re-annealing until the mismatch repair process is completed (Binz, Lao, Lowry, & Wold, 2003). RPA is also able to interact with other DNA processing proteins and adopt different binding modes in order to facilitate the repairing process (Fanning, Klimovich, & Nager, 2006). For example, human RPA70 displays OB fold properties
including oligomerization via its OB folds. This changes the binding mode of RPA to nucleic acids, with stronger preference to ssDNA and RNA over dsDNA as well as the length of the oligonucleotides (Bochkareva, Belegu, Korolev, & Bochkarev, 2001; Kim, Paulus, & Wold, 1994; Theobald et al., 2003). Most importantly, it also has the ability to destabilize and unwind dsDNA to initiate DNA replication (Brosh et al., 2000; Theobald et al., 2003; Treuner, Ramsperger, & Knippers, 1996). Recently a study from Yan illustrated some key features of RPA displayed in the HIN200 domain of IFI16 but further investigation is needed (Yan et al., 2008). Yan showed the HIN200 domain binds preferably to ssDNA with G,C rich oligonucleotides as assayed by fluorescence quenching. Yan also showed the HIN200 domain could oligomerize upon binding to ssDNA as well as compact and extend the ssDNA. However as most of the experiments in Yan’s paper suggest, IFI16 HIN200 domain is a nucleic acid binding protein with RPA-like OB fold, but a critical property is missing. If the assumption of IFI16 is a DNA repair protein like RPA, it should at least bind ssDNA and destabilize dsDNA as those properties are often found on DNA repair proteins. Nevertheless, IFI16 HIN200 domain was found to slightly stabilize dsDNA instead of destabilizing it. Despite IFI16-HIN200 domain lacks the property to destabilize duplex DNA, it well resembles an RPA protein and points toward it being another DNA repair protein. The lack of the property to destabilize duplex DNA raised the possibility that other domains of IFI16 can destabilize dsDNA.
1.3.4 The PAAD domain of IFI16 is a novel ssDNA binding domain

The PAAD domain is another conserved domain found throughout the HIN200 family, which also belongs to the death domain super family. Every member of the death domain super family has a conserved domain composed of 85-100 amino acids that is involved in protein-protein interactions to execute apoptosis and initiate the inflammation pathway (Liu, Rojas, Ye, & Godzik, 2003). Members of the death domain super family include the death domain (DD), the death effector domain (DED), the caspase recruitment domain (CARD) and the Pyrin, AIM, ASC, Death domain like (PAAD). X-ray crystallography revealed that the death domains have a conserved 6 helix bundle folded into a Greek key motif (Liang & Fesik, 1997). Each member with this typical structural motif is found to dimerize or interact with other members of the same sub-family. For example, the CARD domain of caspase-9 is able to dimerize when inactive as well as interact with the CARD domain of Apaf-1 during caspase activation, a well regulated apoptosis mechanism via protein-protein interaction (Bao & Shi, 2007). However, as there is no structure solved for the PAAD domain of IFI16, comparative modelling predicted that it has a six helix bundle motif compatible with other death domain with the exception of a disordered helix 3 (Dalal, 2006). The function of the PAAD domain sub-family is well characterized and as a member of the Death Domain superfamily, it is not surprising that it has been shown to be involved in protein-protein interactions during apoptosis and inflammation (Pawlowski, Pio, Chu, Reed, & Godzik, 2001). As discussed in Section 1.3.2, the PAAD domain interacts with other cell cycle regulatory factors causing cell cycle arrest via protein-protein interactions. However in
addition to these protein-protein interactions, it has recently been shown that the PAAD domain of IFI16 could also interact with nucleic acids (Dalal, 2006). The DNA binding characteristics of the PAAD domain are similar to the HIN200 domain of IFI16. Both domains recognize single stranded DNA equally or better than double stranded DNA and are able to oligomerize upon binding to ssDNA. The PAAD domain is able to compact ssDNA and Dalal et al. proposed a model where multiple copies of the PAAD domain are wrapped by the ssDNA (Dalal, 2006). A very important aspect of the PAAD domain in interacting with nucleic acid is that it can destabilize duplex DNA, which is the key property of RPA’s OB fold that is missing in the HIN200 domain. Here we show evidence that the lack of the function of destabilization in the HIN200 domain maybe complemented by the PAAD domain of IFI16. It is a good example of a protein that functions as a whole by combining functions from different domains. A very interesting feature of the PAAD domain of IFI16 is the partially folded structure. As mentioned previously, PAAD domains have similar structure compared to other death domains except for NALP1, which is another member of the PAAD domain sub-family but possess only five helices. Helix 3 in NALP1 is disordered and therefore it cannot fold properly into an α-helix. The model of the PAAD domain of IFI16 used the NALP1 NMR structure as the best template according to MODELLER (Dalal & Pio, 2006). Dalal has shown that the PAAD domain of IFI16 adopts a partially disordered structure as characterized by thermodynamics and circular dichroism (Dalal & Pio, 2006). The same author suggested that the partially folded structure might confer the ability to bind nucleic acids. This hypothesis can be validated and we can look at the conservation of these properties in
the HIN-200 family by comparing their nucleic acid binding properties with another PAAD domain family member, MNDA.

1.4. Myeloid Nuclear Differentiation Antigen (MNDA)

1.4.1 MNDA expression and interaction profile

MNDA belongs to the HIN-200 family as it contains one copy of the HIN-200 domain, the α domain. MNDA also belongs to the death domain super family because it has a PAAD domain in its N-terminus and potentially links to other death domains involved in signalling pathways in apoptosis and inflammation (R. C. Briggs et al., 2006). A unique characteristic of MNDA that is not found in other HIN-200 family members is its pattern of expression. MNDA expression is restricted specifically to the hematopoietic system such as granulocytes and monocytes, as previous studies showed that MNDA is involved in myeloid differentiation but not expressed in lymphoid cells (R. Briggs et al., 1994). MNDA expression has been characterized using human leukemic cell line HL-60 (human promyelocytic leukemia cells), which Goldberger and his colleagues have found that MNDA expression increased during granulocyte differentiation triggered by retinoic acid but not found in less differentiated HL-60 cells (Goldberger, Brewer, Hnilica, & Briggs, 1984), indicating lineage specific expression. Cousar and his colleagues went even further to prove that MNDA is involved in myeloid differentiation. Cousar showed that MNDA expression correlates with myeloid differentiation status using human acute leukaemia cells, where MNDA expression gets consistently stronger as the myeloblastic cells mature (Cousar & Briggs, 1990). It is important to note that MNDA is only expressed in hematopoietic
cell lines but not non-hematopoietic cells, in contrast to IFI16 expression, which is detected in both cell lines.

MNDA possesses most of the structural properties that are found within the HIN-200 family. The NMR structure of the N-terminal PAAD domain is solved (2DBG) and determined to have the 6 helix bundle in its monomeric form, and the full length protein is able to dimerize via its imperfect leucine zipper motif (Xie, Briggs, & Briggs, 1997). The PAAD domain of IFI16 and MNDA share similar structure with roughly 56% sequence identity. The MNDA PAAD domain has a well-folded, highly helical content with an imperfect leucine zipper whereas IFI16 PAAD domain has a partially folded structure. The model of the PAAD domain of IFI16 illustrates that it has a disordered region between helix 2 and 4 generating a large loop but the PAAD domain of MNDA has a complete helix 3. Structural superimposition of the two PAAD domains clearly reveals the disordered region (Figure 1.3). Since the loop between the two helices has an important role in binding to ligands such as the helix-loop-helix motif of transcription factors that binds to nucleic acids. It may suggest a difference of protein-ligand interactions between the PAAD domain of IFI16 and MNDA. Furthermore, MNDA is found to bind to nuclear proteins such as nucleolin (C23) and nucleophosmin (Xie, Briggs, Morris et al., 1997), which interact with transcription factor YY1. Xie showed that upon binding of MNDA to YY1, the DNA binding activity of YY1 is enhanced. The same author also showed that MNDA can bind directly to YY1 independent of the MNDA or YY1 interaction with nucleolin and nucleophosmin and only the PAAD domain of MNDA is sufficient to enhance YY1 DNA binding activity (Xie, Briggs, & Briggs, 1998). The ability of MNDA to interact with transcription
factors and nucleic acids has opened a new area of research for MNDA regarding transcriptional modulation, while the mechanism of how MNDA regulates the affinity of transcription factors to nucleic acids remains elusive. Surprisingly, the recent discoveries of IFI16 as a single stranded nucleic acid binding protein (Yan et al., 2008) and MNDA as a transcriptional modulator (Xie et al., 1998; Xie, Briggs, Olson, Sipos, & Briggs, 1995) suggest these two proteins share a common property since they both interact with nucleic acids. However, their function may diverge as IFI16 was shown to be involved in DNA repair but no literature up to date has suggested MNDA involvement in DNA repair. A detailed examination of the nucleic acid binding property of IFI16 and MNDA may shed light on their respective functional role as a nucleic acid binding protein.
Figure 1.3. Modelling of PAAD domain of IFI16 suggest that MNDA has a six helix bundle motif but IFI16 has a disordered region.

This is the superimposition of the NMR structure of MNDA PAAD (2DBG; Blue) and model of IFI16 PAAD (Orange) built from NALP1 PAAD (1PN5) as the template. The superimposition was performed using Pymol align command. RMSD of the superimposition is 3.9Å.
1.5. Overview of objectives

The PAAD domain sub-family members have different structural properties compared to other Death domain super family proteins and this deviation may lead to different functions. The recent discovery that the PAAD domain of IFI16 has ssDNA binding properties agrees with our hypothesis that the PAAD domain may adopt many different ligand interactions. Since among the different 3D structure of PAAD domain family members there are different levels of disorder in their unstructured region, their ability to interact with other ligands may be reflective of their thermodynamic signatures. Moreover, our data (Yan et al., 2008) showing that IFI16-HIN200 is an RPA like protein that exhibits most of the OB fold DNA binding characteristics has strengthened the link between IFI16 and DNA repair as previously suggested (Aglipay et al., 2003; Choubey & Gutterman, 1996; Choubey & Lengyel, 1995). The fact that IFI16 has similar functions to RPA leads us to investigate whether these properties are conserved within the HIN200 family. To determine if the OB fold properties are conserved within the HIN200 family, especially the ability to destabilize DNA duplex, and to identify the structural/stability difference between members of the PAAD domain, we ask the following questions.

1. Are the newly identified OB fold-like properties of IFI16 conserved within the HIN200 family?

2. What are the amino acid side chains involved in the single stranded nucleic acids binding in the HIN200 family?

3. What is the correlation between double stranded nucleic acid destabilization and secondary structure and stability of the PAAD domain?
To address these questions, we chose the PAAD domain of MNDA for the structural and thermodynamic comparison with the PAAD domain of IFI16. Circular dichroism, fluorescence and dynamic light scattering were used to characterize the structural properties and stability of MNDA-PAAD. The DNA binding properties of the PAAD and HIN200 domain were examined by Electrophoretic Mobility Shift Assay (EMSA), UV cross linking and tyrosine fluorescence quenching. Protein-protein complex formation was studied using chemical cross linking. Fluorescence Resonance Energy Transfer (FRET) was used to determine if the protein is able to compact or extend single stranded DNA. More importantly, the ability to destabilize DNA duplex was determined using melting depression experiments. Moreover, Nuclear Magnetic Resonance (NMR) was used to investigate whether specific amino acids are responsible for the interaction with ssDNA. Finally, we investigated partially folded structure of IFI16-PAAD by GFP (Green Fluorescence Protein) directed evolution and circular dichroism.
2: MATERIALS AND METHODS

2.1. Cloning and expression

The DNA fragments that encode the amino acids sequence of the PAAD domain of MNDA (1-90, NP_002423) and IFI16 (1-102, CAI15085) and HIN200 domain of MNDA (214-398, NM_002432) were amplified using PCR and inserted into expression vector pET28b via 5' BamH1 and 3' Ndel cut site. The HIN200 domain of MNDA (205-398, NP_002423) was also PCR amplified and inserted into pET28b but via 5'Nde1 and 3' EcoR1 restriction sites. The proteins were produced as a fusion protein that fused with a 6x histidine tag in their N-terminus using bacterial Escherichia coli BL21(DE3) as their host strain. The culture was induced with 1 mM IPTG when the O.D$_{600nm}$ reached 0.6 and grown for 4 hours at 37°C. The cells were pelleted by centrifuging at 4000 rpm with JLA10.5 (Beckman) at 4°C for 20 minutes. The cell pellet was resuspended in lysis buffer (50 mM NaH$_2$PO$_4$, 300 mM NaCl, 10 mM imidazole, 100 µM PMSF, pH 8.0) and sonicated at approximately 20 Volts. The lysate was further centrifuged at 30000 xg for 1 hour at 4°C and the cleared lysate was purified according to the QIA-expressionist protocol (QIAGen) using Ni-NTA beads. The proteins were allowed to bind on to the beads, washed with 25 mM imidazole, and eluted with 250 mM imidazole. The proteins after elution were immediately subjected to dialysis overnight with the buffer that gave the optimal solubility to avoid precipitation and aggregation.
The PAAD domain of IFI16 was dialyzed against 50 mM sodium acetate, 14.4 mM β-mercaptoethanol, pH 4.0 and both MNDA-PAAD and MNDA-HIN were dialyzed with 5 mM Tris, 100 mM NaCl, 14.4 mM β-mercaptoethanol, pH 8.0. IFI16-PAAD was subjected to a second round of purification using cation exchange. The running buffer used was 50 mM sodium acetate, pH 4.0 and a gradient from 0 M to 2 M NaCl over 100 minute was used to elute the protein. Eluted proteins were immediately dialyzed with 50 mM sodium acetate, 14.4 mM β-mercaptoethanol, pH 4.0 and stored at -80°C. MNDA-PAAD was further purified using size exclusion chromatography with Superdex 75 (Amersham). The running buffer was 5 mM Tris, 100 mM NaCl, pH 8.0. 14.4 mM β-mercaptoethanol was added after the protein has been eluted and the protein was stored in -80°C. All proteins were concentrated to approximately 8mg/ml using an Amicon Ultra centrifugation filter with a 5K Dalton molecular weight cut off.

2.2. Mass spectroscopy (ESI and MALDI-TOF)

The mass spectra were collected from two sources, Chemistry Department in Simon Fraser University and Genome BC Proteomics Center in University of Victoria. The ESI (Electrospray ionization) was performed in SFU using Varian 4000 GC/MS/MS and the MALDI-TOF was performed in UVic using Applied Biosystems 4800 Voyager MALDI-TOF/TOF mass spectrometers. In the ESI experiment, approximately 50 mM of IFI16-PAAD was loaded with 5 mM of sodium acetate, pH 4.0 and trace amount of β-mercaptoethanol in a total volume of 20 ul. The MALDI-TOF used to detect protein-DNA complex used 140 pM of IFI16-PAAD and 150 pM of MNDA-PAAD. The same molar concentration of ssDNA was UV cross linked (same
as section 2.7) to the protein prior to mass spectroscopy, T25 and T13, respectively. The matrix used was sinipinic acid and cyano-4-hydroxycinnamic acid (CHCA). Some complexes were prepared and concentrated using C18 Zip-tip and eluted using 1 μl of 90% acetonitrile and the 1μl was loaded onto the matrix.

2.3. Circular dichroism (CD)

JASCO-J-810 spectropolarimeter equipped with a peltier type PFD-425S constant temperature cell holder was used to collect all CD spectra of PAAD and HIN200 domains. IFI16-PAAD was diluted to 40 μM using 5 mM sodium acetate pH 4.0 while MNDA-PAAD and MNDA-HIN200 were diluted to 46 μM and 22 μM, respectively, using 5 mM Tris 100 mM NaCl, pH 8.0 for far UV measurement. Far UV was measured from 260 nm to 190 nm at 200 nm/min scan rate with 100 mdeg sensitivity and 0.1s response time. The temperature was kept constant at 20°C over the whole process to minimize stability fluctuation and a 0.05 cm path length quartz cell was used. The units obtained from the raw measurement were converted into mean residue ellipticity and expressed in deg cm$^2$ decimol$^{-1}$. A negative control with buffers only was subtracted from all spectra to correct the baseline. The secondary structure assessment was performed using cdPRO with three different algorithms, CDSSTR, CONTINLL, and SELCON3. CDSSTR was liable to all secondary structure assessment due to its high accuracy with $\alpha$ helix prediction of proteins with known structures. cdPRO gave an assessment of $\alpha$ helical content with two different categories, regular helix and disordered helix. The percent $\alpha$ helix in our results were the sum of the two categories.
2.4. Intrinsic tyrosine fluorescence

Tyrosine fluorescence was obtained using SLM4800 spectrofluorimeter equipped with a single wavelength monochromator. 7 μM of MNDA-PAAD in 5 mM Tris, 100 mM NaCl, and pH 8.0 was used in the experiment. The protein was excited at 275 nm (λex of tyrosine) and the emission spectrum was monitored from 250 nm to 350 nm. Denatured protein sample contained the same initial material with an addition of 6 M urea. Denatured protein sample was incubated overnight at 25°C to allow equilibrium to establish at standard temperature. The cuvette path length was 1 cm and the data were collected at a gain of 100, 600 Volts with bandwidths of 8nm in all four slits. A baseline was subtracted from both the native and denatured sample with the respective buffers.

2.5. Comparative modelling

The comparative model of IFI16 was built by Dalal based on the template from NALP1 (PDB code 1PN5) (Dalal, 2006). The NMR structure of MNDA-PAAD (PDB code 2DBG) was superimposed on the model of IFI16-PAAD using the align command in PyMol.

2.6. Electrophoretic mobility shift assay

Prior to perform electrophoretic mobility shift assay (EMSA), oligonucleotides were labelled with $^{32}$P using T4 kinase from Invitrogen. The Oligonucleotides used in EMSA, UV cross linking and fluorescence quenching are listed in Table 2-1. The labelling reaction composed of 500 pmol oligonucleotides in a 1x forward kinase buffer, 50 μCi of $^{32}$P γ-ATP and 10
units of T4 kinase. The labelling reaction took 1 hour at 37°C and was stopped by heating the sample to 65°C for 10 minutes or adding 1 mM EDTA. The reaction mixtures were purified by loading into a denaturing polyacrylamide gel consisting of 8 M urea with 20% acrylamide and 1x TBE. The gel ran at 250 Volts until the dye front reached 1/5 from the bottom of the gel. UV shadowing was used to verify the position of the oligonucleotides the band was excised, crushed and soaked in TE buffer with 300 mM NaCl at 37°C overnight. Gel debris was removed from the solution and the eluted DNA was cleaned by ethanol precipitation. Air dried oligonucleotides were re-suspended in 1x TE buffer and the concentration of the labelled oligonucleotides was approximately 5-10 μM. To prepare for the shift assay, a two fold serial dilutions of MNDA-PAAD and MNDA-HIN200 from 210 μM (2.1 nmol) to 0.83 μM (8.3 pmol) using 20 mM HEPES, pH 7.5 were made and each dilution was incubated with a fixed amount of labelled probe (360 nM / 3.6 pmol) for 30 minutes at room temperature. The negative control having the probe only along with other samples were loaded into a 10%, 1x TBE native polyacrylamide gel and ran for 1 hour at 100 Volts. The gel was disassembled from the gel apparatus and wrapped with saran wrap, placed in a cassette and exposed to a phosphor screen (Amersham) overnight. The phosphor screen was scanned using a Typhoon 9410 variable mode imager (Molecular Dynamics) with highest sensitivity and varied resolution (200-50). The images were loaded into ImageQuant software version 2.2 for an integration of the shifted band to obtain the relative intensity. The intensity of the shifted bands were corrected with the background and the relative value of each band was converted into fraction bound. The fraction bound was
plotted against the concentration of the protein in the corresponding fraction and the plot was analyzed using GraphPad Prism version 4.03.
<table>
<thead>
<tr>
<th>Name of oligonucleotides</th>
<th>Sequence</th>
</tr>
</thead>
<tbody>
<tr>
<td>T25</td>
<td>TTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT</td>
</tr>
<tr>
<td>A25</td>
<td>AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA</td>
</tr>
<tr>
<td>GC - 5</td>
<td>5'-GGAAGAAGGAAGTGATGGGATCAGGATCCGCTGGCTCC-3'</td>
</tr>
<tr>
<td>GC - 3</td>
<td>5'-GGAGCCAGCGGGATCCTGATCCACTTGGTTGTTC-3'</td>
</tr>
</tbody>
</table>

Table 2-1. Oligonucleotides used in this study
2.7. UV cross linking

All UV cross linking of proteins to nucleic acids were conducted in the same fashion. 2 μg of proteins were individually mixed with 0, 0.1, 0.2, 0.5, 1, 2, 5 molar concentration of nucleic acids in 20 mM HEPES, pH 7.5 and the mixtures were incubated at room temperature for 10 minutes. The mixtures were put into the Stratalinker 1800 cross linker and irradiated with UV light at 254 nm at a distance of 20 cm for 15 minutes at 1000 J/min. The samples were loaded into a 15% SDS-PAGE gel and ran for 50 minutes at 200 Volts. The gels were visualized by silver staining. The ladder used in each gel was PAGE ruler from Fermantas and the relative mobility of each ladder in each gel was plotted in a calibration curve (Log molecular weight vs. Rf). In each gel, the molecular weight of each species was identified by measuring the Rf and plotted in the corresponding calibration curve.

2.8. Chemical cross linking

Protein-protein interactions were promoted using two different cross linker, glutaldehyde and formaldehyde. 10 μg of proteins were mixed with 0, 0.1, 0.2, 0.5, 1, 2, 5 molar concentration of T25 and T70 oligonucleotides in 20 mM HEPES, pH 7.5. The total volume of the mixture was 200 μl, they were allowed to incubate at room temperature for 10 minutes, and 8 μl of 25% w/v chemical cross linker was added into the mixtures. The cross linking reactions were incubated for 2 minutes and stopped by adding 10 μl of freshly prepared quenching buffer (2 M NaBH₄, 0.1 M NaOH) and incubated for 20 minutes. Nine μl of trichloroacetic acid (78% w/v) was added to precipitate the proteins for 5 minutes at 4°C and the solution was centrifuged in a table top centrifuge.
at 13000 rpm for 10 minutes. Pellets were washed with ice cold acetone two times and centrifuged in between washes. The pellets were air dried and re-suspended using 1x laemmli buffer and loaded into 15% SDS-PAGE gel. Gels were ran at 200 Volts for 1 hour and silver stained for visualization.

2.9. Fluorescence quenching assay

A tyrosine quenching assay was performed using a SLM4800 spectofluorimeter equipped with a single wavelength monochomator. Seven μM of MNDA-PAAD in 5 mM Tris, 100 mM NaCl, and pH 8.0 was used in the experiment. MNDA-PAAD was titrated with single stranded DNA A25, T25, GC-5, GC-3 and double stranded DNA AT25, and GC5-3 in 20 mM HEPES, pH 7.5. Titration started at 0 molar concentrations of nucleic acids and increased up to 4 molar concentrations. 0.1 molar concentration increments were used from 0 – 1 and 1 molar concentration increments were used from 1 – 4. Each sample’s relative tyrosine fluorescence emission was measured by first exciting the sample with $\lambda_{\text{ex}}$ 275 nm and monitoring the emission spectra from 250 nm to 350 nm. Each curve was plotted and the relative value at $\lambda_{\text{em}}$ 304 nm was selected and converted to percentage quenched, the percentage quenched values were plotted against the concentration of nucleic acids titrated and curve fitting was done using this plot with GraphPad Prism version 4.03.

The plot was fitted into either a one site binding model ($Y = \frac{B_{\text{max}}X}{(K_D+X)}$) or a two site binding model ($Y = \frac{B_{\text{max1}}X}{(K_{D1}+X)} + \frac{B_{\text{max2}}X}{(K_{D2}+X)}$) using non-linear regression fit, where $B_{\text{max}}$ correspond to maximal binding and $K_D$ correspond to dissociation constant. An F-test was
performed to determine the preferred binding model. The detail of the statistical analysis is listed in GraphPad PRISM manual (Motulsky, 2003).

### 2.10. Double stranded DNA melting

The guanine and cytosine rich oligonucleotides were selected to test the ability of the proteins to destabilize the DNA duplex. GC-5 and GC-3 are single stranded oligonucleotides that are complementary to each other. They were annealed by mixing equal molar concentrations of each oligonucleotides in an annealing buffer (20 mM HEPES, 133 mM NaCl, pH 7.5), heated at 95°C for 10 minutes, then gradually cooled down to 20°C at a rate of 1°C per minute. A Cary 300 Bio UV-Visible spectrophotometer equipped with a temperature regulator was used to measure all melting depression experiments, which monitored the absorbance at 260 nm from 20°C to 100°C at 1°C per minute. Ten μM (1 nmol) of dsDNA was mixed with 0, 1, 2, 3, 4, 5 molar concentration of proteins in each individual run. The hyperchromicity of the dsDNA was measured at UV 260 nm at each titration. The absorbance at 260 nm that contributed to the hyperchromicity of the dsDNA was corrected by subtracting to the absorbance at 260 nm of the blank containing protein only. The absorbance at 260 nm was converted into fraction denatured by making the assumption that the lowest value correspond to all double stranded forms and the highest value correspond to all single stranded form. The melting temperature was estimated by plotting the 50% fraction denatured corresponding value.
2.11. Fluorescence resonance energy transfer

Fluorescence resonance energy transfer (FRET) was used to monitor the change in ssDNA conformation, with specially designed FRET oligonucleotides used to demonstrate the extension or compaction of the ssDNA. A schematic representation of the FRET oligo is in Figure 3.18.

Briefly, an 18 base pair oligonucleotides is labelled with an acceptor fluorophore at its 5'end (sequence: 5' QUASAR670 – GCCTCGCTGCGTGC CA-3') and a 58 base pair oligonucleotides is labelled with a donor fluorophore at its 3' end (sequence: 5' TGGCGACGGCAGCGAGGC-(T)₄₀ – QUASAR570 T3') labelled oligos were generated at the University of Calgary DNA Core Services. The two oligonucleotides are partially complementary to each other with the exception of the stretch of polydT₄₀ tail on the 3' labelled oligo. The fluorescence property of the donor QUASAR 570 has a λ_ex = 548 nm and λ_em = 567 nm, while the acceptor QUASAR 670 has a λ_ex = 648 nm and λ_em = 667 nm. Equal molar concentrations of the two oligonucleotides are mixed in an annealing buffer (20 mM Tris, 400 mM NaCl, pH 8.0) and annealed by heating to 95°C for 10 minutes and gradually cool down to 20°C at 1°C per minute. The annealing of the DNA was confirmed by running each single stranded form with the annealed product in 15% native gel. The gel was ethium bromide stained and visualized (data not shown). The FRET measurements were monitored by exciting the sample at 548 nm (λ_ex of donor) and the emission spectrum was monitored from 500 nm to 700 nm. The cuvette path length was 1 cm and the data were collected at a gain of 100, 600 Volts with bandwidths of 8 nm in all four slits. A baseline was subtracted from all spectra with the
respective buffers. 300 nM of annealed FRET oligonucleotides in 20 mM HEPES, pH 7.5 was used and it was titrated with increasing molar concentration of protein with 1 molar concentration increment from 0 - 20.

2.12. Site-directed mutagenesis

A cysteine residue was mutated to a serine in MNDA-PAAD to remove the single free cysteine that potentially increases the amount of aggregation. The MNDA-PAAD mutant was generated as described in the QuickChange® II Site-Directed Mutagenesis Kit manual. The sequence of the mutant was verified by sequencing at Macrogen (Korea). Forward primer (5' - GCGTTGCCTCTCTAGACAAACTAA TAGAACTTGCCAAAGATATG - 3'), reverse primer (5' - CATATCTTTGGCAAGTTCTATTAGTTTGTCTAGAGGCAACGC - 3').

2.13. Nuclear magnetic resonance

NMR spectroscopy is a common approach used to investigate protein-nucleic interaction by looking in the binding interface and binding affinity. NMR spectroscopy provides structural information by determining the chemical environment of the magnetic nuclei and correlates the chemical shift from nuclei to nuclei. In the context of protein-ligand interactions, the receptor protein is often labelled with stable isotope such as $^{15}$N or $^{13}$C. The binding of the ligand alters the binding site chemical environment and therefore the chemical shift of the particular nuclei near the binding site are perturbed. These perturbations are best seen in a correlation spectra such as $^{13}$C/$^1$H or $^{15}$N/$^1$H, known as 2D HSQC (Heteronuclear Single Quantum Correlation) chemical shift mapping. In this case, ligands can slowly titrate into the
receptor protein and after each titration, a full scan of 2D HSQC is obtained. As a result, superimposing the titrated 2D HSQC will reveal any cross peaks being perturbed. Generally, there are two types of perturbation, fast exchange (weak affinity) or slow exchange (strong affinity). If the ligand binding to the receptor is in a slow exchange process compared to the chemical shift in frequency units, the perturbation of the chemical shift will appear as two distinct sets of resonances for the free and bound ligands. On the other hand, if the process is fast exchange, a single set of NMR line resonating at the average chemical shift of the free and bound ligands. For simplicity, $K_D < 10^{-5}$ is considered as fast exchange, $K_D = 10^{-6}$ to $10^{-7}$ is intermediate between fast or slow exchange, and $K_D$ smaller than $10^{-8}$ is considered as slow exchange (Pellecchia, Sem, & Wuthrich, 2002).

The cysteine free MNDA-PAAD construct was transformed into BL21(DE3) and the pre-culture was grown in LB media overnight to allow it to reach stationary phase. Cells were pelleted by brief centrifugation in an Eppendorf 5810R centrifuge with rotor A-4-62 at 4000 rpm (~5000 xg) for 20 minutes. 5 ml of fresh LB was used to re-suspend the pellet and subsequently used to inoculate into M9 media with $^{15}$N labelled ammonium chloride. 1 mM IPTG was added to induced protein production when O.D reached 0.2 and the induction took 4-8 hours at 37°C. Cells were lysed and purified as described in section 2.1.

The NMR spectrum with protein only was recorded at 25°C on a Bruker Avance-600 NMR spectrometer. The reference point of the 2D HSQC on the Bruker 600MHz was set using a $^{15}$N labelled T4 lysozyme. 1.7 mg/ml (0.13 mM) of MNDA-PAAD placed in 5 mM Tris, 100 mM NaCl, pH 8, 10% $D_2$O and
subjected to $^{1}H/^{15}N$ 2D HSQC using water flip-back pulse. The NMR spectra of the protein titration with ssDNA were recorded at 25°C on a Varian Inova 500 NMR spectrometer ran VNMRJ 1.1D, with 4 channels, equipped with an XYZ-gradient or Z-gradient HCN 5mm probe. Pulse sequence was used as listed in BioPack by professor Lewis Kay (Kay, Keifer, & Saarinen, 1992). 5mM HCN was used to set as the reference point for the 2D HSQC. 0.21 mM of MNDA-PAAD was titrated with 0.1, 0.2, 0.3, and 0.4 molar concentration of polydT13 ssDNA and each titration the 2D HSQC was recorded.

2.14. Flow diagram for *in vitro* GFP directed evolution of IFI16-PAAD

IFI16-PAAD was selected to perturb its structure to observe the effects on folding, stability and DNA binding properties. The wild type IFI16-PAAD was introduced to random point mutation using error prone PCR. The mutants will be fused to a C-terminus GFP protein and the brighter fluorescence mutant will be selected to further DNA shuffling to recombine the mutation and eliminates non-essential mutation. Dalal completed the mutagenesis with 3 rounds of DNA shuffling and I have completed the mutants’ expression and thermodynamic measurements.
Error prone PCR

PCR both wild type and PAAD mutant

DNA digestion and reassembly

Insert DNA fragment into pET28b expression vector

Transform mutant library into bacterial cell strain BL21(DE3)

Insert into GFP fusion vector

Isolate plasmid DNA from 40 brightest colonies

Express mutant proteins and select 40 brightest colonies

Insert into GFP fusion vector, transform into BL21(DE3)

PCR of reassembled templates

Express protein from 3x shuffled DNA and select 40 brightest colonies

Isolate plasmid from each colony, excise PAAD DNA fragment using NdeI and BamHI

Express PAAD mutants, protein purification, stability and structural analysis

3 rounds DNA shuffling
2.15. Mutant expression

Expression of IFI16-PAAD mutants generated from directed evolution are the same as listed in section 2.1. Only affinity chromatography using Ni-NTA was performed to purify the proteins.

2.16. Far UV and thermodynamic parameters assessment of IFI16-PAAD mutants using circular dichroism

Same as section 2.3.

2.17. Chemical stability assessment of IFI16-PAAD using circular dichroism

Same as section 2.3.
3: RESULTS

3.1 OB fold properties are conserved within the PAAD domain family

3.1.1 Understanding the spectroscopic property and stability of MNDA-PAAD

3.1.1.1 Rationale

The PAAD domain of IFI16 was thought to have a disordered helix 3 by comparative modelling (Dalal & Pio, 2006). Details about the secondary structure and stability of the protein were determined using circular dichorism by Dalal (Dalal & Pio, 2006). Dalal concluded that IFI16-PAAD follows a two state folding mechanism but possess a partially folded structure with an overall lower stability of the protein compared to other death domain family members. The structure of MNDA-PAAD has been solved (2DBG) and MNDA-PAAD has a well ordered helix 3. By comparing these two PAAD domains, we can gain insight into the relationship between IFI16-PAAD’s stability and folding of helix 3.

3.1.1.2 Expression and purification of MNDA-PAAD

The human DNA sequence of MNDA-PAAD and IFI16-PAAD were PCR amplified and cloned into expression vector pET28-b. The PAAD proteins were fused to a 6x-histidine tag at the N-terminus and expressed in *Escherichia coli* bacteria BL21(DE3). The fusion proteins were first purified using Ni-NTA column. However, the purity of the proteins was not optimal
because other contamination bands could be seen on a 15% SDS-PAGE. Therefore, each protein was subjected to second round of purification. The PAAD of IFI16 was subjected to cation-exchange chromatography and the PAAD of MNDA was subjected to size exclusion chromatography. Other purification strategies were used in the second round of purification but the chosen one gave us the maximum recovery of the proteins. The purity of the proteins after two rounds of purification was estimated by SDS-PAGE. Figure 3.1 illustrated the size exclusion chromatography and purity of MNDA-PAAD. The expected molecular weight of MNDA-PAAD is 12.9 kDa and IFI16-PAAD is 15.7 kDa. A sharp peak was observed near 14 ml of elution and the fractions containing this peak were collected and determined to be the correct size of MNDA-PAAD by SDS-PAGE. On the other hand, cation exchange chromatogram of IFI16-PAAD is shown in Figure 3.2. Similarly, the largest peak collected at ~33 ml of elution with 50% NaCl (1 M) corresponds to IFI16-PAAD as determined by SDS-PAGE. The SDS-PAGE gels of the two purified protein are displayed as an insert in the chromatograms.

3.1.1.3 Secondary structure determination using Circular Dichorism

The structure of MNDA-PAAD as determined by NMR spectroscopy and IFI16-PAAD by comparative modelling, consists of 6 helices linked with loops, the helices are arranged in anti-parallel fashion recognized as a Greek key motif. The secondary structure of a protein can be determined using circular dichroism and this technique can be used to make a comparison between structural features of the two PAAD domains. Circular dichroism uses a circularly polarized light beam to radiate a protein solution. The circularly polarized light consists of equal magnitudes of left polarized and
right polarized light. One of the two types will absorb more than the other and cause a deviation from linear polarized light, the magnitude of the deviation is converted into molar ellipticity and the value of the molar ellipticity at various wavelengths is what is used to determine the secondary structure of the protein. Circular dichroism requires sample to be chiral for absorption of the polarized light and amino acids are chiral molecules. Proteins have three fundamental secondary structure motifs, which are α-helix, β-strand and random coil. Each type of secondary structure has a distinct pattern of circular dichroism spectrum measuring from 260 nm to 190 nm (far UV range). The PAAD domains contain only α-helix and random coil, where α-helix gives two local minima at 222 nm and 208 nm, and random coil gives a local maximum at 216 nm and a local minimum at 198 nm. Here, far UV spectra were made between 190 nm and 260 nm at 20°C. The protein concentration used in the far UV experiment was 46 μM and the degree of ellipticity was transformed into mean residue ellipticity by this formula $\Theta \times \frac{100 \times Mr}{c \times l \times Na}$, which $\Theta$ is the raw value from the CD spectra, Mr is the molecular weight of the protein, c is the concentration of the protein, l is the light path length and Na is the number of amino acids. The respective far UV spectra of MNDA and IFI16 PAAD were superimposed and it clearly showed that there were two local minima at 222 nm and 208 nm in both spectra indicating the protein contains high helical content (Figure 3.3). However, the CD amplitude of IFI16-PAAD is shallower than MNDA-PAAD denoting less helical content. The secondary structure content of the proteins was estimated by cdPRO (Sreerama & Woody, 2000). This program calculates the amount of secondary structure based on 3 individual algorithms - CDSSTR, SELCON3, and CONTINLL. All three
individual algorithms agreed that MNDA-PAAD contains ~60% \( \alpha \)-helix and that IFI16-PAAD has ~40% \( \alpha \)-helix.
Figure 3.1. Superdex 75, SEC of MNDA-PAAD.

The peaks were annotated by taking the elution volume and plotted in the calibration curve to obtain the molecular weight of the largest peak. MNDA-PAAD eluted at 14.6 ml and a single sharp peak was observed. The peak fractions were collected and the sample was run into a 15% SDS-PAGE gel. The gel is shown as an insert.
Figure 3.2. Cation exchange chromatography of IFI16-PAAD.

IFI16-PAAD is subjected to second round of purification using cation exchange chromatography. The majority of the protein eluted in 33 ml with 50% NaCl (=1 M) and this fraction is collected and determined to be IFI16-PAAD.
Figure 3.3. CD spectra of the PAAD domains.

A superimposition of IFI16-PAAD and MNDA-PAAD’s far UV spectra reveals a higher α-helical content of MNDA than IFI16. The α-helical content is estimated by observing the two local minima at 208 nm and 222 nm and their respective mean residue ellipticity.
3.1.1.4 Thermodynamic parameters of the PAAD domain

The thermodynamic properties of MNDA-PAAD and IFI16-PAAD were obtained by thermal denaturation and measuring circular dichroism (CD) signal at 222 nm over a temperature range from 20°C to 100°C. All curves were fitted to a two state equilibrium-folding model (Figure 3.4). MNDA-PAAD has a melting temperature of 58.7 ± 0.4 °C, \( \Delta H = -36.7 \pm 1.1 \) (kcal mol\(^{-1}\)), \( \Delta S = -0.111 \pm 0.003 \) (kcal mol\(^{-1}\) K\(^{-1}\)) and \( \Delta G_{folding}^{25°C} = -3.7 \) (kcal mol\(^{-1}\)) and IFI16-PAAD has a Tm of 42.4 ± 0.5 °C, \( \Delta H = -27.2 \pm 2.5 \) (kcal mol\(^{-1}\)), \( \Delta S = -0.086 \pm 0.08 \) (kcal mol\(^{-1}\) K\(^{-1}\)) and \( \Delta G_{folding}^{25°C} = -1.5 \) (kcal mol\(^{-1}\)). These thermodynamic parameters were extracted from the fitting of the denaturation curve to a two state folding model and the linear extrapolation method (Myers, Pace, & Scholtz, 1995; Pace, 1986). The \( \Delta G_{folding}^{25°C} \) and \( \Delta H \) for MNDA-PAAD are lower than IFI16-PAAD indicating MNDA-PAAD requires less energy at 25°C to fold properly. The higher melting temperature of MNDA-PAAD indicates it requires more energy to unfold it. At equilibrium, proteins undergo a reversible process of folding and unfolding between their native and denatured states, but different transition intermediate states may be involved (McCully, Beck, & Daggett, 2008). A protein can go from its native state into a partially denatured state (nearly native state), then into the denatured state. To monitor these states when proteins unfold and refold, a thermal denaturation along with secondary structure determination can be used. Here, the reversibility of the folding-unfolding transition of the PAAD domains was also tested using circular dichroism (Figure 3.5). The PAAD of IFI16 was able to reverse its unfolded state into its native state after thermal denaturation and the Tm obtained from the denatured and renatured curves were nearly
identical (Dalal, 2006). This suggests the folding and unfolding pathway of IFI16-PAAD follows a two state mechanism. However, this is not the case for the PAAD domain of MNDA. The renatured curve of MNDA-PAAD showed that MNDA-PAAD did not renature into its initial state, but instead lost some of its secondary structure and ellipticity compared to the initial native state observed at 20°C. These results suggest that MNDA-PAAD folding/unfolding mechanism involved more than two states, where MNDA-PAAD has been trapped in an intermediate state after heat denaturation. Although MNDA-PAAD has higher stability and secondary structure compare to IFI16-PAAD, MNDA-PAAD does not refold after being thermal-denatured. The high percentile of helical content and increased thermal stability of MNDA-PAAD over IFI16-PAAD was thought to be the difference in helix 3. Proteins of the death domain super family such as NALP1 and ASC have the same overall structure as MNDA but all differ in helix 3. Effectively, NALP1 has a disordered region between helix 2 and 4 but MNDA has a complete helix 3. Until the structure of MNDA-PAAD was solved in 2006, IFI16-PAAD was predicted to have a disordered region much like NALP1. Our data also supports the hypothesis that IFI16-PAAD has a disordered region because of it lower helical content and stability compared to MNDA-PAAD.
Figure 3.4. Thermal denaturation of the PAAD domain.

The CD signal of far UV at 222 nm is monitored over a temperature ranging from 20°C to 100°C. The ellipticity is transformed into fraction unfolded of the protein where 1 = 100% unfolded. Both PAAD domains unfolded via a two state unfolding mechanism where IFI16-PAAD is less stable than MNDA-PAAD. Their respective thermodynamic parameters are listed in Table 3-4 and Table 3-5.
Figure 3.5. MNDA-PAAD did not regain secondary structure after thermal denaturation.

MNDA-PAAD is first heated from 20°C to 100°C and then cooled back to 20°C. The ellipticity at 222 nm is monitored over the entire process. The protein is not able to regain its original amount of secondary structure after heated to 100°C.
3.1.1.5 Intrinsic fluorescence – tertiary structure unfolding of MNDA-PAAD

The intrinsic fluorescence properties could serve as an indication of the folding of the protein. When the fluorescent amino acids (tryptophan, tyrosine and phenylalanine) buried inside the hydrophobic core of the protein become exposed to solvent during denaturation, the residues experience a change in polarity in the environment surrounding them. This can cause a red or blue shift in the fluorescence spectra and changes the quantum yield of the emission intensity (Royer, 2006). Typically, tryptophan gives the most reliable and strongest measurement of the intrinsic fluorescence property of a protein, however there is no tryptophan in the protein sequence of MNDA-PAAD. Therefore, we measured the unfolding process of the protein using 6 M urea while monitoring the tyrosine fluorescence. Tyrosine absorbs light at 275 nm and emits at 304 nm and aromatic residues absorb and emit in a higher efficiency when they are buried inside the protein core due to their non-polarity. When the aromatic residues are exposed to solvent during denaturation or in close proximity with a quencher, the fluorescence emission will be quenched by the intermolecular interactions (Zhuang et al., 2000). Urea is used as the denaturant and the tyrosine emission of the protein in the absence and presence of urea is measured. Urea is very effective in disrupting non-covalent bonds in protein and we observed a slight decrease in MNDA-PAAD tyrosine emission at 304 nm when it was denatured with 6 M urea. In both the native and denatured protein fluorescence spectra, we observed a single peak at 304 nm (Figure 3.6). The peak corresponds to the denatured protein was observed to have the emission maximum lower than 304 nm and this denotes a red shift, which indicates the tyrosine experienced
a change in the polarity of the environment and is likely that the protein underwent a conformational change. Since the decrease in fluorescence is only a small scale after denaturation, it suggests that the protein retained its tertiary structure and remain well folded. Perhaps another denaturant such as guanidine hydrochloric acid would have helped to disrupt the hydrogen bonds in the protein that maintain tertiary structure. IFI16-PAAD denaturation also gave similar results (Dalal, 2006).
Figure 3.6. Tyrosine spectroscopic property of MNDA-PAAD.

The tyrosine fluorescence emission spectra of native and denatured protein. The protein is exciting at 275 nm and monitoring at 304 nm (λex = 275 nm, λem = 304 nm). The protein is denatured using 6 M urea. There is a slight decrease in fluorescence emission peak of tyrosine after the protein being denatured, indicating the tertiary structure if the protein is altered. Spectra are averaged with two independent trials (n = 2).
3.1.6 Polydispersity determination of MNDA-PAAD using dynamic light scattering

The hydrodynamic radius, polydispersity, and presence of aggregation of MNDA-PAAD were determined using dynamic light scattering. These parameters are essential in protein crystallography in helping to determine the crystallization conditions as well as providing a better understanding of how the protein behaves in various solutions. Dynamic light scattering measures the light that is scattered from the protein sample, caused by the brownian motion of the proteins in solution and the fluctuations of the scattering intensity can be observed. This scattered light intensity correlates to the apparent size of the protein in solution (Wilson, 2003). Here, 0.5 mg of MNDA-PAAD was separated by size exclusion chromatography using a Superdex 75 column. The size exclusion column was attached to a dynamic light scattering apparatus and the light scattered by the eluted proteins was detected. The size exclusion chromatogram observed a single sharp peak and the light scattered from this peak was calculated to have an apparent molecular mass of $1.29 \times 10^4$ g/mol, which correspond to the monomer of MNDA-PAAD (Figure 3.7). According to the light scattering data, greater than 90% of the light scatter came from the monomeric form of MNDA-PAAD and the remaining percentile came from a shallow peak that corresponds to the dimeric form of MNDA-PAAD when referenced back to the size exclusion chromatography. The polydispersity value was 1 and this indicates the protein was mono-dispersed in solution. Lastly, the hydrodynamic radius of MNDA-PAAD was calculated to be 1.9 nm in its monomeric form.
Figure 3.7. Dynamic light scattering of MNDA-PAAD.

A single sharp peak observed during the size exclusion chromatography in DLS. The protein determined to be monodispered with 1.9 nm of hydrodynamic radius.
3.1.1.7 Comparative modelling reveals disordered region in IFI16-PAAD

Previously circular dichroism indicates that there are differences in folding within the PAAD domain family, namely between IFI16-PAAD and MNDA-PAAD (Dalal & Pio, 2006). We illustrate this point again by superimposing their structures for visual comparison. However, the structure of IFI16-PAAD has not been solved so a comparative model of IFI16-PAAD was made. Dalal and Pio determined that using MODELLER, the best template to model IFI16-PAAD is PDB code: 1PN5, the NMR structure of another PAAD protein, NALP1 (Dalal & Pio, 2006). Using Swisspdb viewer, the superimposition of IFI16-PAAD model to MNDA-PAAD structure (PDB code: 2DBG) was made and the RMSD is calculated to be 0.3 Å between all the alpha carbon (Figure 3.8). It is clear to see that MNDA-PAAD has a complete helix 3 but IFI16-PAAD is predicted to have random coil in the corresponding region. Combining the results from circular dichroism and comparative modelling, Dalal and Pio concluded that IFI16-PAAD is most likely to have a disordered helix 3.
Figure 3.8. Superimposition of IFI16-PAAD and MNDA-PAAD showing the six helix bundle and disordered helix region.

The model of IFI16-PAAD is generated as a comparative model using NALP1-PAAD (1PN5) as the template. PDB ID of MNDA-PAAD is 2DBG. The superimposition, done using SwissPDBViewer has a RMSD = 0.3 Å. Blue cartoon represent 2DBG and the model of IFI16-PAAD is in orange cartoon. The disordered helix 3 is shown on the most right of the figure.
3.1.2 The OB fold binding properties of MNDA

3.1.2.1 Rationale

As shown by Xie (Xie et al., 1998; Xie et al., 1995), MNDA binds to transcription factor YY1 and nucleic acid only with its N-terminal half, presumably the PAAD domain region. The interaction between MNDA and nucleic acid has not been investigated in extensive details. Johnstone et al. (Johnstone et al., 2000) in 2000 showed that IFI16 binds to dsDNA and Aglipay et al. (Aglipay et al., 2003) in 2003 showed evidence and linked IFI16 to DNA repair. The PAAD domain of IFI16 was characterized by Dalal et al. (Dalal, 2006) and the author determined to be a single stranded DNA binding protein. The HIN200 domain was previously characterized by Yan et al., who concluded that it is a RPA like protein that possessed most of the OB fold DNA recognition properties (Yan et al., 2008). With the discovery that IFI16 recognizes nucleic acids similar to RPA OB fold, it would be worthwhile to know if these properties are conserved across the HIN-200 family because MNDA and IFI16 share above 50% sequence identity for their individual domain. The key features of RPA with OB fold include binding to ssDNA with nucleotide preference, oligomerization upon DNA binding, destabilization of dsDNA, compacting and extending of DNA. By investigating the nucleic acids binding properties of MNDA, we may gain a better understanding of how the HIN-200 family functions in general and determine if MNDA is a RPA-like single stranded binding protein.
3.1.2.2 MNDA-PAAD binds to nucleic acids

Electrophoretic mobility shift assay (EMSA) was performed to verify MNDA-PAAD could form protein-nucleic acid complexes. We used different single stranded oligonucleotides such as polydA25 (A25), polydT25 (T25), guanine rich 5’ 35mer (GC - 5) and cytosine rich 3’ 35mer (GC - 3) in this assay. These oligos used were $^{32}$P labelled on their 5’OH group using T4 polynucleotide kinase. The reaction mixtures were prepared by mixing a fixed amount of $^{32}$P oligonucleotides (3.6 pmol) with a series of 2 fold serial dilution of MNDA-PAAD (210 μM – 0.83 μM) (2.1 nmol – 8.3 pmol). The EMSA showed that despite any type of oligonucleotides, the addition of MNDA-PAAD resulted in a shifted band compared to the free probe (Figure 3.9). Although the shifted band only migrated 5 mm into the gel, the intensity of the shifted band decreased as the amount of protein decreased, suggesting that the shifted band comprised of MNDA-PAAD and the labelled oligonucleotides. Moreover, as IFI16-PAAD is able to bind to RNA, we further tested whether MNDA-PAAD also has the ability to bind to RNA. We used a $^{32}$P random RNA library (79mer) for another shift assay. The result indicates MNDA-PAAD could also bind to RNA regardless of the RNA sequence. Although IFI16-PAAD used a different random RNA library (49mer), similar shifted bands were seen with all probes tested in the EMSA experiment. It is possible that at high concentration of protein, some of the protein-DNA complexes may precipitated and would not penetrate the matrix. Therefore, the DNA did not show up in wells at high protein concentration. An estimation of the binding affinities to each of the oligonucleotides was obtained by analyzing the EMSA shifted bands using ImageQuant 2.2. The software integrated the bands
correspond to the bound fraction and expresses the values in percentage. The percent bound was plotted against their respective MNDA-PAAD concentrations and the curves were input into GraphPad. GraphPad fitted all curves into a one site binding model and calculated the statistics for the non-linear regression fit (Figure 3.10). The results indicate MNDA-PAAD has slight preference to polythymine and RNA nucleotide as their affinities appear greater than the other oligos tested ($K_D = 9.05 \, \mu M$ and $K_D = 1.76 \, \mu M$, respectively). Other oligonucleotides affinities are; A25 = 88.08 $\mu M$, GC - 5 = 34.66 $\mu M$, GC - 3 = 146 $\mu M$. Although no precise conclusion can be made with these numbers, more characterization of MNDA-PAAD affinity to different nucleotides were performed using fluorescence quenching (See section 3.1.2.4).
Figure 3.9. MNDA-PAAD binds to different single stranded DNA by EMSA.
Lane 1 represents 7-10 pmol of free $^{32}$P end labelled oligonucleotides. Lane 2 and beyond are 2 fold serial dilution of MNDA-PAAD starting with 210 µM to 0.83 µM mixed with a constant amount of probe (varies from 7-10 pmol depending on which oligonucleotide). F – free probe, C – Complex. (A) polydT(25), (B) polydA(25), (C) Guanine rich 5' 35mer, (D) Cytosine rich 3' 35mer, (E) random 79mer RNA library.
Figure 3.10. Binding curves and one site binding $K_0$ calculations from EMSA experiments of MNDA-PAAD.

Each graph represent one oligonucleotide used in the EMSA experiment as in Figure 3.8. The intensity of the shifted band (upper) in each lane was integrated and expressed in percent bound. (A) polydT(25) $K_0 = 9.05 \mu M$, (B) polydA(25) $K_0 = 88.08 \mu M$, (C) Guanine rich 5' 35mer $K_0 = 34.66 \mu M$, (D) Cytosine rich 3' 35mer $K_0 = 146 \mu M$, (E) random 79mer RNA library $K_0 = 1.76 \mu M$. 
3.1.2.3 The PAAD domain forms a stable complex with ssDNA as determined by UV cross linking, mass spectroscopy and western blot

The results from the EMSA experiment may not suggest that MNDA-PAAD interacts with nucleic acid. To corroborate these findings, we used UV cross linking. UV cross linking causes two transient interactions covalently linked by first irradiating the molecule to form free radical to allow for covalent bond with nearby molecule, making the interaction easier to detect. Here, MNDA-PAAD and nucleic acid are UV cross linked and their interactions are further confirmed. The same sets of oligonucleotides were used along with UV to examine the ability of MNDA-PAAD to bind to nucleic acid. In this experiment, 2 µg of protein (153 pmol) was titrated against 0, 0.1, 0.2, 0.5, 1.0, 2.0 and 5.0 molar ratio of ssDNA. The reactions were exposed to UV light at 254 nm for 15 minutes at 1000 J/min. Each reaction mixture was loaded on a 15% SDS PAGE along with a standard protein ladder and visualized using silver stain. The molecular weight of the bands that corresponded to MNDA-PAAD monomer, dimer and protein-nucleic acid complex were calculated based on the calibration curve generated by measuring the Rf of the standard protein ladder (Figure 3.12).
Figure 3.11. UV cross linking of MNDA-PAAD with ssDNA.

The reaction mixtures were ran in SDS-PAGE and silver stained for visualization. In all gels, lane 1 represents 2 μg (150 pmol) of MNDA-PAAD without cross linking, lane 2 is MNDA-PAAD cross linking without DNA, lane 3-8 represent MNDA-PAAD cross linking with ssDNA in 0.1:1, 0.2:1, 0.5:1, 1:1, 2:1, 5:1 DNA:protein molar ratios. M – Monomer, D – Dimer, T – Trimer, C – Complex, T70 – PolydT(70)
Figure 3.12. Gel calibration curve for molecular weight determination of UV cross linking species of MNDA-PAAD with ssDNA.

Each calibration curve correspond to the protein ladder ran in each gel represented in Figure 3.11. (A) T25, (B) A25, (C) GC-5, (D) GC-3, (E) T70
### Table 3-1. Experimental molecular weight determination of MNDA-PAAD-nucleic acid complexes after SDS-PAGE analysis and UV cross linking with ssDNA.

Columns are experimental molecular weight for: (M) MNDA-PAAD monomer; (D) MNDA-PAAD dimer; (O) oligonucleotide; (C) MNDA-PAAD-nucleic acid complex. (Expected) Expected MNDA-PAAD-nucleic acid molecular weight as determined by the sum of experimentally calculated M with O; (Deviation) Absolute difference between C and Expected molecular weight protein-ssDNA complexes. (NO) Not Observed

<table>
<thead>
<tr>
<th>Oligonucleotide</th>
<th>M (kDa)</th>
<th>D (kDa)</th>
<th>O</th>
<th>C (kDa)</th>
<th>Expected Complex Weight (kDa)</th>
<th>Deviation from expected weight (kDa)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A25</td>
<td>13.2</td>
<td>27.3</td>
<td>NO</td>
<td>21.6</td>
<td>20.7</td>
<td>0.9</td>
</tr>
<tr>
<td>T25</td>
<td>13.2</td>
<td>27.1</td>
<td>NO</td>
<td>22.3</td>
<td>20.5</td>
<td>1.8</td>
</tr>
<tr>
<td>GC5</td>
<td>13.1</td>
<td>26.9</td>
<td>NO</td>
<td>26.4</td>
<td>23.8</td>
<td>2.6</td>
</tr>
<tr>
<td>GC3</td>
<td>13.3</td>
<td>27.2</td>
<td>NO</td>
<td>25.0</td>
<td>23.5</td>
<td>1.5</td>
</tr>
<tr>
<td>T70</td>
<td>13.4</td>
<td>27.0</td>
<td>25.8</td>
<td>42.5</td>
<td>37.2</td>
<td>5.3</td>
</tr>
<tr>
<td>dimer + T70</td>
<td>13.4</td>
<td>27.0</td>
<td>NO</td>
<td>53.7</td>
<td>53.3</td>
<td>0.4</td>
</tr>
<tr>
<td>trimer + T70</td>
<td>13.4</td>
<td>27.0</td>
<td>NO</td>
<td>64.3</td>
<td>62.8</td>
<td>1.5</td>
</tr>
</tbody>
</table>
The calculated molecular weight of each species fell in the range of their expected molecular weight. Table 3-1 summarized the expected and experimental determined molecular weight of each species. The gels clearly showed that the MNDA-PAAD/oligo complex started to form at 0.1:1 molar ratio of DNA:protein. The cross linking of MNDA-PAAD with A25, T25, GC5 and GC3 generated only one complex band that can be seen on the SDS-PAGE (Figure 3.11). The lower band corresponds to the monomeric form of MNDA-PAAD, which has an experimental molecular weight of 13.0 kDa. The top band corresponds to the dimeric form of MNDA-PAAD as the experimental molecular weight determined to be 26.0 kDa. The band in between the monomer and dimer form of MNDA-PAAD is the protein-DNA complex where the protein monomer binds to one molecule of oligonucleotide.

There was a different staining pattern on the T70 gel that presumably corresponds to poly dT70 oligonucleotides, while other nucleotides (A25, T25, GC5, and GC3) staining can also be seen on the 15% SDS PAGE if the dye front did not run off the gel (data not shown). The complex band’s intensity increased as the concentration of ssDNA increased indicating the formation of the complex. MNDA-PAAD was UV cross linked to a 70mer of polythymine in an attempt to find out if more than 1 molecule of MNDA-PAAD could bind onto a single oligonucleotide. Interestingly, multiple bands showed up in the SDS-PAGE gel and their respective molecular weights correspond to monomer (12.9 kDa), dimer (25.8 kDa), monomer plus T70 (12.9 kDa + 24 kDa), dimer plus T70 (25.8 kDa + 24 kDa), trimer plus T70 (38.7 kDa + 24 kDa). This indicates that more than 1 PAAD domain is bound onto a single oligonucleotides of 70mer and this result is consistent with IFI16-PAAD (Dalal,
To rule out the possibility that contamination or aggregation caused the complex formation, we used mass spectroscopy to determine the exact size of the different species after UV cross linking and western blot to verify the complex contain the desired protein. In this case, we first examine the complex between IFI16-PAAD and nucleic acid, followed by MND-PAAD. Fifty μM of IFI16-PAAD was subjected to electrospray-ionization mass spectrometry. The molecular weight of IFI16-PAAD was determined to be 14.4 kDa, which is 0.1 kDa less than the expected (data not shown). We tried to perform ESI of IFI16-PAAD after UV cross linked to T25 but the complex did not ionized for unknown reason, therefore we switched to MALDI-TOF. The result obtained from MALDI-TOF has a slight discrepancy compared to ESI, which the monomer is now determined to be 15.7 kDa, roughly 1.2 kDa higher than expected. Nevertheless, a small but significant peak observed at a higher mass to charge ratio, which was expected to be the complex (Figure 3.13 B). The experimental weight of the T25/protein complex is 23.7 kDa and the expected is 22.4 kDa, again the weight is roughly 1.3 kDa off. Moreover, MND-PAAD was subjected to MALDI-TOF after UV cross linked to a different length of polythymine, T13 (Figure 3.13 A). The expected size of the MND-PAAD is 12.9 kDa and T13 is 3.9 kDa. As expected, monomeric MND-PAAD gave a molecular weight of 12.9 kDa and the complex was 16.9 kDa. The error of the complex is roughly within 1 kDa for the 2 PAAD domains after UV cross linked to oligonucleotides as determined by mass spectroscopy. These results indicate that the protein/nucleic acid complex is not caused by aggregation and we will verify the complex further by immunoblotting.
Figure 3.13. Mass Spectroscopy of the PAAD domain with ssDNA.

50 µM of MNDA-PAAD and IFI16-PAAD have been UV cross link to equal molar ratio of T13 or T25, respectively. Panel A is the mass spectrum of MNDA-PAAD and panel B is IFI16-PAAD. In both spectra, the sharp and highest peak is observed and the calculated molecular weight of that peak is (A) 12.9 kDa and (B) 15.7 kDa. The complex molecular weight of MNDA-PAAD to T13 is observed and detected to be at 16.9 kDa. The 25.8 kDa peak is the dimer of MNDA-PAAD. The complex molecular weight of IFI16-PAAD to T25 is observed at 23.7 kDa.
In the immuno-blotting, the PAAD domain of both MNDA and IFI16 were UV cross linked to A25, T25, GC5, and GC3 and ran on a 15% SDS-PAGE gel. Using anti-histidine antibody, we managed to visualize the monomer and dimer of each PAAD domain as well as the complex with A25 and T25 (Figure 3.14). However, western blot did not detect the complex with GC5 and GC3 and this could be because when the nucleic acids bound on the protein, it also blocks the antibody binding site. Nevertheless, the complex formation of the PAAD domain with nucleic acid is not a cause of aggregation or contamination as demonstrated by mass spectroscopy and western blot, respectively.
Figure 3.14. Western blot of the PAAD domain with ssDNA.

Western blot of the PAAD domain using anti-histidine tag antibody. Proteins were subjected to UV cross linking as previously described (Figure 3.10) at 1:1 molar ratio. In the absence of oligonucleotides, no complex is detected in both SDS-PAGE and western blot in the absence and presence of UV. Complex is detected only when both oligonucleotides and UV cross linking applied. C – Complex.
To further characterize the ability of MNDA-PAAD to interact with nucleic acid, we used different length of polythymine ranging from 19 mer to 5 mer and apply UV cross-linking as described above to determine the minimal length of single stranded nucleic acid that MNDA-PAAD recognizes. Proteins that interact with nucleic acid have a tendency to bind to certain length of the DNA or specific nucleotide sequence to function optimally, which is considered as the interaction specificity (Mazina & Mazin, 2004). Similar types of bands patterning were obtained for T19 to T5 compared to T25 except with the band that corresponds to the complex, which would migrate differently depending on the size of the oligonucleotides. We also observed the complex band intensity got weaker and eventually disappeared as the oligonucleotides length got shorter indicating no complex formed (Figure 3.16). The weakest intensity of the complex could be seen with 11mer of ssDNA, which suggests that one molecule of MNDA-PAAD could bind to one molecule of ssDNA with at least 11 mer in length. Although it is hard to determine if the protein-DNA complex bands exist in the T7 and T5 gel, we assume MNDA-PAAD is binding to 9 nucleotides because on the electrostatic potential surface of MNDA-PAAD, we found a basic patch on the surface that spans approximately 30 Å (Figure 3.15). The length of 9 nucleotides is 30.6 Å and therefore the ssDNA could be binding on that particular basic patch on MNDA-PAAD.
Figure 3.15. Electrostatic potential of MNDA-PAAD.

The electrostatic potential of MNDA-PAAD generated by Delphi and visualized in Pymol. Blue color represents basic surface and red color represents acidic surface. The arrow is pointing towards the basic surface where the 9 nucleotides oligo is potentially binding.
Figure 3.16. Length dependence of MNDA-PAAD with ssDNA by UV cross linking.

Each gel represents a different length of polydT and UV cross linking is applied in similar fashion as Figure 3.10. The mixtures were loaded on SDS-PAGE and silver stained for visualization. No complex can be seen in T7 and T5 even the ssDNA is at 5 molar excess.
3.1.2.4 MNDA-PAAD binding constant determination using fluorescence quenching assay

Previously we showed that the tyrosine fluorescence in MNDA-PAAD could be quenched upon denaturation. In general, the binding of a ligand can also quench intrinsic fluorescence, therefore the binding affinity can be determined by curve fitting of the saturation curve to obtain the binding mode. The saturation curve was plotted by the concentration of the ligand against the amount quenched. According to Yan and Dalal, the PAAD and HIN200 domain of IFI16 have at least micro molar range affinity for nucleic acid as shown by tyrosine fluorescence quenching (Dalal, 2006; Yan et al., 2008).

Here we performed the same approach by quenching the tyrosine fluorescence with ssDNA by titration to obtain a more accurate binding constant for MNDA-PAAD-nucleic acid complexes. The concentration of MNDA-PAAD was kept constant at 0.1 mg/ml (7.8 mM) and it was titrated against an increasing concentration of oligonucleotides from 0:1 to 4:1 molar ratio with 0.1 molar increments. The samples were excited at 275 nm and the emission spectrum from 250 nm to 350 nm was recorded. The emission peaks at 304 nm were transformed into relative percentage quenched and plotted against their respective molar ratio of nucleic acid added. The resulting curves were analyzed by GraphPad to curve fit it into a one site or two site binding model and determine the binding constant (Figure 3.17). All the curves were preferentially fitted to a one site binding model with non-linear regression. MNDA-PAAD binds to each of the oligonucleotides with similar $K_D$ with micro molar range affinity, which is considered low affinity and could be non-specific binding. There is a $B_{max}$ value indicating the percentage of substrate bound with each $K_D$ obtained. However, MNDA-PAAD only fits to
one site binding model regardless of the different nucleotides used, therefore the \( B_{\text{max}} \) value for each \( K_D \) with each nucleotide is almost 1. Each of the oligo nucleotides used in the experiment are designed to be thymine rich, adenine rich, cytosine rich, guanine rich, single stranded or duplex, yet their affinity to MNDA-PAAD are similar. This denotes MNDA-PAAD has no preference for nucleotides base and the ratio of single stranded to double stranded is approximately 1 indicating MNDA-PAAD is a double stranded binding protein, which some of the results contrasted to the results found with EMSA. It is important to note that the two experimental techniques have an essential different pitfall in determining the binding affinity, which will be discussed in subsequent section (Section 4.2). The saturation binding curve of the tyrosine fluorescence showed a plateau region when it reached 100% (Figure 3.17), meaning all of the tyrosines’ fluorescence in MNDA-PAAD have been quenched. This quenching suggests that the tyrosine residues of MNDA-PAAD were involved in the DNA interaction. Berg, Winter and Von Hippel described a DNA binding model where protein binds to DNA with low affinity and non-specifically first by permitting the protein to sample many binding sites per DNA encountered (Berg, Winter, & von Hippel, 1981; Winter, Berg, & von Hippel, 1981), followed by intramolecular translocation to the specific binding site (Halford & Marko, 2004). Kalodimos also suggested that protein sequentially binds to DNA with low affinity and then a conformational change to the protein promotes high affinity with specificity (Kalodimos et al., 2004). The conformational change of MNDA-PAAD upon binding to nucleic acid is further characterized by NMR and is discussed in subsequent section. However, IFI16-PAAD was shown to prefer GC rich sequence and has two
binding modes. The calculated $K_D$ and $B_{\text{max}}$ values for MNDA-PAAD and each nucleotide is listed in Table 3-2.
<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>$K_0$ (µM)</th>
<th>$B_{\text{max}}$ (% x 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A25</td>
<td>1.725 ± 0.053</td>
<td>1.117 ± 0.030</td>
</tr>
<tr>
<td>T25</td>
<td>1.453 ± 0.091</td>
<td>1.103 ± 0.027</td>
</tr>
<tr>
<td>GC-3 (Mut3)</td>
<td>1.470 ± 0.506</td>
<td>1.114 ± 0.036</td>
</tr>
<tr>
<td>GC-5 (Mut5)</td>
<td>1.269 ± 0.301</td>
<td>1.108 ± 0.022</td>
</tr>
<tr>
<td>dsAT</td>
<td>1.804 ± 0.132</td>
<td>1.017 ± 0.025</td>
</tr>
<tr>
<td>dsGC</td>
<td>1.783 ± 0.120</td>
<td>0.997 ± 0.023</td>
</tr>
</tbody>
</table>

Table 3-2. Nucleic acid binding parameters of MNDA-PAAD by tyrosine fluorescence quenching.

All tyrosine fluorescence quenching data were collected with three independent trials (n = 3). The $K_0$ and $B_{\text{max}}$ are calculated by averaging the three trials and the standard deviations are included.
Figure 3.17. Dissociation constant of MNDA-PAAD with nucleic acids determined by tyrosine fluorescence quenching.

MNDA-PAAD tyrosines were titrated by (A) A25, (B) T25, (C) GC-5, (D) GC-3, (E) dsGC, (F) dsAT from 0 to 15 µM of DNA. The emission peak at 304 nm of tyrosine was plotted against the concentration of DNA in each titration and their affinity was analyzed using GraphPad. The insert is the scatchard plot representing the preferential binding mode for each set of oligo.
3.1.2.5 **MNDA-PAAD does not compact or extend ssDNA by Fluorescence Resonance Energy Transfer (FRET)**

The ability to compact or extend DNA is a common property exhibits by single stranded nucleic acids binding protein (Robbins *et al.*, 2005). We recently showed that the PAAD domain of IFI16 was able to compact ssDNA suggesting that IFI16-PAAD is a ssDNA binding protein (Dalal, 2006). Although from fluorescence quenching experiment that MNDA-PAAD does not show any preference to single stranded nucleic acid, the ability to modify ssDNA conformation may be retained in the family. To investigate this property, we used the fluorescence resonance energy transfer (FRET) approach and discovered that MNDA-PAAD does not compact or extend ssDNA. In this approach, a 58 bp oligonucleotides that consist of an 18 bp double strand and 40mer polythymine 3' overhang was used. The 5' end of the 18 bp oligonucleotides was labelled with QUASAR 670 fluorescence tag and the 3' polythymine overhang of the 58 bp oligonucleotides was labelled with QUASAR 570 fluorescence tag. A schematic representation of the FRET oligo is shown in Figure 3.18. The corresponding excitation and emission wavelength of QUASAR 670 and QUASAR 570 are 650 nm 670 nm and 550 nm 570 nm, respectively. The QUASAR 570 was used as the donor and QUASAR 670 was used as the acceptor. The energy transfer was measured between the donor and acceptor by exciting the donor at 548 nm and recording the emission spectra at 667 nm. A fixed amount of FRET oligo (300 nM) was titrated with MNDA-PAAD in 300 nM increments to a final molar ratio of 20:1 PAAD:oligo. The polythymine overhang is flexible and could wrap around the protein or extend depending on the ability of the protein to deform.
it. RPA protein has the property to compact or extend the ssDNA to which it binds; therefore, we speculated that MNDA-PAAD could act as a ssDNA binding protein and could also compact or extend ssDNA like RPA. If MNDA-PAAD wraps the polythymine overhang, the energy transferred from the donor to the acceptor should increase. On the other hand, if it is being stretched, the energy transfer should decrease with the addition of protein because the distance between the donor and acceptor increased. The result showed the donor fluorophore energy was quenched in a dose dependent manner with increasing protein concentration and the energy from the excited donor was lost with no specific reason (Figure 3.19). It could be that the light energy has changed into other forms of energy such as heat, which was not measured. We also tested the transfer from the donor to the acceptor in the absence of protein suggesting that the polythymine distance is close enough to have energy transferred.
Figure 3.18. Oligonucleotides design to study compaction or extension of ssDNA by MNDA using FRET.

18 nucleotides region of the FRET oligo is double stranded and 40 nucleotides are single stranded. The scenario on the left represents compacting of the ssDNA (wrapping) and therefore the acceptor and donor come close together and energy can be transferred. The scenario on the right represents extending (stretching) of the ssDNA where the two fluorophores distance increased, hence decreased in energy transfer.
Figure 3.19. Fluorescence resonance energy transfer suggests that MNDA-PAAD does not compact or extend single stranded DNA.

(A) FRET measurements were performed at a constant concentration of 300 nM double-stranded oligonucleotide in response to increasing MNDA-PAAD concentration. Total Fluorescence scan at different wavelength and PAAD:dsDNA ratio.
(B) The maximum emission of the donor Quasar 570 (at 564 nm) and acceptor Quasar 670 (at 667 nm) is plotted against the molar ratio of MNDA-PAAD reveals no energy was transferred during the titration.
3.1.2.6 PAAD domains destabilize dsDNA by hyperchromicity assay

The HIN200 domain of IFI16 was previously shown to be a RPA like protein but lacks the ability to destabilize the duplex DNA (Yan et al., 2008), however this property is found in the PAAD domain of IFI16 (Dalal, 2006) and the two domains could function by complementing each other. It is not known if duplex DNA destabilization is general to the PAAD domain family, therefore we tested with MNDA-PAAD. This was done by measuring the melting point of a duplex in the absence and presence of MNDA-PAAD. The duplex DNA used in this experiment was 35mers and prepared by annealing 100 μM GC-5 and 100 μM GC-3 in an annealing buffer. The spectrophotometer records the hyperchromicity of 66 μM dsDNA (dsGC) at UV 260 nm over a temperature range from 20°C to 100°C. The melting point of the duplex decreased as more MNDA-PAAD was added and at 5 molar excess of MNDA-PAAD, the melting point showed a 20 degree depression (Figure 3.20). This clearly illustrated that MNDA-PAAD has the ability to destabilize double stranded DNA and we speculated that the more folded and stable PAAD domain can destabilize duplex DNA better. It is because IFI16-PAAD does not destabilize duplex DNA as potent as MNDA-PAAD. This hypothesis will be examined further in section 3.3.
Double strand GC was prepared first by heating to 95°C for 10mins and cool down to 20°C at approximately 0.5°C/min. Increasing MNDA-PAAD was used in each round of melting depression and the Tm of the duplex was 61.5°C without protein. At 5 molar excess of protein, the Tm of the duplex drops to 41.9°C.
3.1.2.7 Oligomerization of the PAAD domain determine ssDNA binding property

The PAAD domain family is found to be able to form a homodimer for activation, and oligomerization is a common phenomenon when protein interacts with nucleic acid. The PAAD and HIN200 domains of IFI16 were shown to form oligomers upon binding to ssDNA and this property suggesting IFI16 is a ssDNA binding protein like RPA. However, previous results from MNDA-PAAD showed very distinct nucleic acid binding properties compared to IFI16-PAAD, making it ambiguous to claim that MNDA-PAAD is a ssDNA binding protein. Another characteristic of a single stranded DNA binding protein is the ability to oligomerize into protein-protein complexes (Robbins et al., 2005; Robbins et al., 2004). Previously we observed that a small portion of MNDA-PAAD formed a dimer after UV-cross linking, indicating the ability to form dimer is retained. Therefore, we tested if MNDA-PAAD could oligomerize upon binding to nucleic acid in the presence of chemical cross-linker when bound with T25 and T70. We used formaldehyde as the cross linker since formaldehyde cross links between the nitrogen atom at the end of the side-chain of lysine and the nitrogen atom of a peptide linkage. 2 μg of MNDA-PAAD was incubated with 1% v/v formaldehyde in the absence and presence of T25 and T70. The samples were loaded on a 15% SDS PAGE gel and visualized by silver staining. When MNDA-PAAD is chemically cross linked in the absence of ssDNA, the dimeric and trimeric form of the protein showed an increased intensity indicating the protein has been cross linked into multimeric form. The oligomerization state of MNDA-PAAD in the presence of ssDNA remains the same suggests that unlike IFI16-PAAD, MNDA-PAAD does not
form oligomers upon binding to ssDNA regardless of the length of the ssDNA (Figure 3.21).
Figure 3.21. Chemical cross linking of MNDA-PAAD to polythymine 70.

Protein without cross linker in lane 1. Protein with cross linker without ssDNA in lane 2. Lane 3 has both the cross linker and ssDNA but MNDA-PAAD does not increase its oligomerization state. M – MNDA-PAAD monomer, D – MNDA-PAAD dimer, Tr – MNDA-PAAD trimer.
3.1.2.8 MNDA-HIN200 also has RPA like function

The HIN200 domains of IFI16 was identified to be an RPA like protein that recognizes ssDNA better, oligomerizes upon binding, compacts and extends ssDNA and has the same binding polarity as RPA (Yan et al., 2008). We sought to characterize the ssDNA binding property of MNDA-HIN200 by first asking the question, does MNDA-HIN200 behave like RPA? MNDA-HIN200 was expressed and purified in a similar fashion as MNDA-PAAD except only one round of affinity chromatography was done to purify MNDA-HIN200. The purity of MNDA-HIN200 can be seen in Figure 3.22 and by naked eye the purity of the protein is greater than 90% when visualized on SDS-PAGE with coomassie blue stained. We verified that MNDA-HIN200 is able to bind to ssDNA as illustrated by EMSA and UV cross linking. In the EMSA experiment, $^{32}$P γ-ATP is used to label these oligonucleotides; polydT(25), polydA(25), guanine rich 5’ 35mer, cytosine rich 3’ 35 mer. The titration used a fixed amount of radio labelled probe (~10 μM/10 pmol) and titrated against a 2 fold serial dilutions of MNDA-HIN200 from 210 μM to 0.83 μM (21 nmol to 8.3 pmol). A shifted band is observed in the first few titration points, where there were enough protein to form complex with the probe and cause a change in migration in the gel. This shifting pattern is similar to the EMSA of MNDA-PAAD (Figure 3.9, Figure 3.23), which a common problem arose was the complex did not penetrate into the gel matrix. Therefore, we used UV cross linking to assess the complex formation of MNDA-HIN200 to ssDNA. As expected, a band corresponding to the complex molecular weight was observed in the UV cross linking gel. Besides the formation of complex, MNDA-HIN200 also formed dimer without UV cross linking but the portion of
dimer increased when MNDA-HIN200 was UV cross linked. The complex band was observed between the monomer and dimer of MNDA-HIN200, where the intensity of the complex band increased as the amount of ssDNA increased (Figure 3.24). Moreover, we showed that MNDA-HIN200 does not oligomerize upon binding to ssDNA using chemical cross linking (Figure 3.26). MNDA-HIN200 is chemically cross linked using formaldehyde in the presence or absence of ssDNA. Again, we used two different length of oligonucleotides to determine firstly if the presence of ssDNA promotes protein-protein interactions and secondly if the protein-protein interactions are oligonucleotides length dependent? Surprisingly that MNDA-HIN200 does not oligomerize upon the addition of ssDNA, the same result with MNDA-PAAD where no protein-protein interaction was promoted by ssDNA. To go on further characterizing the DNA binding properties of MNDA-HIN200, we used the same approaches such as fluorescence quenching and FRET. In fluorescence quenching, a minor changed was made which we used tryptophan fluorescence instead of tyrosine because MNDA-HIN200 has one tryptophan. As mentioned earlier, tryptophan fluorescence is by far the most sensitive to chemical environment with a higher quantum yield compare to tyrosine and phenylalanine. The tryptophan fluorescence quenching experiment with MNDA-HIN200 was conducted using the same six sets of oligonucleotides used in the previous studiesA25, T25, GC-5, GC-3, dsGC and dsAT. The spectra were obtained by exciting the protein with 280 nm and measured the emission from 300 nm to 400 nm, where 348 nm is the maximum emission of tryptophan. A gradual decrease of fluorescence intensity at 348 nm was observed while DNA is being added into the protein
solution, indicating that the microenvironment is altered near the tryptophan. Each titration spectra was fitted into one site or two site binding model as before using GraphPad Prism. The curve fittings are showed in Figure 3.25 and the binding constants are listed in Table 3-3. All curves were best fitted into one site binding model and their respective binding affinity falls in the micro molar range, suggesting weak affinity and possibly non-specific binding. The highest affinity that MNDA-HIN200 recognizes determined by tryptophan fluorescence quenching is ssDNA T25 with 57.6 nM. The data also indicates that MNDA-HIN200 does not prefer single stranded to double stranded. Lastly, FRET was used to determine if MNDA-HIN200 also has the ability to extend or compact ssDNA like IFI16-HIN200 domain. Using the same FRET oligo as illustrated in Figure 3.18 and the same experimental procedures as listed in section 2.11, we found that MNDA-HIN200 has the ability to first compact the ssDNA and then extend it. When MNDA-HIN200 was first introduced into the FRET oligo solution, the acceptor fluorescence emission at 667 nm started to increase. Each subsequent addition of MNDA-HIN200 increased the acceptor emission until the protein was at 11 molar excess to the FRET oligo, after which the acceptor emission started to decrease. It is unfortunate that we cannot plot the emission of the donor because some spectra near 580 nm were out of the dynamic range measurement of the instrument. The increase of acceptor fluorescence indicates the distance between the donor and acceptor is shortened and therefore the energy from the donor is more efficient to transfer to the acceptor, where a decrease in acceptor fluorescence indicates the exact opposite.
Summarizing the DNA binding results from MNDA-HIN200, this domain recognizes nucleic acid as demonstrated by UV cross linking and EMSA. It binds both ssDNA and dsDNA with a slight prefer for single stranded polythymine ($K_D = 57.6 \text{nM}$). It does not form protein-protein interactions upon binding to ssDNA but it does modulate ssDNA conformation by extension or compaction. A detailed comparison of the DNA binding properties between IFI16 and MNDA is listed in Table 3-7.
Figure 3.22. SDS-PAGE analysis of MNDA-HIN200.

The protein is estimated to have a molecular weight of 26 kDa and a prominent band is observed at the correct molecular weight. Approximately 60 µg of protein was loaded on the gel.
Figure 3.23. MNDA-HIN200 recognizes ssDNA by EMSA.

Each panel represents a set of oligonucleotides used in the EMSA experiment. In all gels, the left most lane contains free probe alone. Protein concentration decreases by half in each lane from left to right starting from 210 mM. The shifted band on the top of the gel corresponds to the protein-DNA complex. F – Free probe, C – Complex. (A) polydT25, (B), polydA(25), (C) Guanine rich 5’ 35mer, (D) Cytosine rich 3’ 35 mer.
Figure 3.24. UV cross linking of MNDA-HIN200 with ssDNA.

In all gels, lane 1 represents 2 μg (77 pmol) of MNDA-HIN200 without cross linking, lane 2 is MNDA-HIN200 cross linking without DNA, lane 3-8 represent MNDA-HIN200 cross linking with ssDNA in 0.1:1, 0.2:1, 0.5:1, 1:1, 2:1, 5:1 DNA:protein molar ratios. (A) polydT(25), (B) polydA(25), (C) Guanine rich 5’ 35mer, (D) Cytosine rich 3’ 35mer. M – Monomer, D – Dimer, T – Trimer, C – Complex.
C

% Quenching

[GC-3] (μM)

D

% Quenching

[GC-5] (μM)
Figure 3.25. Dissociation constant determination of MNDA-HIN200 by Tryptophan fluorescence quenching.

MNDA-HIN200 tryptophan was titrated by (A) A25, (B) T25, (C) GC-3, (D) GC-5, (E) dsGC, (F) dsAT from 0 to 15 μM of DNA. The emission peak at 348 nm of tryptophan was plotted against the concentration of DNA in each titration and their affinity was analyzed using GraphPad. The insert is the scatchard plot represents the preferential binding mode for each set of oligo.
<table>
<thead>
<tr>
<th>Oligonucleotides</th>
<th>$K_D$ (μM)</th>
<th>$B_{\text{max}}$ (% x 100)</th>
</tr>
</thead>
<tbody>
<tr>
<td>A25</td>
<td>$2.758 \pm 0.278$</td>
<td>$0.862 \pm 0.034$</td>
</tr>
<tr>
<td>T25</td>
<td>$0.576 \pm 0.060$</td>
<td>$1.070 \pm 0.024$</td>
</tr>
<tr>
<td>GC-3 (Mut3)</td>
<td>$1.997 \pm 0.477$</td>
<td>$1.129 \pm 0.095$</td>
</tr>
<tr>
<td>GC-5 (Mut5)</td>
<td>$1.920 \pm 0.390$</td>
<td>$1.026 \pm 0.073$</td>
</tr>
<tr>
<td>dsAT</td>
<td>$3.722 \pm 0.467$</td>
<td>$0.948 \pm 0.052$</td>
</tr>
<tr>
<td>dsGC</td>
<td>$1.700 \pm 0.129$</td>
<td>$0.913 \pm 0.025$</td>
</tr>
</tbody>
</table>

Table 3-3. **Nucleic acid binding parameters of MNDA-HIN200 by tryptophan fluorescence quenching.**

All tryptophan fluorescence quenching data were collected with one trial and the standard deviations were calculated by fitting the saturation curve to the binding models in Graphpad.
Figure 3.26. Chemical cross linking of MNDA-HIN200 to T70.

Protein without cross linker in lane 1. Protein with cross linker but without ssDNA in lane 2. Lane 3 has both the cross linker and ssDNA but MNDA-HIN200 does not increase its oligomerization state in the presence of ssDNA. M – MNDA-HIN200 monomer, D – MNDA-HIN200 dimer, Tr – MNDA-HIN200 trimer.
Figure 3.27. MNDA-HIN200 compact and extend ssDNA by FRET.

ssDNA is compacted when titrated by protein from 1-10 molar ratio as indicated by an increase in fluorescence at 667 nm, after which a decline in fluorescence suggests ssDNA is being extended.
3.2 Chemical shift mapping of MNDA-PAAD with ssDNA

In an attempt to identify which amino acids of MNDA-PAAD are responsible for the DNA interaction, we performed a chemical shift mapping experiment by NMR. A major challenge in NMR is to obtain a pure sample with relative high concentration (0.3 mM or above) and high stability. In order to reduce the chance of aggregation, we mutated the free cysteine residue in MNDA-PAAD to a serine residue. By doing such a mutation, we hoped to eliminate the potential of non-specific disulfide bridge formation. The secondary structure and thermal stability of the cysteine free MNDA-PAAD were tested by circular dichroism and no major differences were observed compared to wild type. We uniformly labelled the nitrogen atoms in the protein with the $^{15}\text{N}$ isotope in M9 minimal media. The labelled protein was purified using his-tag purification and gel filtration as described previously. We set the reference point of the 2D HSQC on the Bruker 600MHz using $^{15}\text{N}$ labelled T4 lysozyme. In the actual 2D HSQC experiment, 1.7 mg/ml (0.13 mM) of MNDA-PAAD in 5 mM Tris, 100 mM NaCl, pH 8 10% D$_2$O was subjected to $^1\text{H}/^{15}\text{N}$ 2D HSQC using water flip-back at 25°C. Of the 110 amino acids, the cross-peaks of approximately 78 backbone amides in the 2D spectrum were observed. Most of the peaks were evenly distributed, indicating that the protein is folded properly. To determine if any residue in MNDA-PAAD interacts with or is in close proximity to the nucleic acid when bound, a chemical shift mapping of the protein with single stranded DNA was performed and monitored by a series of HSQC spectra. A different instrument was used in the titration because for technical issue. Nevertheless, the Varian 500 MHz gave similar 2D HSQC spectra with MNDA-PAAD alone (0.23 mM).
We titrated the protein using 0.1 molar ratio of ssDNA (0.023 mM) to protein (0.23 mM) as each increment up to 0.4 molar ratio. We selected a shorter ssDNA, T13, to be used in the NMR titration experiment to minimize the noise that could be generated during the NMR measurement. It is clear to see that most of the peaks collapsed towards 8 ppm on the $^1$H dimension at 0.4 molar ratios indicating the protein being denatured/aggregated. As a result the experiment was stopped at 0.4 molar ratio. A superimposition of each of the titration 2D HSQC spectra is shown in Figure 3.28. The overall protein conformation was changed after the second titration and we believed that the protein at that point started to denature. Nonetheless, some peaks remain unchanged after titration and we highlighted 3 peaks that were shifted upon addition of ssDNA. Their orientations are ($^1$H 6.84 ppm, $^{15}$N 123.2 ppm), ($^1$H 7.36 ppm, $^{15}$N 116 ppm), ($^1$H 7.51 ppm, $^{15}$N 117.2 ppm). The fact that some N-H peaks were shifted indicates the nucleic acid added is in close proximity and resulted with a change of the chemical environment of that N-H group. Each amino acid has an amide group on the polypeptide backbone and each N-H peak on the spectra should indicate an amino acid. It should also note that the amide group on lysine and arginine will also show up as a peak on the spectra. The change in chemical shift demonstrated that the binding of nucleic acid is specific to certain amino acids. I was able to map all the peaks chemical shift change on the 2D HSQC titration spectra and traced down their movement by the titration analysis tools in NMRViewJ. All the peaks on the 2D HSQC spectra were assigned with a number and the relative chemical shift change of each peak is represented in a bar graph (Figure 3.30). There are certainly some peaks having greater chemical shift change and for those
peaks having a chemical shift changed, the binding affinity was calculated. The binding affinities were pooled and surprisingly that there are 3 distinct pools, a pool of low binding affinity (M), moderate affinity (μM) and high affinity (fM). This result suggests that some of the N-H group on MNDA-PAAD are binding to the ssDNA tightly and maybe critical to the binding. However, some chemical shift experienced a dramatic change and it is possible that the protein has changed it conformation before being denatured. There could be multiple reasons that contributed to the denaturation of the protein during the NMR experiment, which one of them could be the addition of ssDNA leads to destabilization of the protein. To test if the addition of ssDNA caused the denaturation of the protein, a thermal stability test was performed. Similar to thermal denaturation as previously described, MNDA-PAAD was heated from 20°C to 100°C with equal molar concentration of ssDNA. The melting processed was monitored by circular dichroism at 222 nm (Figure 3.29). In the presence of ssDNA, the melting point of the protein is 55.9°C and $\Delta G_{folding}^{25°C}$ is -2.47 kcal mol$^{-1}$. The $\Delta G_{folding}^{25°C}$ of the protein in the presence of ssDNA is slightly higher compared to the apo-protein indicating the presence of ssDNA lower the protein stability, which could affects the time of the protein to remain soluble in solution. Table 3-4 summarizes the thermodynamic parameter of MNDA-PAAD in the presence and absence of ssDNA.
Figure 3.28. MNDA-PAAD recognizes ssDNA in the fast/medium exchange rate by chemical shift mapping.

Top left: $^1$H/$^{15}$N 2D HSQC titration spectra of 0.23 mM MNDA-PAAD in 5 mM Tris 100 mM NaCl pH 8.0 10% D$_2$O and 25°C. Titration spectra is color coded, where red = no ssDNA added, blue = 0.1 molar ratio of ssDNA added, green = 0.2 molar ratio, purple = 0.3 molar ratio, orange = 0.4 molar ratio. The significantly affected peaks are highlighted and zoomed in on the top right ($^1$H 6.84 ppm, $^{15}$N 123.2 ppm), bottom left ($^1$H 7.36 ppm, $^{15}$N 116 ppm) and bottom right ($^1$H 7.51 ppm, $^{15}$N 117.2 ppm), where their movement are obvious in the first two titration and the movement stopped at the last titration point.
Thermal Denaturation of MNDA-PAAD with ssDNA

Figure 3.29. ssDNA destabilizes MNDA-PAAD structure by thermal denaturation.

The addition of ssDNA to MNDA-PAAD affects the thermostability of the protein slightly. The protein remains a two state unfold mechanism and the thermodynamic parameters are listed in Table 3-4.

<table>
<thead>
<tr>
<th></th>
<th>$T_m$ (°C)</th>
<th>$\Delta G_{25°C}^{\text{folding (kcal/mol)}}$</th>
<th>$\Delta H$ (kcal/mol)</th>
<th>$\Delta S$ (kcal/mol/k)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MNDA-PAAD mutant</td>
<td>53.9 ± 0.4</td>
<td>-3.2</td>
<td>-36.3 ± 0.7</td>
<td>-0.111 ± 0.002</td>
</tr>
<tr>
<td>MNDA-PAAD + T25</td>
<td>55.9 ± 0.5</td>
<td>-2.5</td>
<td>-26.3 ± 0.3</td>
<td>-0.08 ± 0.009</td>
</tr>
</tbody>
</table>

Table 3-4. Thermodynamic parameters of MNDA-PAAD in the absence and presence of ssDNA.

Thermal denaturation data were collected with one trial and the standard deviations were calculated by fitting the denaturation spectra to a two state denaturation model.
Figure 3.30. Relative chemical shift change in the chemical shift mapping of MNDA-PAAD.

Each peak is assigned a number and the relative chemical shift changes are indicated as a bar graph. The KD of each peak that have a change in chemical shift is calculated and the distribution is indicated in the insert. The KD of each peak is split into 3 pools having low binding affinity (M), moderate affinity (μM) and high affinity (fM).
3.3 Correlation of IFI16-PAAD's structure and stability to dsDNA destabilization property

3.3.1 Rationale

Dalal and Pio recently showed that death domains and PAAD domain family members have different thermodynamic and secondary structure signatures that may confer new protein-ligand interactions (Dalal & Pio, 2006). Here we verified that hypothesis by demonstrating that the PAAD domain of IFI16 can bind to ssDNA and is able to destabilize dsDNA. Finally, we also illustrated that such destabilization properties are conserved at least between two different family members MND1 and IFI6. Here we went further by generating mutants with different thermodynamic stability and secondary structure contents of IFI6-PAAD and tested the effect of protein stability and amount of secondary structure to the dsDNA destabilization properties.

3.3.2 Identification of IFI16-PAAD mutations with higher stability and secondary structure content by GFP directed evolution

Previously Dalal provided evidence that IFI16-PAAD has a disordered region, but the reasons causing this disordered region are currently unknown. Dalal investigated the structural basis of IFI16-PAAD and determined that the PAAD domain folds by a two state mechanism and can undergo different partially folded conformations when induced by secondary structure promoting agents such as TFE and SDS (Dalal & Pio, 2006). Specific mutations of the PAAD domain may occur during natural evolution that account for the disordered region. Here we adopted the system developed by Waldo in 2003, in which the author fused the target protein C-terminus to a GFP protein. The
author concluded that the more soluble of the target protein could make the GFP fluoresces better (Waldo, 2003). We attempted to use this method to correlate protein solubility with protein folding and stability (Cabantous et al., 2005; Waldo, 2003; Waldo, Standish, Berendzen, & Terwilliger, 1999). In this study, random PCR mutagenesis and DNA shuffling were performed on IFI16-PAAD to mimic natural protein evolution in vitro in order to identify mutants with increased solubility. Mutants are then fused with a GFP (Green Fluorescence Protein) tag at their C-terminus for increased fluorescence selection (Cabantous et al., 2005; Waldo, 2003; Waldo et al., 1999). For those mutants with increased fluorescence will be tested for its secondary structure and stability by circular dichroism. The secondary structure contents of each mutant were assessed by far UV and cdPRO analysis, and the stability of the mutants was obtained by recording their thermal and chemical denaturation. The denaturation curves were fitted to a two state unfolding model by linear extrapolation. My work on IFI16-PAAD directed evolution has led to the conclusion that mutations (146N, L28F, K88E) generated the most improvement in helical content (% helix = 44.7 ± 4.6) and stability (ΔG_{folding}^{25^\circ C} = -2.44 ± 0.51 kcal/mol). On the other hand, mutations (146N, E15K, L28P) produced the least stable (ΔG_{folding}^{25^\circ C} = -0.77 ± 0.07 kcal/mol) and decreased helicity protein (% helix = 22.9 ± 2.0). These mutations can be seen on the model of IFI16-PAAD illustrated in Figure 3.31. All other mutations’ thermodynamic parameters and percent of secondary structure are listed in Table 3-5.
Figure 3.31. Modelling of IFI16-PAAD mutant with increased secondary structure and stability.

Four mutations generated by random mutagenesis and DNA shuffling tested to have the most significant affect on the structure and stability of IFI16-PAAD are E15 (red), L28 (purple), I46 (brown), K88 (green).
Four mutants thought to influence the structure and stability the most were further tested using chemical denaturation and the CD spectra was monitored at 222 nm. The Gibbs free energy at 25°C in the absence of denaturants (Δ\(C_{H2O}^{25\text{C}}\Delta G\)) of each mutant was calculated using VanHoff’s analysis and the critical denaturant concentration and change in surface area (\(C_m\) and ΔASA) value were obtained using the correlation proposed by Myers (Myers et al., 1995). The most interesting observation is that the mutation that impart increased stability and structure do not lie in helix 3 region, but the neighbouring amino acids, namely L28 and I46 could be responsible for the helix propagation by a structural shift. The thermal and chemical stability parameter of the four mutants are listed in Table 3-6.

3.3.3 IFI16-PAAD mutants with increased stability and secondary structure confer enhanced dsDNA destabilization

It has always been a debatable question of whether protein structure confers function and how the structure of a protein affects its function. Here we showed that by perturbing the protein structure and selecting by GFP fluorescence intensity, we can identify mutants with different thermodynamic signatures. We wanted to extend this structural perturbation to functional studies by looking at the effect of different thermodynamic signatures to dsDNA destabilization. To address this, the four mutants with the most significant impact on IFI16-PAAD secondary structure and stability (mutant 4, 6, 22, and 29) were subjected to double stranded DNA destabilization assay. As described in section 3.1.2.6, IFI16-PAAD destabilizes double stranded DNA (Dalal, 2006) but not as dramatically as MNDA-PAAD does. MNDA-PAAD caused a dramatic 20°C decrease in the melting temperature of the
dsDNA while IFI16-PAAD only decreased the melting temperature of the dsDNA by 10°C. Consistent to our hypothesis that better folded PAAD domain can destabilize duplex DNA better, we found mutant 4 and 6 were more structured and stable and mutant 22 and 29 were the less structured and stable than the wildtype. When these mutants protein were incubated with dsDNA during melting experiment we observed two distinct pools, mutant 4 and 6 shifted the $T_m$ more to the left indicating a lower $T_m$ and mutant 22, 29 and wild type shifted not as far indicating higher $T_m$ (Figure 3.32). These results suggest that the method of GFP directed evolution selects about 50% of improve folding and stability mutants and the better folded mutants have improved the ability to destabilize dsDNA.
Melting depression of IFI16-PAAD mutants

dsDNA melting curves from IFI16-PAAD and its mutants. Mutant 22 and 29 have a similar destabilization activity like the wild type. Mutant 4 and 6 are better folded and more stable protein, here they also correlate to improve dsDNA destabilization.

Figure 3.32. Different thermodynamic signatures of IFI16-PAAD mutants modulate dsDNA destabilization.
<table>
<thead>
<tr>
<th>Clone</th>
<th>% helix</th>
<th>Tm (°C)</th>
<th>ΔH (kcal/mol)</th>
<th>ΔS (kcal/mol/K)</th>
<th>ΔG&lt;sub&gt;25°C&lt;/sub&gt;&lt;sup&gt;folding&lt;/sup&gt; (kcal/mol)</th>
<th>n</th>
<th>Mutation 1</th>
<th>Mutation 2</th>
<th>Mutation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>44.7 ± 4.6</td>
<td>47.80 ± 1.18</td>
<td>-34.80 ± 4.87</td>
<td>-0.109 ± 0.015</td>
<td>-2.44 ± 0.51</td>
<td>3</td>
<td>I46N</td>
<td>L28F</td>
<td>K88E</td>
</tr>
<tr>
<td>6</td>
<td>44.2 ± 1.3</td>
<td>44.78 ± 0.75</td>
<td>-29.41 ± 2.64</td>
<td>-0.091 ± 0.008</td>
<td>-1.84 ± 0.18</td>
<td>3</td>
<td>I46N</td>
<td>V89E</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>44.0 ± 4.3</td>
<td>45.03 ± 2.86</td>
<td>-28.14 ± 2.32</td>
<td>-0.089 ± 0.007</td>
<td>-1.77 ± 0.34</td>
<td>3</td>
<td>I46N</td>
<td>A93T</td>
<td></td>
</tr>
<tr>
<td>17</td>
<td>41.5 ± 2.2</td>
<td>46.70 ± 0.90</td>
<td>-26.94 ± 1.65</td>
<td>-0.084 ± 0.005</td>
<td>-1.83 ± 0.18</td>
<td>3</td>
<td>I46N</td>
<td>T81A</td>
<td>K88E</td>
</tr>
<tr>
<td>15</td>
<td>41.2 ± 0.9</td>
<td>39.47 ± 1.21</td>
<td>-23.44 ± 2.47</td>
<td>-0.076 ± 0.007</td>
<td>-1.36 ± 0.40</td>
<td>3</td>
<td>I46S</td>
<td>R96G</td>
<td></td>
</tr>
<tr>
<td>Wild Type</td>
<td>40.4 ± 0.2</td>
<td>42.43 ± 0.50</td>
<td>-27.17 ± 2.50</td>
<td>-0.086 ± 0.008</td>
<td>-1.52 ± 0.12</td>
<td>3</td>
<td>none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>16</td>
<td>39.8 ± 0.5</td>
<td>33.63 ± 0.88</td>
<td>-27.34 ± 8.95</td>
<td>-0.089 ± 0.029</td>
<td>-0.89 ± 0.20</td>
<td>3</td>
<td>I46S</td>
<td>E15K</td>
<td></td>
</tr>
<tr>
<td>30</td>
<td>39.8 ± 1.0</td>
<td>35.72 ± 0.58</td>
<td>-25.93 ± 5.27</td>
<td>-0.084 ± 0.017</td>
<td>-1.14 ± 0.07</td>
<td>3</td>
<td>I46N</td>
<td>E15K</td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>38.9 ± 0.9</td>
<td>46.74 ± 2.41</td>
<td>-28.32 ± 3.13</td>
<td>-0.089 ± 0.009</td>
<td>-1.94 ± 0.41</td>
<td>3</td>
<td>I46N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>27</td>
<td>38.6 ± 0.5</td>
<td>43.41 ± 0.99</td>
<td>-24.94 ± 8.10</td>
<td>-0.079 ± 0.026</td>
<td>-1.45 ± 0.44</td>
<td>3</td>
<td>E15K</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>22.9 ± 2.0</td>
<td>31.38 ± 0.61</td>
<td>-20.51 ± 1.04</td>
<td>-0.067 ± 0.003</td>
<td>-0.77 ± 0.07</td>
<td>3</td>
<td>I46N</td>
<td>E15K</td>
<td>L28P</td>
</tr>
<tr>
<td>11</td>
<td>20.1 ± 4.2</td>
<td>31.48 ± 2.11</td>
<td>-58.78 ± 56.52</td>
<td>-0.193 ± 0.185</td>
<td>-1.36 ± 1.41</td>
<td>3</td>
<td>I46N</td>
<td>R96G</td>
<td>I17V</td>
</tr>
</tbody>
</table>

Table 3-5. Secondary structure and thermodynamic parameters of IFI16-PAAD mutants after directed evolution.

The percent helix was calculated by cdPRO; far UV spectra. ΔH, ΔS, and ΔG<sub>25°C</sub><sup>folding</sup> were calculated by VanHoff's analysis of thermal denaturation. Thermal denaturation data were collected and averaged with standard deviation with three independent trials (n = 3). Far UV spectra were collected three times (n = 3) and the secondary structure percentage was calculated for each trial and averaged with standard deviation.
<table>
<thead>
<tr>
<th>Clone</th>
<th>$\Delta G_{\text{folding}}^{25^\circ C}$ (kcal/mol)</th>
<th>$\Delta G_{\text{H}_2\text{O}}^{25^\circ C}$ (kcal/mol)</th>
<th>Tm (°C)</th>
<th>$\Delta H$ (kcal/mol)</th>
<th>$\Delta S$ (kcal/mol/K)</th>
<th>$m$ (kcal/mol/M)</th>
<th>$C_m$ (M)</th>
<th>$\Delta ASA$ ($\AA^2$)</th>
<th>Mutation 1</th>
<th>Mutation 2</th>
<th>Mutation 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>-2.44 ± 0.51</td>
<td>-2.22</td>
<td>47.80 ± 1.18</td>
<td>-34.80 ± 4.87</td>
<td>-0.109 ± 0.015</td>
<td>-1.12</td>
<td>1.99</td>
<td>1186</td>
<td>I46N</td>
<td>L28F</td>
<td>K88E</td>
</tr>
<tr>
<td>6</td>
<td>-1.84 ± 0.18</td>
<td>-2.26</td>
<td>44.78 ± 0.75</td>
<td>-29.41 ± 2.64</td>
<td>-0.091 ± 0.008</td>
<td>-1.12</td>
<td>2.01</td>
<td>1186</td>
<td>I46N</td>
<td>V89E</td>
<td></td>
</tr>
<tr>
<td>Wild Type</td>
<td>-1.52 ± 0.12</td>
<td>-1.75</td>
<td>42.43 ± 0.50</td>
<td>-27.17 ± 2.50</td>
<td>-0.086 ± 0.008</td>
<td>-1.7</td>
<td>1.03</td>
<td>3822</td>
<td>none</td>
<td></td>
<td></td>
</tr>
<tr>
<td>22</td>
<td>-1.94 ± 0.41</td>
<td>-1.56</td>
<td>46.74 ± 2.41</td>
<td>-28.32 ± 3.13</td>
<td>-0.089 ± 0.009</td>
<td>-0.96</td>
<td>1.64</td>
<td>3550</td>
<td>I46N</td>
<td></td>
<td></td>
</tr>
<tr>
<td>29</td>
<td>-0.77 ± 0.07</td>
<td>-1.09</td>
<td>31.38 ± 0.61</td>
<td>-20.51 ± 1.04</td>
<td>-0.067 ± 0.003</td>
<td>-0.76</td>
<td>1.44</td>
<td>2640</td>
<td>I46N</td>
<td>E15K</td>
<td>L28P</td>
</tr>
</tbody>
</table>

**Table 3-6.** Thermodynamics and chemical stability parameters of four IFI16-PAAD mutants.

$\Delta G_{\text{folding}}^{25^\circ C}$ of the four selected mutants was obtained by chemical denaturation and monitored by CD at 222 nm. $C_m$ and $\Delta ASA$ were calculated by the correlation proposed by Myers (*Myers et al., 1995*). Thermal denaturation data were collected and averaged with standard deviation with three independent trials ($n = 3$). Chemical denaturation spectra were collected at UV 222 nm for 30 seconds.
Table 3-7. Summary of RPA nucleic acid binding properties for the PAAD and HIN200 domain of MNDA and IFI16. $Dalal M. Sc Thesis.  § Yan M.Sc Thesis

<table>
<thead>
<tr>
<th></th>
<th>MNDA-PAAD</th>
<th>MNDA-HIN200</th>
<th>$IFI16-PAAD</th>
<th>§IFI16-HIN200α</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binds</td>
<td>ssDNA, dsDNA, RNA</td>
<td>ssDNA, dsDNA</td>
<td>ssDNA, dsDNA, RNA</td>
<td>ssDNA, dsDNA, RNA</td>
</tr>
<tr>
<td>ssDNA/dsDNA ratio</td>
<td>~1</td>
<td>~1</td>
<td>$10^1$ ~ $10^2$</td>
<td>$10^2$ ~ $10^3$</td>
</tr>
<tr>
<td>Nucleotide types preference</td>
<td>No</td>
<td>No</td>
<td>G, C</td>
<td>G, C</td>
</tr>
<tr>
<td>Binding mode</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>Oligomerization</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>wrap/stretch ssDNA</td>
<td>No</td>
<td>Compact/extend ssDNA</td>
<td>Compact ssDNA</td>
<td>Compact/extend ssDNA</td>
</tr>
<tr>
<td>Destabilize dsDNA</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>No</td>
</tr>
</tbody>
</table>
4: DISCUSSION

4.1 Different structural properties and stability within the PAAD domain family

Although only a few PAAD domain structures have been solved, they can generally be divided into two groups. The first group has a complete six helix bundle such as MNDA-PAAD and ASC2-PAAD (PDB code: 2DBG, 2HM2) and the second group has a disordered region along with five helices, namely NALP1-PAAD and NALP10-PAAD (PDB code: 1PN5, 2DO9). Previously Dalal has used helix inducers such as SDS and TFE to increase the amount of helical content in IFI16-PAAD (Dalal & Pio, 2006). They also compared the thermodynamic parameters of the PAAD domain to another death domain family, CARD, and concluded that the PAAD domain is generally less stable with higher conformational entropy compared to CARD family members. This is consistent with the fact that some PAAD domains are partially folded. Moreover, it has been speculated that the unique partially folded structure of the PAAD domain might confer new protein-ligand interactions as they discovered that IFI16-PAAD is a novel single stranded DNA binding protein. With the extensive characterization of the PAAD domain of IFI16 by Dalal, we decided to characterize another PAAD domain family member to make a general comparison regarding the structural properties and protein-ligand interactions. We chose the PAAD domain of MNDA because IFI16-PAAD has the highest sequence identity with MNDA-PAAD.
and we would like to know if high sequence identity correlates with similar structural properties, stability and function. The NMR structure of MNDA-PAAD shows that it has a complete six helix bundle, making it an ideal candidate to compare the thermodynamic parameters with IFI16-PAAD. In addition, MNDA-PAAD with moderate expression levels and purification recovery made it a promising candidate for further investigation. As expected, the secondary structure assessment and thermodynamic parameters suggested that MNDA-PAAD has higher helical content as well as higher stability compared to IFI16-PAAD. It is important to note that after thermal denaturation of the PAAD domains, IFI16-PAAD is able to regain its secondary structure when it cooled back to 20°C while MNDA-PAAD cannot. This suggests that MNDA-PAAD may be trapped in the unfolded state upon thermal denaturation, and other folding elements such as chaperones may be needed in order to refold its structure. To characterize the structural integrity of MNDA-PAAD further, we used urea to denature the protein and observed the change in the intrinsic fluorescence property. The decrease in tyrosine fluorescence intensity with a blue shift (304 nm – 302 nm) indicates the protein has denatured and exposed the tyrosine residue to the non-polar environment (Yang, Yang, Zhang, Yu, & An, 2007). Although MNDA-PAAD structure seems to be well folded and stable, we tried to enhance its stability by introducing small molecules such as osmolytes (data not shown). Osmolytes are small organic molecules that are present in the cell to protect and promote proteins from denaturation (Bolen & Rose, 2008). Osmolytes are often used as additives in protein crystallization to stabilize the protein. The secondary structure and thermal stability of MNDA-PAAD remain unchanged.
in the presence of osmolytes, suggesting that the protein is either already well
folded and in its native state or the osmolyte we tested did not bind to the
protein and additional screening is needed. On the other hand, IFI16-PAAD’s
secondary structure and stability can be modified by helix promoting agent
such as SDS or TFE as demonstrated by Dalal, illustrating the flexibility of
IFI16-PAAD. The speculation regarding IFI16-PAAD which adopts a partially
folded structure confers new protein-ligand interaction was also carefully
examined with a comparison to MNDA-PAAD. In brief, the NMR structure of
MNDA-PAAD shows that it has an ordered helix 3 and here we show that it
has the ability to interact with nucleic acids. However, there are distinctive
DNA binding properties between the two PAAD domains and it could be
because of their difference in structural properties. A more detailed discussion
can be seen in section 4.2.

4.2 Comparison of nucleic acid binding properties between IFI16 and MNDA

It was hypothesized by Dalal that IFI16-PAAD may have diverged
function in the PAAD domain family as its structure differs from other PAAD
domain family members. Dalal identified IFI16-PAAD as a novel single
stranded nucleic acid binding protein as IFI16-PAAD exhibits ssDNA binding
protein properties, which include higher affinity to single stranded DNA than
double stranded DNA, having nucleotides preferences, compaction
(wrapping) of ssDNA, oligomerization upon binding to ssDNA, and dsDNA
destabilization. These ssDNA binding properties are conserved for ssDNA
binding proteins with OB fold such as Replication Protein A (RPA). The
HIN200 domain of IFI16 (α domain) was also identified to be a RPA like
protein exhibiting OB fold binding properties (Yan et al., 2008), yet there are several distinct differences of the OB fold DNA binding properties between the two domains. IFI16-HIN200 (α domain) can extend and compact ssDNA and recognizes ssDNA in the same orientation as RPA (3' to 5'). These properties strongly support the conclusion that the HIN200 domain of IFI16 is a RPA-like, nucleic acid binding protein with an OB fold. However, a critical property, the ability to destabilize dsDNA, is missing from IFI16-HIN200 but present in RPA. RPA destabilizes and unwinds dsDNA to the single stranded form during DNA replication initiation (Brosh et al., 2000; Treuner et al., 1996), a common property that most ssDNA binding proteins has. It is important to note that the PAAD domain of IFI16 has the ability to destabilize dsDNA, therefore, the full length of IFI16 would have most of the OB fold properties. IFI16 strong preference to ssDNA and destabilize dsDNA may assist DNA repair in the BASC complex by first releases the damaged DNA into single stranded form and stabilizes the ssDNA until the repair machinery become active. It is not known whether this functional complementation is conserved in other HIN200 family members that contain both PAAD and HIN-200 domains; therefore, we explored the DNA binding properties of another HIN200 family member, MNDA, which has not been shown to be involved in DNA repair or DNA replication.

In the case of MNDA, the PAAD domain also exhibits ssDNA binding properties but differed from IFI16-PAAD. MNDA-PAAD does not have a preference for particular nucleic acids nor single stranded or double stranded DNA, does not oligomerize upon binding to ssDNA, does not extend or pack ssDNA, but it does destabilize dsDNA. Here we demonstrated that two PAAD
domains with high sequence identity and similar structure recognized nucleic acid differently. The findings that both PAAD domains interact with nucleic acid destabilize dsDNA are consistent with the fact that IFI16 is involved in DNA repair and transcriptional repression (Aglipay et al., 2003). Although there is no evidence showing MNDA is involved in DNA repair, Xie et al. provided a link for MNDA to transcriptional repression as MNDA binds to transcription factor YY1 in directing transcriptional repression (Xie et al., 1998). The work I have done with MNDA sheds light on how MNDA interacts with nucleic acid and leads us to conclude that the protein-nucleic acid properties are conserved between the two PAAD domains with minor differences. As previously mentioned, the amino acid sequence between IFI16-PAAD and MNDA-PAAD is well conserved, and this conservation in amino acid sequence underlies their conserved DNA binding property. Luscombe and Thornton demonstrated that specific amino acids that interacts with nucleic acid tends to be well conserved within families (Luscombe & Thornton, 2002). The differences in binding property are thought to be the result of their structural differences, where IFI16-PAAD contains a disordered helix 3 but MNDA-PAAD does not. The disordered helix 3 may generate more flexibility on the surface of the protein to form protein-nucleic acid interaction.

EMSA is widely used to detect protein-nucleic acid interactions because of its simplicity. Protein-nucleic acid complex would cause a shifted band that migrates differently to the probe but for MNDA, the putative complex migrated only at most, 2mm into the gel. However, it is a well-known problem for protein complexes, which could be the reason for the inaccurate binding constants determined in our EMSA experiment. Protein aggregation,
oligomerization, charges of the complex can be the cause but this problem has recently been solved by using a blue native PAGE gel (Hon, 2009) that can change the charge of the complex (Wittig, Braun, & Schagger, 2006). To verify the existence of protein-nucleic acid complex we used UV cross linking. UV irradiates thymine to form free radicals and these free radicals allow for covalent bond with adjacent amino acids to generate DNA-protein cross links (Perrier et al., 2006). The UV cross linking experiment shows MNDA-PAAD only forms complex with ssDNA but not dsDNA. However, fluorescence quenching and dsDNA melting experiments support the capability of MNDA-PAAD binds to dsDNA. A plausible explanation to the failure to detect dsDNA complex with UV cross linking is that the bases in a double stranded nucleic acids hold together by base stacking and hydrogen bonding. Therefore, it makes UV harder to irradiate the paired thymine to form free radicals and cross link to the protein. The affinity of MNDA-PAAD to ssDNA and dsDNA determined from fluorescence quenching experiments suggests that MNDA-PAAD has roughly the same affinity for both (micro molar range). In addition, MNDA-PAAD causes more than 20°C decrease in dsDNA melting point, which denotes the ability of MNDA-PAAD to bind to dsDNA despite the fact that UV failed to cross link MNDA-PAAD to dsDNA. It is also important to note that the binding constants determined by fluorescence quenching might not represent the real binding constants as other factors could also have contributed to the quenching of the fluorescence during titration. One example could be the addition of ligand induced a conformational change of the protein, hence exposing the fluorescence residues without direct interactions with the ligands. We cannot rule out this possibility because NMR
spectroscopy clearly showed the protein was denatured after the addition of ssDNA. Nevertheless, we ruled out the possibility that contamination or aggregation is the cause of the protein-DNA complex formation by western blot and mass spectrometry. Moreover, I have demonstrated the ssDNA binds to specific region on MNDA-PAAD by NMR titration. The peaks gradually moved as MNDA-PAAD is titrated and this gradual movement is a characteristic of fast exchange, which indicates the ligand’s binding affinity is in the micro molar range and is considered as weak binding (Lian, Barsukov, Sutcliffe, Sze, & Roberts, 1994). This result is consistent with the finding from fluorescence quenching experiment as both techniques agree that MNDA-PAAD weakly binds to nucleic acid.

Here we also characterized the DNA binding properties of the HIN200 domain of MNDA. The results indicate that MNDA-HIN200 recognize both single stranded and double stranded DNA, however MNDA-HIN200 has no preference to either one. MNDA-HIN200 shows a slightly higher binding affinity to polythymine nucleotides, which may suggest that MNDA-HIN200 binds to regions rich with thymine, such as a TATA box. MNDA-HIN200 also shows no oligomerization upon binding to ssDNA but it can extend and compact ssDNA. We chose to characterize these properties because IFI16 HIN200 domain was demonstrated to be an RPA like protein and these properties are signatures of RPA protein. IFI16-PAAD was shown to have functional complementary with IFI16 HIN200 domain regarding the ability to destabilize dsDNA, which fulfilled the properties for IFI16 to become an ssDNA binding protein like RPA. However, I have found that MNDA-PAAD and MNDA-HIN200 both exhibit very distinct DNA binding properties.
compared to RPA. Nevertheless, MNDA-PAAD failed to extend or compact ssDNA but this property is redeemed by MNDA-HIN200. Although MNDA may not act as a RPA like protein similar to IFI16 for DNA repair, MNDA with its unique DNA binding properties may have an important role in transcription process as MNDA was found to interact with transcription factor YY1.

During transcription, transcription factors recruit other regulatory proteins to make up the transcription machinery such as transcriptional activator or repressor, and RNA polymerase. For example, in yeast, the transcription machinery compose of RNA polymerase II, transcription factors II, suppressor of RNA polymerase B mutations and Gal11, each of which has their own function that contributes to the initiation of transcription (Ptashne & Gann, 1997). The RNA polymerase in eukaryotes does not unwind the DNA duplex to initiate transcription, instead the transcription factor performs the duty, specifically transcription factor II H (Lee & Young, 1998; Ptashne & Gann, 1997). Transcription factors generally consist of TBP (TATA Binding Protein) and TAFs (TBP Associated Factors). The role of TBP is to bind to the core promoter region including the TATA element where TAFs assist TBP binding in this process in basal transcription. TAFs can serve as coactivators, function in promoter recognition or modify general transcription factors (GTFs) to facilitate complex assembly and transcription initiation (Thomas & Chiang, 2006). MNDA binds to transcription factor YY1 and nucleic acids, and only the N-terminal region (presumably the PAAD domain) of MNDA is required for the enhanced affinity of YY1 to DNA (Xie et al., 1998). Here, we provide evidence that MNDA-PAAD can destabilize/separate double stranded DNA, therefore MNDA may act as a general cofactor to assist strand separation for
transcription factor YY1 to bind to ssDNA during transcription initiation. Although MNDA-PAAD recognizes nucleic acid with weak affinity (μM) and no preference to single stranded DNA, a recently proposed DNA binding model of the Lac repressor shows that non-specific DNA binding was a necessary step in order to position the DNA binding domain for a high affinity interaction with the cognate recognition sequence in the Lac operator (Kalodimos et al., 2004).

4.3 IFI16-PAAD mutants with different thermodynamics signature obtained by directed evolution can modulate nucleic acids destabilization

This project was initially started by Dalal and the initial goal was to find mutations in IFI16-PAAD that could re-order the disordered region. Proteins require a properly folded three-dimensional structure to function, and mutations that alter the protein structure can be deleterious to the protein function. Generally, mutations occur naturally under selective pressure to enhance the protein activity or stability. However the native conformation of IFI16-PAAD can be modulated by different chemicals, which suggests its flexibility in its helix 3 region (Dalal & Pio, 2006). Here we have identified four amino acid mutations that have the greatest impact on IFI16-PAAD secondary structure and stability using a combination of error prone PCR, DNA shuffling and GFP selection. The generation of IFI16-PAAD mutants DNA library was completed by Dalal and my contribution was to verify each mutant’s secondary structure and stability using both circular dichorism and fluorescence measurement. The selection of mutants was based on increased fluorescence intensity of the GFP since fluorescence intensity usually
attributes to solubility. Only the 40 brightest colonies were selected and moved into the next round of mutagenesis to achieve maximal but manageable sample size. The mutagenesis stopped when the mutations converged towards specific locations, but it is important to note that none of the mutations lies in the helix 3 region. The four mutations (E15K, L28P, I46N and K88E) identified to have the highest impact on IFI16-PAAD structure and stability could change the helix propensity, hence promote a more structured protein. Strikingly, most of the disorder region prediction software such as Globplot, DISOPRED, VL3H-Disport, DisEMBL, RONN (See Appendix I) predicted that IFI16-PAAD contains no disordered helix 3. These data may imply that helix 3 has the potential to fold into an α-helix but is prevented by the lack of entropic effect. If the disordered helix 3 of IFI16-PAAD is expected to refold into an α-helix, its secondary structure should be similar to MNDA-PAAD. However, the changes of the secondary structure from the mutants are subtle and helix 3 may not yet refold into an α-helix. Instead, the subtle increase in helical content of mutant 4 can be explained by helix capping. The N-terminus of helix 1 is very close to the K88E mutation and the negatively charged glutamic acids may neutralize the helix dipole moment created by the carbonyl group of the peptide bond, therefore stabilizing the helix. I46N/S is presented in all of the mutants regardless of increased or decreased structure and stability, which suggests the importance of this amino acid. An intriguing mutation that leads to decrease secondary structure and stability is E15K. The positively charged lysine is positioned very close to the C-terminus of disordered helix 3 and this in fact could form favourable C-terminal helix capping by stabilizing the negative charge dipole at the C-terminal of alpha
helix. However, these results suggest that E15K does not improve the protein structure but degrades it, which also indicates that improving solubility does not always translate into better folded structure. Nevertheless, a correlation between improving protein structure with enhancing stability can be seen from both thermo and chemical stability data. In an attempt to determine if more stable proteins confer higher functional capacity, we demonstrated that the two mutants with the most improved structure and stability also destabilize DNA duplex to a greater extent compared to the wild type of IFI16-PAAD. If the preceding speculation of helix capping is true, it also implies that helix 1 may be important in DNA duplex destabilization.
My effort in characterizing the structural property of the PAAD domain of MNDA and IFI16 has provided insight to the cause of the partially folded structure of IFI16-PAAD and the effect of the partially folded structure to the nucleic acids destabilization property. I went even further in characterizing the nucleic acids binding properties of the PAAD domain and HIN200 domain of MNDA to shed light on the conservation property between the two HIN-200 family members. Previously in the literature, the HIN-200 family member, MNDA, was only recorded to bind to nucleic acids non-specifically and without any extensive details regarding the interaction. However, the recent discovery that another HIN-200 family member, IFI16, interacts with nucleic acids similarly to ssDNA binding protein RPA has motivated my investigation on MNDA nucleic acids binding property.

The PAAD domain is very intriguing because of its different levels of folding. IFI16-PAAD was thought to have a partially folded structure and it is further supported by the thermodynamic parameters and comparative modelling comparison to MNDA-PAAD. The solved structure of MNDA-PAAD (PDB code: 2DBG) contains a globular structure with six helices packed in a greek key motif is superimposed to IFI16-PAAD model and it clearly reveals the disordered helix 3 of IFI16-PAAD. It is remarkable that two proteins with high sequence identity adopt such folding difference. The folding difference
seen in the structure also translates into the difference in thermodynamic parameters, where IFI16-PAAD was found to be lower in stability and helical content.

Besides characterizing the structural property, the interaction with nucleic acids was also studied extensively in this thesis. Although it is not the first report to illustrate MNDA has the ability to interact with nucleic acids, my work has provided details on MNDA binding properties with its individual domains. Our discovery that the PAAD domain also interacts with nucleic acids breaks the barrier that the PAAD domain was solely a protein-protein interaction domain involved in signalling pathway during apoptosis and inflammation. Understanding the DNA binding properties of the HIN-200 members allows us to demonstrated functional complementation of the HIN-200 family as shown by the ability to destabilize dsDNA of IFI16-PAAD but not IFI16-HIN200. Although neither of the MNDA domains exhibits functional complementation regarding dsDNA destabilization, functional complementation remains in MNDA with respect to compact and extend of ssDNA from the HIN200 domain but not the PAAD domain. Moreover, we showed the individual domain of MNDA recognizes nucleic acids by EMSA and UV cross linking and the ability to extend or compact ssDNA by FRET. Fluorescence quenching has given us the range of affinity of the protein to nucleic acids. Our effort also prompts us to investigate critical residue of the PAAD domain involved in the DNA interactions. Although more work must be done in order to identify the amino acids involved, we revealed by NMR that the MNDA-PAAD DNA binding is specific to certain amino acids and the affinity is between $10^{-7}$ to $10^{-6}$ M. Comparing the DNA binding properties of
IFI16 to MNDA reveals not only functional complementation, but the ability to oligomerize may account for the difference in binding constants.

In an attempt to identify residues that are responsible for the unstructured region of the PAAD domain of IFI16, we identified four amino acids (E15, L28, I46, K88) that have the most impact on IFI16-PAAD structure as well as the ability to destabilize dsDNA. We found that mutants that are more structured and with higher stability are consistently more potent in the ability to destabilize dsDNA and vice versa. These mutants need to be further characterized regarding the nucleic acids binding properties.

The literatures has uncovered the involvement of HIN-200 proteins in many biological pathways such as cell cycle arrest, DNA repair, DNA replication and transcription initiation (Asefa et al., 2004; Ludlow et al., 2005; Y. Wang et al., 2000; Xie et al., 1998; Xin et al., 2003). Here we have provided direct evidences showing the HIN200 proteins interact with nucleic acids with specificities. Chemical shift mapping has provided a preliminary result that ssDNA interacts with PAAD domain in a specific region. To further identify the critical residues responsible for the interaction, 3D NMR by labelling the protein with $^{13}$C, $^{15}$N and $^1$H will reveal the amino acids identity of the shifted peaks. The fact that HIN-200 proteins have the ability to destabilize dsDNA raises the possibility that it also unwinds dsDNA. Since there is evidence showing IFI16 is involved in BASC complex, an unwinding assay to demonstrate the ability to unwind dsDNA as well as to remove DNA secondary structure will provide a strong linkage to DNA repair pathway. This is currently undertaken in our laboratory and our preliminary data suggests that there is no correlation between the ability to unwind and to destabilize
dsDNA (Hon, 2009). Moreover, the tumor suppressor BRCA1 accumulates rapidly in nuclear foci in response to induction DNA breaks suggests that BRCA1 may function in early step in repair pathway (Schlegel, Jodelka, & Nunez, 2006). Since IFI16 interacts with BRCA1 (Aglipay et al., 2003), having an antibody to target IFI16 and ssDNA and look for co-localization in cells may provide further insight about the role of IFI16 in DNA repair pathway (Dalal, 2006).
APPENDICES

Appendix A: EMSA curve fitting parameters of MNDA-PAAD.

A25
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value Not Necessary
Conclusion (alpha = 0.05) Simpler model fits better
Preferred model One site binding (hyperbola)

F (DFn, DFd)

One site binding (hyperbola)  Two site binding (hyperbola)
Best-fit values
BMAX  1.306  BMAX10.6029
KD  88.08  KD1  88.60
Std. Error
BMAX  0.1225  BMAX20.7027
KD  21.59  KD2  87.59
95% Confidence Intervals
BMAX  1.023 to 1.588  BMAX1 0.0 to 3.473e+006
KD  38.30 to 137.9  KD1  0.0 to 5.739e+006
Goodness of Fit
Degrees of Freedom  8  Degrees of Freedom  6
R squared   0.9703  R squared   0.9703
Absolute Sum of Squares  0.03823  Absolute Sum of Squares  0.03823
Sy.x  0.06913  Sy.x  0.07982
Constraints
BMAX  BMAX > 0.0  BMAX1BMAX1 > 0.0
KD  KD > 0.0  KD1 KD1 > 0.0

Data
Number of X values  10
Number of Y replicates  1
Total number of values  10
Number of missing values  0
Comparison of Fits

Null hypothesis: One site binding (hyperbola)
Alternative hypothesis: Two site binding (hyperbola)
P value: 0.0651
Conclusion (alpha = 0.05): Do not reject null hypothesis
Preferred model: One site binding (hyperbola)

F (DFn, DFd): 4.456 (2,6)

One site binding (hyperbola)
Best-fit values
BMAX: 0.9204
KD: 9.049
Std. Error
BMAX: 0.1139
KD: 4.860
95% Confidence Intervals
BMAX: 0.6577 to 1.183
KD: 0.0 to 20.26
Goodness of Fit
Degrees of Freedom: 8
R squared: 0.7208
Absolute Sum of Squares: 0.2516
Sy.x: 0.1773
Constraints
BMAX > 0.0
KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX: 10.4165
KD1: 0.9081
BMAX: 20.9326
KD2: 144.5
Std. Error
BMAX: 0.1779
KD1: 1.324
BMAX: 20.3751
KD2: 211.6
95% Confidence Intervals
BMAX: 0.0 to 0.8517
KD1: 0.0 to 4.149
BMAX: 20.01463 to 1.850
KD2: 0.0 to 662.2
Goodness of Fit
Degrees of Freedom: 6
R squared: 0.8876
Absolute Sum of Squares: 0.1012
Sy.x: 0.1299
Constraints
BMAX1 > 0.0
BMAX2 > 0.0
KD1 > 0.0
KD2 > 0.0

Data
Number of X values: 10
Number of Y replicates: 1
Total number of values: 10
Number of missing values: 0
## GC-3

### Comparison of Fits

- **Null hypothesis**: One site binding (hyperbola)
- **Alternative hypothesis**: Two site binding (hyperbola)
- **P value**: Not Necessary
- **Conclusion (alpha = 0.05)**: Other model didn’t converge

**Preferred model**: One site binding (hyperbola)

**F (DFn, DFd)**

<table>
<thead>
<tr>
<th>Model</th>
<th>BMAX</th>
<th>KD</th>
<th>Std. Error</th>
<th>95% Confidence Intervals</th>
<th>Goodness of Fit</th>
</tr>
</thead>
<tbody>
<tr>
<td>One site binding</td>
<td>1.466</td>
<td>146.0</td>
<td>0.1608</td>
<td>1.073 to 1.859</td>
<td>Degrees of Freedom 6</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>35.83</td>
<td>58.33 to 233.7</td>
<td>R squared 0.9807</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absolute Sum of Squares 0.02141</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sy.x 0.05973</td>
</tr>
<tr>
<td>Two site binding</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Constraints**

- BMAX > 0.0
- KD > 0.0

---

### Data

- **Number of X values**: 8
- **Number of Y replicates**: 1
- **Total number of values**: 8
- **Number of missing values**: 0
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value 0.3116
Conclusion (alpha = 0.05) Do not reject null hypothesis
Preferred model One site binding (hyperbola)

One site binding (hyperbola)

Best-fit values
BMAX 1.108
KD 34.66
Std. Error
BMAX 0.05986
KD 6.260
95% Confidence Intervals
BMAX 0.9661 to 1.249
KD 19.86 to 49.47

Goodness of Fit
Degrees of Freedom 7
R squared 0.9765
Absolute Sum of Squares 0.02388
Sy.x 0.05841

Constraints
BMAX BMAX > 0.0
KD KD > 0.0

Two site binding (hyperbola)

Best-fit values
BMAX1 0.07180
KD1 1.000e-007
BMAX2 21.079
KD2 46.42

Std. Error
BMAX1 0.09615
KD1 2.344
BMAX2 0.08620
KD2 16.74

95% Confidence Intervals
BMAX1 0.0 to 0.3190
KD1 0.0 to 6.025
BMAX2 0.8569 to 1.300
KD2 3.390 to 89.45

Goodness of Fit
Degrees of Freedom 5
R squared 0.9853
Absolute Sum of Squares 0.01498
Sy.x 0.05474

Constraints
BMAX1 BMAX1 > 0.0
KD1 KD1 > 0.0
BMAX2 BMAX2 > 0.0
KD2 KD2 > 0.0

Data
Number of X values 9
Number of Y replicates 1
Total number of values 9
Number of missing values 0
Random RNA

Comparison of Fits

Null hypothesis  One site binding (hyperbola)
Alternative hypothesis  Two site binding (hyperbola)
P value  0.1068
Conclusion (alpha = 0.05)  Do not reject null hypothesis
Preferred model  One site binding (hyperbola)

F (DFn, DFd)  8.366 (2,2)

One site binding (hyperbola)

Best-fit values
BMAX  1.007
KD  1.764

Std. Error
BMAX  0.1044
KD  0.6782

95% Confidence Intervals
BMAX  0.7170 to 1.296
KD  0.0 to 3.646

Goodness of Fit
Degrees of Freedom  4
R squared  0.8351
Absolute Sum of Squares  0.04627
Sy.x  0.1076

Constraints
BMAX  BMAX > 0.0
KD  KD > 0.0

Two site binding (hyperbola)

Best-fit values
BMAX1  0.3528
KD1  0.000e-007
BMAX2  0.9915
KD2  11.20

Std. Error
BMAX1  0.2095
KD1  0.3762
BMAX2  0.1801
KD2  11.22

95% Confidence Intervals
BMAX1  0.0 to 1.254
KD1  0.0 to 1.619
BMAX2  0.2165 to 1.767
KD2  0.0 to 59.46

Goodness of Fit
Degrees of Freedom  2
R squared  0.9824
Absolute Sum of Squares  0.004940
Sy.x  0.04970

Constraints
BMAX1  BMAX1 > 0.0
KD1  KD1 > 0.0
BMAX2  BMAX2 > 0.0
KD2  KD2 > 0.0

Data
Number of X values  10
Number of Y replicates  1
Total number of values  6
Number of missing values  4
Appendix B: Tyrosine fluorescence quenching curve fitting parameters of MNDA-PAAD.

A25 – First trial
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value Not Necessary
Conclusion (alpha = 0.05) Other model didn’t converge
Preferred model One site binding (hyperbola)

F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
  BMAX 1.138
  KD 1.762
Std. Error
  BMAX 0.04318
  KD 0.1984
95% Confidence Intervals
  BMAX 1.044 to 1.232
  KD 1.330 to 2.194
Goodness of Fit
  Degrees of Freedom 12
  R squared 0.9789
  Absolute Sum of Squares 0.02566
  Sy.x 0.04624
Constraints
  BMAX BMAX > 0.0
  KD KD > 0.0

Two site binding (hyperbola) Didn’t Converge
Best-fit values
  BMAX1
  KD1
  BMAX2
  KD2
Std. Error
  BMAX1
  KD1
  BMAX2
  KD2
95% Confidence Intervals
  BMAX1
  KD1
  BMAX2
  KD2
Goodness of Fit
  Degrees of Freedom
  R squared
  Absolute Sum of Squares
  Sy.x
Constraints
  BMAX1 BMAX1 > 0.0
  KD1 KD1 > 0.0
  BMAX2 BMAX2 > 0.0
  KD2 KD2 > 0.0
Data
  Number of X values 14
  Number of Y replicates 1
  Total number of values 14
  Number of missing values 0
A25 – Second trial
One site binding (hyperbola)

Best-fit values
- BMAX 1.095
- KD 1.687

Std. Error
- BMAX 0.01813
- KD 0.08439

95% Confidence Intervals
- BMAX 1.055 to 1.134
- KD 1.504 to 1.871

Goodness of Fit
- Degrees of Freedom 12
- R squared 0.9955
- Absolute Sum of Squares 0.004704
- Sy.x 0.01980

Constraints
- BMAX BMAX > 0.0
- KD KD > 0.0

Data
- Number of X values 14
- Number of Y replicates 1
- Total number of values 14
- Number of missing values 0
### T25 - First trial

Comparison of Fits

<table>
<thead>
<tr>
<th>Null hypothesis</th>
<th>One site binding (hyperbola)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alternative hypothesis</td>
<td>Two site binding (hyperbola)</td>
</tr>
<tr>
<td>P value</td>
<td>Not Necessary</td>
</tr>
<tr>
<td>Conclusion (alpha = 0.05)</td>
<td>Other model didn't converge</td>
</tr>
<tr>
<td>Preferred model</td>
<td>One site binding (hyperbola)</td>
</tr>
</tbody>
</table>

#### F (DFn, DFd)

<table>
<thead>
<tr>
<th>One site binding (hyperbola)</th>
<th>Two site binding (hyperbola)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Best-fit values</td>
<td>Didn't Converge</td>
</tr>
<tr>
<td>BMAX 1.122</td>
<td>BMAX1</td>
</tr>
<tr>
<td>KD 1.517</td>
<td>KD1</td>
</tr>
<tr>
<td>Std. Error</td>
<td>BMAX2</td>
</tr>
<tr>
<td>BMAX 0.03049</td>
<td>KD2</td>
</tr>
<tr>
<td>KD 0.1298</td>
<td>Std. Error</td>
</tr>
<tr>
<td>95% Confidence Intervals</td>
<td>BMAX1</td>
</tr>
<tr>
<td>BMAX 1.056 to 1.188</td>
<td>KD1</td>
</tr>
<tr>
<td>KD 1.234 to 1.799</td>
<td>BMAX2</td>
</tr>
<tr>
<td>Goodness of Fit</td>
<td>KD2</td>
</tr>
<tr>
<td>Degrees of Freedom 12</td>
<td>95% Confidence Intervals</td>
</tr>
<tr>
<td>R squared 0.9873</td>
<td>BMAX1</td>
</tr>
<tr>
<td>Absolute Sum of Squares 0.01463</td>
<td>KD1</td>
</tr>
<tr>
<td>Sy.x 0.03491</td>
<td>BMAX2</td>
</tr>
<tr>
<td>Constraints</td>
<td>KD2</td>
</tr>
<tr>
<td>BMAX BMAX &gt; 0.0</td>
<td>Goodness of Fit</td>
</tr>
<tr>
<td>KD KD &gt; 0.0</td>
<td>Degrees of Freedom</td>
</tr>
<tr>
<td></td>
<td>R squared</td>
</tr>
<tr>
<td></td>
<td>Absolute Sum of Squares</td>
</tr>
<tr>
<td></td>
<td>Sy.x</td>
</tr>
<tr>
<td></td>
<td>Constraints</td>
</tr>
<tr>
<td></td>
<td>BMAX1 BMAX1 &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td>KD1 KD1 &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td>BMAX2 BMAX2 &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td>KD2 KD2 &gt; 0.0</td>
</tr>
<tr>
<td>Data</td>
<td></td>
</tr>
<tr>
<td>Number of X values 14</td>
<td></td>
</tr>
<tr>
<td>Number of Y replicates 1</td>
<td></td>
</tr>
<tr>
<td>Total number of values 14</td>
<td></td>
</tr>
<tr>
<td>Number of missing values 0</td>
<td></td>
</tr>
</tbody>
</table>
T25 – Second trial
Comparison of Fits
Null hypothesis   One site binding (hyperbola)
Alternative hypothesis   Two site binding (hyperbola)
P value   Not Necessary
Conclusion (alpha = 0.05) Simpler model fits better
Preferred model   One site binding (hyperbola)
F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
BMAX  1.084
KD 1.388
Std. Error
BMAX  0.009645
KD 0.04037
95% Confidence Intervals
BMAX  1.063 to 1.105
KD 1.300 to 1.476
Goodness of Fit
Degrees of Freedom 12
R squared  0.9985
Absolute Sum of Squares 0.001579
Sy.x 0.01147
Constraints
BMAX BMAX > 0.0
KD KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX1 0.5110
KD1 1.384
BMAX2 0.5728
KD2 1.391
Std. Error
BMAX1 0.009645
KD1 0.04037
BMAX2 0.009645
KD2 0.04037
95% Confidence Intervals
BMAX1 0.0 to 666560
KD1 0.0 to 8366
BMAX2 0.0 to 666561
KD2 0.0 to 7541
Goodness of Fit
Degrees of Freedom 10
R squared  0.9985
Absolute Sum of Squares 0.001579
Sy.x 0.01257
Constraints
BMAX1 BMAX1 > 0.0
KD1 KD1 > 0.0
BMAX2 BMAX2 > 0.0
KD2 KD2 > 0.0

Data
Number of X values 14
Number of Y replicates 1
Total number of values 14
Number of missing values 0
GC-3 First trial
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value Not Necessary
Conclusion (alpha = 0.05) Other model didn't converge
Preferred model One site binding (hyperbola)
F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
BMAX 1.088
KD 1.112
Std. Error
BMAX 0.01266
KD 0.04651
95% Confidence Intervals
BMAX 1.061 to 1.116
KD 1.010 to 1.213
Goodness of Fit
Degrees of Freedom 12
R squared 0.9969
Absolute Sum of Squares 0.003264
Sy.x 0.01649
Constraints
BMAX BMAX > 0.0
KD KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX1
KD1
BMAX2
KD2
Std. Error
BMAX1
KD1
BMAX2
KD2
95% Confidence Intervals
BMAX1
KD1
BMAX2
KD2
Goodness of Fit
Degrees of Freedom
R squared
Absolute Sum of Squares
Sy.x
Constraints
BMAX1 BMAX1 > 0.0
KD1 KD1 > 0.0
BMAX2 BMAX2 > 0.0
KD2 KD2 > 0.0
Data
Number of X values 14
Number of Y replicates 1
Total number of values 14
Number of missing values 0
**GC-3 Second trial**

Comparison of Fits

- **Null hypothesis**: One site binding (hyperbola)
- **Alternative hypothesis**: Two site binding (hyperbola)
- **P value**: Not Necessary
- **Conclusion (alpha = 0.05)**: Other model didn't converge
- Preferred model: One site binding (hyperbola)

<table>
<thead>
<tr>
<th>Function</th>
<th>Best-fit values</th>
<th>Std. Error</th>
<th>95% Confidence Intervals</th>
<th>Goodness of Fit</th>
<th>Constraints</th>
</tr>
</thead>
<tbody>
<tr>
<td>One site binding (hyperbola)</td>
<td>BMAX 1.139</td>
<td>KD 1.828</td>
<td>BMAX 1.067 to 1.210</td>
<td>Degrees of Freedom: 12</td>
<td>BMAX &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td>Std. Error</td>
<td></td>
<td>KD 1.493 to 2.162</td>
<td>R squared: 0.9881</td>
<td>KD &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absolute Sum of Squares: 0.01420</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sy.x: 0.03440</td>
<td></td>
</tr>
<tr>
<td>Two site binding (hyperbola) Didn't Converge</td>
<td>BMAX1</td>
<td>KD1</td>
<td>BMAX2</td>
<td>Degrees of Freedom:</td>
<td>BMAX1 &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td>KD2</td>
<td></td>
<td>R squared:</td>
<td>KD1 &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Absolute Sum of Squares:</td>
<td>BMAX2 &gt; 0.0</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Sy.x: 0.03440</td>
<td>KD2 &gt; 0.0</td>
</tr>
</tbody>
</table>

**Constraints**

- BMAX > 0.0
- KD > 0.0

**Data**

- Number of X values: 14
- Number of Y replicates: 1
- Total number of values: 14
- Number of missing values: 0
GC-5 First trial

Comparison of Fits

Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value Not Necessary

Conclusion (alpha = 0.05) Other model didn't converge
Preferred model One site binding (hyperbola)

F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
  BMAX 1.092
  KD 1.056
Std. Error
  BMAX 0.04079
  KD 0.1452
95% Confidence Intervals
  BMAX 1.003 to 1.181
  KD 0.7401 to 1.373
Goodness of Fit
  Degrees of Freedom 12
  R squared 0.9697
  Absolute Sum of Squares 0.03525
  Sy.x 0.05420
Constraints
  BMAX BMAX > 0.0
  KD KD > 0.0

Two site binding (hyperbola)
Didn't Converge
Best-fit values
  BMAX1
  KD1
  BMAX2
  KD2
Std. Error
  BMAX1
  KD1
  BMAX2
  KD2
95% Confidence Intervals
  BMAX1
  KD1
  BMAX2
  KD2
Goodness of Fit
  Degrees of Freedom
  R squared
  Absolute Sum of Squares
  Sy.x
Constraints
  BMAX1BMAX1 > 0.0
  KD1 KD1 > 0.0
  BMAX2BMAX2 > 0.0
  KD2 KD2 > 0.0

Data
  Number of X values 14
  Number of Y replicates 1
  Total number of values 14
  Number of missing values 0
**GC-5 Second trial**

One site binding (hyperbola)

Best-fit values
- BMAX 1.123
- KD 1.482

Std. Error
- BMAX 0.02404
- KD 0.1009

95% Confidence Intervals
- BMAX 1.071 to 1.176
- KD 1.262 to 1.702

Goodness of Fit
- Degrees of Freedom 12
- R squared 0.9918
- Absolute Sum of Squares 0.009275
- Sy.x 0.02780

Constraints
- BMAX BMAX > 0.0
- KD KD > 0.0

Data
- Number of X values 14
- Number of Y replicates 1
- Total number of values 14
- Number of missing values 0
dsAT
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value 0.5401
Conclusion (alpha = 0.05) Do not reject null hypothesis
Preferred model One site binding (hyperbola)
F (DFn, DFd) 0.6555 (2,10)

One site binding (hyperbola)
Best-fit values
  BMAX  101.7
  KD 1.804
Std. Error
  BMAX  2.523
  KD 0.1317
95% Confidence Intervals
  BMAX 96.21 to 107.2
  KD 1.517 to 2.091
Goodness of Fit
  Degrees of Freedom 12
  R squared 0.9902
  Absolute Sum of Squares 85.72
  Sy.x 2.673
Constraints
  BMAX  BMAX > 0.0
  KD  KD > 0.0

Two site binding (hyperbola)
Best-fit values
  BMAX1 15.115
  KD1 1.000e-007
  BMAX2 298.75
  KD2 2.124
Std. Error
  BMAX1 20.05
  KD1 0.042
  BMAX2 17.33
  KD2 0.8207
95% Confidence Intervals
  BMAX1 0.0 to 49.78
  KD1 0.0 to 2.321
  BMAX2 60.15 to 137.3
  KD2 0.2958 to 3.953
Goodness of Fit
  Degrees of Freedom 10
  R squared 0.9914
  Absolute Sum of Squares 75.78
  Sy.x 2.753
Constraints
  BMAX1  BMAX1 > 0.0
  KD1  KD1 > 0.0
  BMAX2  BMAX2 > 0.0
  KD2  KD2 > 0.0

Data
  Number of X values 14
  Number of Y replicates 1
  Total number of values 14
  Number of missing values 0
dsGC

Comparison of Fits

Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value 0.0014
Conclusion (alpha = 0.05) Reject null hypothesis
Preferred model Two site binding (hyperbola)

F (DFn, DFd) 13.56 (2,10)

One site binding (hyperbola)
Best-fit values
  BMAX 99.71
  KD 1.783
Std. Error
  BMAX 2.274
  KD 0.1201
95% Confidence Intervals
  BMAX 94.76 to 104.7
  KD 1.521 to 2.044
Goodness of Fit
  Degrees of Freedom 12
  R squared 0.9914
  Absolute Sum of Squares 70.37
  Sy.x 2.422
Constraints
  BMAX BMAX > 0.0
  KD KD > 0.0

Two site binding (hyperbola)
Best-fit values
  BMAX1 111.13
  KD1 1.000e-007
  BMAX2 293.67
  KD2 2.613
Std. Error
  BMAX1 8.141
  KD1 0.2087
  BMAX2 26.625
  KD2 0.4965
95% Confidence Intervals
  BMAX1 0.0 to 29.27
  KD1 0.0 to 0.4649
  BMAX2 78.91 to 108.4
  KD2 1.506 to 3.719
Goodness of Fit
  Degrees of Freedom 10
  R squared 0.9977
  Absolute Sum of Squares 18.95
  Sy.x 1.377
Constraints
  BMAX1 BMAX1 > 0.0
  KD1 KD1 > 0.0
  BMAX2 BMAX2 > 0.0
  KD2 KD2 > 0.0

Data
  Number of X values 14
  Number of Y replicates 1
  Total number of values 14
  Number of missing values 0
Appendix C: Tryptophan fluorescence quenching curve fitting parameters of MNDA-HIN200

A25
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value Not Necessary
Conclusion (alpha = 0.05) Other model didn't converge
Preferred model One site binding (hyperbola)
F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
BMAX 86.15
KD 2.758
Std. Error
BMAX 3.433
KD 0.2776
95% Confidence Intervals
BMAX 78.67 to 93.63
KD 2.153 to 3.363

Goodness of Fit
Degrees of Freedom 12
R squared 0.9832
Absolute Sum of Squares 103.1
Sy.x 2.932

Constraints
BMAX BMAX > 0.0
KD KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX1
KD1
BMAX2
KD2
Std. Error
BMAX1
KD1
BMAX2
KD2
95% Confidence Intervals
BMAX1
KD1
BMAX2
KD2

Goodness of Fit
Degrees of Freedom
R squared
Absolute Sum of Squares
Sy.x

Constraints
BMAX1 BMAX1 > 0.0
KD1 KD1 > 0.0
BMAX2 BMAX2 > 0.0
KD2 KD2 > 0.0

Data
Number of X values 14
Number of Y replicates 1
Total number of values 14
Number of missing values 0
**T25**
One site binding (hyperbola)

**Best-fit values**
- BMAX 107.0
- KD 0.5757

**Std. Error**
- BMAX 2.402
- KD 0.06000

**95% Confidence Intervals**
- BMAX 101.7 to 112.3
- KD 0.4437 to 0.7078

**Goodness of Fit**
- Degrees of Freedom 11
- R squared 0.9866
- Absolute Sum of Squares 130.1
- Sy.x 3.439

**Runs test**
- Points above curve 7
- Points below curve 5
- Number of runs 3
- P value (runs test) 0.0151E1

**Constraints**
- BMAX BMAX > 0.0
- KD KD > 0.0

**Data**
- Number of X values 13
- Number of Y replicates 1
- Total number of values 13
- Number of missing values 0
Comparison of Fits

Null hypothesis  One site binding (hyperbola)
Alternative hypothesis  Two site binding (hyperbola)
P value  Not Necessary
Conclusion (alpha = 0.05) Simpler model fits better
Preferred model  One site binding (hyperbola)
F (DFn, DFd)

One site binding (hyperbola)

Best-fit values
  BMAX  112.9
  KD 1.997

Std. Error
  BMAX  9.532
  KD 0.4771

95% Confidence Intervals
  BMAX  92.17 to 133.7
  KD 0.9571 to 3.036

Goodness of Fit
  Degrees of Freedom 12
  R squared  0.9163
  Absolute Sum of Squares 1111
  Sy.x 9.622

Constraints
  BMAX  BMAX > 0.0
  KD KD > 0.0

Two site binding (hyperbola)

Best-fit values
  BMAX1 192.65
  KD1 1.988
  BMAX2 220.35
  KD2 2.052

Std. Error
  BMAX1 12.633e+007
  KD1 11636
  BMAX2 22.633e+007
  KD2 50499

95% Confidence Intervals
  BMAX1 0.0 to 5.866e+007
  KD1 0.0 to 25927
  BMAX2 0.0 to 5.866e+007
  KD2 0.0 to 112515

Goodness of Fit
  Degrees of Freedom 10
  R squared  0.9163
  Absolute Sum of Squares 1111
  Sy.x 10.54

Constraints
  BMAX1 BMAX1 > 0.0
  KD1 KD1 > 0.0
  BMAX2 BMAX2 > 0.0
  KD2 KD2 > 0.0

Data
  Number of X values 14
  Number of Y replicates 1
  Total number of values 14
  Number of missing values 0
Comparison of Fits

Null hypothesis  One site binding (hyperbola)
Alternative hypothesis  Two site binding (hyperbola)
P value  Not Necessary

Conclusion (alpha = 0.05) Simpler model fits better
Preferred model  One site binding (hyperbola)

F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
BMAX  102.6
KD  1.920
Std. Error
BMAX  7.251
KD  0.3899
95% Confidence Intervals
BMAX  86.79 to 118.4
KD  1.071 to 2.770
Goodness of Fit
Degrees of Freedom  12
R squared  0.9382
Absolute Sum of Squares  667.3
Sy.x  7.457
Constraints
BMAX  BMAX > 0.0
KD  KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX1 183.48
KD1 1.915
BMAX2 219.14
KD2 1.950
Std. Error
BMAX1 17.805e+007
KD1 23818
BMAX2 27.805e+007
KD2 100680
95% Confidence Intervals
BMAX1 10.0 to 1.73ge+008
KD1 0.0 to 53069
BMAX2 10.0 to 1.73ge+008
KD2 0.0 to 224318
Goodness of Fit
Degrees of Freedom  10
R squared  0.9382
Absolute Sum of Squares  667.3
Sy.x  8.169
Constraints
BMAX1  BMAX1 > 0.0
KD1  KD1 > 0.0
BMAX2  BMAX2 > 0.0
KD2  KD2 > 0.0

Data
Number of X values  14
Number of Y replicates  1
Total number of values  14
Number of missing values  0
Comparison of Fits

Null hypothesis  One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value Not Necessary
Conclusion (alpha = 0.05) Simpler model fits better
Preferred model  One site binding (hyperbola)

F (DFn, DFd)

One site binding (hyperbola)
Best-fit values
BMAX  94.48
KD 3.722
Std. Error
BMAX  5.159
KD 0.4667
95% Confidence Intervals
BMAX  83.24 to 105.7
KD 2.705 to 4.739
Goodness of Fit
Degrees of Freedom 12
R squared  0.9764
Absolute Sum of Squares 161.8
Sy.x 3.672
Constraints
BMAX  BMAX > 0.0
KD KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX1 142.80
KD1 3.645
BMAX2 251.66
KD2 3.778
Std. Error
BMAX1 11.167e+007
KD1 22066
BMAX2 11.167e+007
KD2 17183
95% Confidence Intervals
BMAX1 0.0 to 2.600e+007
KD1 0.0 to 49167
BMAX2 0.0 to 2.600e+007
KD2 0.0 to 38287
Goodness of Fit
Degrees of Freedom 10
R squared  0.9764
Absolute Sum of Squares 161.8
Sy.x 4.023
Constraints
BMAX1BMAX1 > 0.0
KD1 KD1 > 0.0
BMAX2BMAX2 > 0.0
KD2 KD2 > 0.0
Data
Number of X values 14
Number of Y replicates 1
Total number of values 14
Number of missing values 0
dsGC
Comparison of Fits
Null hypothesis One site binding (hyperbola)
Alternative hypothesis Two site binding (hyperbola)
P value 0.6330
Conclusion (alpha = 0.05) Do not reject null hypothesis
Preferred model One site binding (hyperbola)
F (DFn, DFd) 0.4814 (2,9)

One site binding (hyperbola)
Best-fit values
BMAX 91.26
KD 1.700
Std. Error
BMAX 2.487
KD 0.1288
95% Confidence Intervals
BMAX 85.78 to 96.73
KD 1.417 to 1.984
Goodness of Fit
Degrees of Freedom 11
R squared 0.9919
Absolute Sum of Squares 50.49
Sy.x 2.142
Constraints
BMAX BMAX > 0.0
KD KD > 0.0

Two site binding (hyperbola)
Best-fit values
BMAX183.09
KD1 1.463
BMAX2700525
KD2 1.123e+006
Std. Error
BMAX126.71
KD1 0.5421
BMAX29.062e+009
KD2 1.454e+010
95% Confidence Intervals
BMAX122.67 to 143.5
KD1 0.2370 to 2.690
BMAX20.0 to 2.050e+010
KD2 0.0 to 3.288e+010
Goodness of Fit
Degrees of Freedom 9
R squared 0.9927
Absolute Sum of Squares 45.61
Sy.x 2.251
Constraints
BMAX1BMAX1 > 0.0
KD1 KD1 > 0.0
BMAX2BMAX2 > 0.0
KD2 KD2 > 0.0
Data
Number of X values 13
Number of Y replicates 1
Total number of values 13
Number of missing values 0
Appendix D: Circular dichroism spectra of IFI16-PAAD mutants with increased secondary structure
Appendix E: Circular dichroism spectra of IFI16-PAAD mutants with decreased secondary structure

![Circular dichroism spectra](image)
Appendix F: Thermal denaturation of IFI16-PAAD mutants with increased thermostability
Appendix G: Thermal denaturation of IFI16-PAAD mutants with decreased thermostability
Appendix H: Chemical denaturation of IFI16-PAAD mutants

![Graph showing chemical denaturation of IFI16-PAAD mutants. The x-axis represents [GuHCl] (M), ranging from 0 to 6, and the y-axis represents fraction unfolded, ranging from 0 to 1. The graph includes lines for Wild type, Mutant 4, Mutant 6, Mutant 22, and Mutant 29.]
Appendix I: IFI16-PAAD disorder region prediction

GlobPlot

globplot.embl.de

DISOPRED

bioinf.cs.ucl.ac.uk/disopred/
VL3H - Disprot prediction

ist.temple.edu/disprot/predictor.php

DisEMBL
Appendix J: MNDA-PAAD 2D HSQC titration spectra

No ssDNA added (Biospin processed)

No ssDNA added (NMRDraw processed)
1 molar ratio concentration of ssDNA added

2 molar ratio concentration of ssDNA added
3 molar ratio concentration of ssDNA added

4 molar ratio concentration of ssDNA added
REFERENCE LIST


Hon, B. K. (2009). IFI16 exhibits RPA function suggesting a role in DNA repair in BASC complex. Simon Fraser University, Vancouver.


Motulsky, H., Christopoulos, A. (2003). GrapPad PRISM v4 Fitting Models to Biological Data using Linear and Nonlinear Regression


