

**THE INTERFERON-INDUCIBLE PROTEIN 16 IS A
SINGLE-STRANDED NUCLEIC ACID BINDING PROTEIN
WITH UNWINDING AND ENDONUCLEASE ACTIVITIES**

by

Benjamin Hon
B.Sc., MBB, Simon Fraser University, 2006

THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF SCIENCE

In the Department of
Molecular Biology and Biochemistry

© Benjamin Hon 2010
SIMON FRASER UNIVERSITY
Spring 2010

All rights reserved. However, in accordance with the *Copyright Act of Canada*, this work may be reproduced, without authorization, under the conditions for *Fair Dealing*. Therefore, limited reproduction of this work for the purposes of private study, research, criticism, review and news reporting is likely to be in accordance with the law, particularly if cited appropriately.

APPROVAL

Name: Benjamin Hon
Degree: Master of Science
Title of Thesis: The Interferon-inducible Protein 16 is a Single-stranded Nucleic Acid Binding Protein with Unwinding and Endonuclease Activities

Examining Committee:

Chair: Dr. Lynne Quarmby
Professor, Department of Molecular Biology and Biochemistry

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

Dr. Rosemary B. Cornell
Supervisor
Professor, Department of Molecular Biology and Biochemistry

Dr. David Vocadlo
Supervisor
Assistant Professor, Department of Chemistry

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

Date Defended/Approved: April 12, 2010



SIMON FRASER UNIVERSITY
LIBRARY

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

ABSTRACT

IFI16 belongs to a protein family (HIN200) that contains a characteristic 200-amino-acid protein domain with unknown function. Members of this family were initially described as transcriptional regulators but IFI16 is also involved in other processes such as DNA repair and programmed cell death. To gain further insights into the biochemical mechanisms underlying these functions we made bioinformatics predictions followed by biochemical validation. We predicted and verified experimentally that the 200-amino-acid domain of IFI16 is structurally and functionally similar to replication protein A (RPA), a protein that removes secondary structure on single-stranded DNA during DNA replication and repair. In addition, I found that IFI16 can unwind double-stranded DNA and has an endonuclease activity. This later unexpected discovery reveals a new function for IFI16. Finally, I further discuss the implications of this discovery in regard to the current knowledge about the function of this protein as well as new directions of studies.

Keywords: IFI16; HIN200 family; Transcriptional regulation; DNA repair; Programmed cell death; 200-amino-acid domain; Replication Protein A; DNA unwinding activity; Endonuclease activity

This thesis is dedicated to

my Mom and Dad,

who have always supported me.

ACKNOWLEDGEMENTS

I would like to thank my senior supervisor, Dr. Frederic Pio, for giving me the opportunity to work in his laboratory as an Individual Study Semester student, then a volunteer, and finally a graduate student during the past five years. I have learned a lot from him, and this work would not have been possible without his guidance and support. I am very grateful to my supervisory committee members, Dr. Rosemary Cornell and Dr. Vocadlo, for their input and direction on my research project. I would also like to thank Dr. Lisa Craig for being my internal examiner and Dr. Lynne Quarmby for chairing my thesis defense. Many thanks to all past and present Pio Lab members, particularly Kush Dalal, Karen Yan, and Desmond Lau, for their assistance and friendship. Special thanks to Dr. Scott Briscoe for hiring me as a teaching assistant for his course and I appreciate his support throughout my graduate career.

I would like to express my deepest gratitude to my parents and my fiancée Elsa Wong for their encouragement. I would also like to acknowledge Alex Yuen, Gabriel Alfaro, Lorena Braid, David Shen, Lester Poon, Gurpreet Sekhon, and Dr. Janet Huang for their assistance to my research. Thanks to all graduate students and friends who have made my graduate career a memorable experience, particularly Kerry Chan, Desmond Lau, Alex Yuen, Dixon Ng, and Sampson Wu.

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 Introduction to the HIN200 protein family	1
1.2 Interferon-inducible protein 16 (IFI16)	5
1.2.1 Domain organization of IFI16.....	5
1.2.2 Cell cycle regulation and transcriptional repression by IFI16.....	8
1.2.3 Role of IFI16 in autoimmune disease.....	10
1.2.4 IFI16 is involved in DNA repair and apoptosis	11
1.2.5 The PAAD domain of IFI16 is a novel ssDNA binding protein	13
1.3 Replication protein A (RPA)	13
1.3.1 Structure of RPA.....	14
1.3.2 RPA contains OB folds	15
1.3.3 Function of RPA	16
1.4 Similarities between IFI16 and RPA	17
1.4.1 Structural similarities of IFI16-HIN200A and RPA70AB	17
1.4.2 IFI16 and RPA are involved in similar DNA pathways and have similar binding partners.....	19
1.5 An overview of objectives, rationale of my research, and results summary	20
2: Materials and Methods.....	24
2.1 Gene cloning.....	24
2.1.1 Cloning for gene expression	24
2.1.2 Cloning for yeast two-hybrid	25
2.2 Protein Expression	26
2.2.1 Bacterial expression system	26
2.2.2 Yeast expression system.....	27
2.2.3 Mammalian expression system.....	27
2.3 Protein Purification	28
2.3.1 Purification of His-tagged proteins	28
2.3.2 Ion-exchange chromatography	29
2.3.3 Gel filtration	30
2.4 UV cross-linking of HIN200A-DNA complexes	30

2.5	Chemical cross-linking of protein complexes	31
2.6	Tyrosine fluorescence quenching assay	32
2.7	Melting depression of dsDNA	32
2.8	Western blotting	33
2.9	Fluorescence Resonance Energy Transfer (FRET)	34
2.9.1	DNA compaction and extension	34
2.9.2	DNA polarity assay	36
2.9.3	Site-directed mutagenesis for DNA polarity assay	37
2.10	PLATE transformation	38
2.11	Yeast two-hybrid	39
2.12	Radioactive labelling DNA with ³² P	40
2.12.1	5' end labelling – kinase assay	40
2.12.2	3' end labelling	41
2.13	Determination of labelled oligo concentration	41
2.14	DNA unwinding assay	41
2.15	DNA cleavage assay	44
3:	Results	45
3.1	The first HIN200 domain of IFI16 has similar DNA binding properties as RPA	45
3.1.1	IFI16-HIN200A binds to single-stranded nucleic acid	45
3.1.2	IFI16-HIN200A has higher binding affinity for GC-rich ssDNA than dsDNA	49
3.1.3	IFI16-HIN200A oligomerizes, compacts and extends ssDNA after binding	52
3.1.4	IFI16-HIN200A does not destabilize dsDNA	55
3.1.5	ssDNA-binding polarity of IFI16-HIN200A	56
3.2	IFI16 unwinds DNA	58
3.2.1	IFI16-HIN200A has DNA unwinding activity	58
3.2.2	DNA unwinding activity of IFI16-HIN200A is ATP independent and requires divalent cation Mg ²⁺	64
3.2.3	Temperature dependence of IFI16-HIN200A unwinding activity	65
3.2.4	A ssDNA segment on the probe does not enhance IFI16-HIN200A unwinding effect	66
3.3	IFI16 has endonuclease activity on ssDNA	69
3.3.1	IFI16-HIN200A cleaves ssDNA	69
3.3.2	Non-sequence specific endonuclease activity of IFI16-HIN200A	70
3.3.3	The DNA endonuclease activity of IFI16-HIN200A is Mg ²⁺ and temperature dependent but ATP independent	76
3.4	Studies on full-length IFI16	80
3.4.1	Yeast 2-hybrid indicates no interaction between IFI16 and BLM	80
3.4.2	IFI16 full-length protein can be over-expressed in HEK293T cells	83
4:	Discussion	85
4.1	IFI16-HIN200 is an RPA-like ssDNA binding protein with unwinding and endonuclease activity	85
4.2	DNA unwinding and endonuclease activities: new functions to IFI16	89
4.3	IFI16 in transcriptional repression and apoptosis	92

5: Conclusion and future directions	98
Appendices.....	101
Appendix A: Proteins used in DNA unwinding and DNA cleavage assay.....	101
Appendix B: ProtParam of IFI16-HIN200A	102
Appendix C: Structural Alignment of RPA70AB (PDB: 1JMC) and IFI16-HIN200A (PDB:2OQ0) by EBI-SSM	104
Appendix D: Secondary structure prediction on oligonucleotides using mfold	107
Appendix E: Vector plasmids.....	111
Appendix F: Synthetic Complete (SC) and Drop-out media	115
Reference List	117

LIST OF FIGURES

Figure 1-1. Domain structures of murine and human proteins of the HIN200 family.	3
Figure 1-2. A closer look at the domain structure of IFI16.....	6
Figure 1-3. Schematic diagram of cell cycle regulation by IFI16.	9
Figure 1-4. Domain structure of RPA (Fanning <i>et al.</i> 2006).	14
Figure 1-5. Crystal structures of RPA70AB and IFI16-HIN200A.	19
Figure 2-1. DNA compaction and extension by IFI16-HIN200A (Yan <i>et al.</i> 2008).	35
Figure 2-2. DNA polarity assay.....	37
Figure 2-3. DNA substrates used in unwinding assay.....	43
Figure 3-1. Protein(IFI16-HIN200A)-ssDNA complex was detected after UV cross-linking (Yan <i>et al.</i> 2008).....	47
Figure 3-2. IFI16-HIN200A also binds RNA (Yan <i>et al.</i> 2008).....	49
Figure 3-3. IFI16-HIN200A has a higher binding preference to GC-rich DNA (Yan <i>et al.</i> 2008).	51
Figure 3-4. IFI16-HIN200A forms different oligomers upon DNA binding (Yan <i>et al.</i> 2008).....	53
Figure 3-5. IFI16-HIN200A can compact and extend ssDNA (Yan <i>et al.</i> , 2008).....	54
Figure 3-6. IFI16-HIN200A does not destabilize dsDNA (Yan <i>et al.</i> 2008).....	55
Figure 3-7. Double-stranded DNA destabilization by IFI16-PAAD (Dalal K, M.Sc thesis, 2006 with permission).	56
Figure 3-8. IFI16-HIN200A binds the 18-mer ssDNA in 3'-5' orientation (Yan <i>et al.</i> 2008).....	57
Figure 3-9. IFI16-HIN200A unwinds Φ X174 partial duplex.	59
Figure 3-10. IFI16-PAAD has no DNA unwinding activity.	60
Figure 3-11. HIN200B of IFI16 unwinds Φ X174 partial duplex.....	62
Figure 3-12. DNA unwinding activity was inhibited in IFI16-2HIN, which contains both HIN200A and HIN200B.	63
Figure 3-13. IFI16-HIN200A requires magnesium to unwind DNA.....	65
Figure 3-14. DNA unwinding of IFI16-HIN200A is temperature sensitive.....	66
Figure 3-15. IFI16-HIN200A unwinds bubble, 5'T and 3'T like 30-mer.....	68
Figure 3-16. IFI16-HIN200A cleaves DNA.....	70

Figure 3-17. Endonuclease activity of IFI16-HIN200A.	72
Figure 3-18. Cleavage assay of IFI16-HIN200A on 5'end and 3'end-labelled 5'T oligonucleotide.	73
Figure 3-19. Cleavage assay of IFI16-HIN200A on 5'end and 3'end-labelled T ₇₀	74
Figure 3-20. Cleavage on GC-5, GC-3 by IFI16-HIN200A showed different pattern of cleaved product.	75
Figure 3-21. Endonuclease activity of IFI16-HIN200A is magnesium dependent.	77
Figure 3-22. Temperature effect on IFI16-HIN200A cleavage activity.	79
Figure 3-23. Yeast 2-hybrid showed no direct interaction between IFI16 and BLM.	82
Figure 3-24. Full-length IFI16 expression by different expression systems.	84
Figure 4-1. Transcriptional regulation of c-Myc (Brooks <i>et al.</i> 2009).	95

LIST OF TABLES

Table 2-1. Sequences of primers used for cloning.....	26
Table 2-2. Oligonucleotides used in UV cross-linking and tyrosine fluorescence quenching	31
Table 2-3. Antibodies used in western blots	34
Table 2-4. Oligonucleotides used in FRET assays	36
Table 2-5. Mutagenic primers for mutant IFI16-HIN200A and their sequences.....	38
Table 2-6. Oligonucleotides used in DNA unwinding assay	43
Table 2-7. Oligonucleotides used in DNA cleavage assay	44
Table 3-1. Experimental molecular weights of IFI16-HIN200A, DNA, and protein-nucleic acid complex from SDS-PAGE (Yan <i>et al.</i> 2008).....	48
Table 3-2. Dissociation constants (K_d) and B_{max} values for IFI16-HIN200A in tyrosine fluorescence quenching (Yan <i>et al.</i> 2008).....	52
Table 4-1. Summary of DNA-interacting properties of IFI16, MNDA, and RPA	89

GLOSSARY

AD	<u>A</u> ctivation <u>D</u> omain
AIM2	<u>A</u> bsent in <u>M</u> elanoma protein <u>2</u>
ASC	<u>A</u> poptosis associated <u>S</u> peck like protein with <u>C</u> aspase Recruitment Domain
ATM	<u>A</u> taxia <u>T</u> elangiectasia <u>M</u> utated protein kinase
BASC	<u>B</u> RCA1 <u>A</u> ssociated Genome <u>S</u> urveillance <u>C</u> omplex
BLAST	<u>B</u> asic <u>L</u> ocal <u>A</u> lignment <u>S</u> earch <u>T</u> ool
BLOTTO	<u>B</u> ovine <u>L</u> acto <u>T</u> ransfer <u>T</u> echnique <u>O</u> ptimizer
BLM	<u>B</u> loom Syndrome protein
BRCA1	<u>B</u> reast <u>C</u> ancer <u>A</u> ssociated Protein <u>1</u>
BSA	<u>B</u> ovine <u>S</u> erum <u>A</u> lbumin
CARD	<u>C</u> aspase <u>R</u> ecruitment <u>D</u> omain
CDK2	<u>C</u> yclin- <u>d</u> ependent <u>K</u> inase <u>2</u>
DMEM	<u>D</u> ulbecco's <u>M</u> odified <u>E</u> agle <u>M</u> edium
Da	<u>D</u> alton
DBD	<u>D</u> NA- <u>b</u> inding <u>D</u> omain
dsDNA	<u>D</u> ouble- <u>s</u> tranded <u>D</u> eoxyribonucleic <u>A</u> cid
DTT	<u>D</u> ithiothreitol
EBI	<u>E</u> uropean <u>B</u> ioinformatics <u>I</u> nstitute
EMSA	<u>E</u> lectrophoretic <u>M</u> obility <u>S</u> hift <u>A</u> ssay

FBS	<u>F</u> etal <u>B</u> ovine <u>S</u> erum
FRET	<u>F</u> örster <u>R</u> esonance <u>E</u> nergy <u>T</u> ransfer
HEK293	<u>H</u> uman <u>E</u> mbryonic <u>K</u> idney <u>293</u> cells
HIN200	<u>H</u> aematopoietic <u>I</u> nterferon-inducible <u>N</u> uclear protein with <u>200</u> -amino-acid motif
HMW	<u>H</u> igh <u>M</u> olecular <u>W</u> eight
IFI16	<u>I</u> nter <u>f</u> eron- <u>i</u> nducible Protein <u>16</u>
IFIX	<u>I</u> nter <u>f</u> eron- <u>i</u> nducible Protein <u>X</u>
IFN	Interferon
IPTG	<u>I</u> sopropyl- β - <u>t</u> hiogalactopyranoside
IR	<u>I</u> onizing <u>R</u> adiation
LMW	<u>L</u> ow <u>M</u> olecular <u>W</u> eight
MEF	<u>M</u> ouse <u>E</u> mbryonic <u>F</u> ibroblast
MNase	<u>M</u> icrococcal <u>N</u> uclease
MNDA	<u>M</u> yeloid <u>N</u> uclear <u>D</u> ifferentiation <u>A</u> ntigen
MW	<u>M</u> olecular <u>W</u> eight
NHEIII ₁	<u>N</u> uclease <u>H</u> ypersensitive <u>E</u> lement
NLS	<u>N</u> uclear <u>L</u> ocalization <u>S</u> ignal
Ni-NTA	<u>N</u> ickel – <u>N</u> itrilo <u>t</u> riacetic acid
OB fold	<u>O</u> ligonucleotide/Oligosaccharide <u>B</u> inding fold
p53	Tumor protein <u>53</u>
p21	Tumor protein <u>21</u>
PAAD	<u>P</u> yrin, <u>A</u> IM, <u>A</u> SC, <u>D</u> eath domain-like domain

PAGE	<u>P</u> oly <u>a</u> cryl <u>a</u> mid <u>e</u> <u>G</u> el <u>E</u> lectrophoresis
PDB	<u>P</u> ro <u>t</u> ein <u>D</u> ata <u>B</u> ank
PLATE	<u>P</u> EG, <u>L</u> ithium <u>A</u> cetate, <u>T</u> ris, and <u>E</u> DTA
PMSF	<u>P</u> henyl <u>m</u> ethane <u>s</u> ulphonyl <u>f</u> luoride
RPA	<u>R</u> eplication <u>P</u> rotein <u>A</u>
Rb	<u>R</u> etinoblastoma protein
RMSD	<u>R</u> oot <u>M</u> ean <u>S</u> quare <u>D</u> eviation
SC Medium	<u>S</u> ynthetic <u>C</u> omplete Medium
SCOP	<u>S</u> tructural <u>C</u> lassification <u>O</u> f <u>P</u> roteins
SLE	<u>S</u> ystemic <u>L</u> upus <u>E</u> rythematosis
SNase	<u>S</u> taphylococcal <u>N</u> uclease
SSB	<u>S</u> ingle-stranded DNA <u>b</u> inding protein
SSc	<u>S</u> ystemic <u>S</u> clerosis
ssDNA	<u>S</u> ingle-stranded <u>D</u> eoxyribonucleic <u>A</u> cid
SV40	<u>S</u> imian <u>V</u> irus 40
TCA	<u>T</u> richloroacetic <u>a</u> cid
UV	<u>U</u> ltraviolet
Y2H	<u>Y</u> east <u>2</u> -hybrid

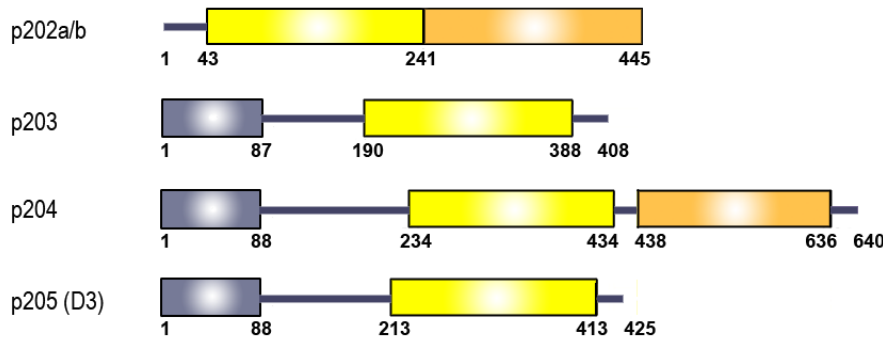
1: INTRODUCTION

1.1 Introduction to the HIN200 protein family

The HIN200 family refers to a group of genes that contain either a single or a tandem repeat encoding a characteristic domain of 200 amino acids that is also called HIN200. The name of this family came from several features that were originally discovered among its members: first, these genes were found to be expressed in haematopoietic cells. Second, they were transcriptionally activated following interferon treatment. Interferons (IFN) are cytokines that induce cells to resist viral replication and regulate cell growth. Third, all gene expressions were localized in the nucleus. Lastly, the characteristic HIN200 motif is 200 amino acids long (hence HIN200). In recent years however, a number of studies have demonstrated that the characteristics of these proteins had broader functions. For example, it was found that these proteins actually have a wider cellular distribution and subcellular localization than initially thought. Effectively, gene expressions of some members were also found in non-haematopoietic tissues such as skin epithelial cells, gastrointestinal tract, urogenital tract and glands, and ducts of breast tissue (Gariglio *et al.* 2002). They can be regulated by a wide range of exogenous stimuli other than IFN, and their localizations are not restricted to the nucleus (Dawson *et al.* 1995; Flati *et al.* 2001). Despite these new findings, the original family name has been maintained for this group of proteins.

The chromosomal localization of all HIN200 family members is very well conserved. Currently, there are five murine members and four human members in the HIN200 family, and they are all located at the q21-23 region of chromosome 1 in both mouse and human genome (Briggs *et al.* 1994; Trapani *et al.* 1992; Landolfo *et al.* 1998). Murine HIN200 proteins include p202a, p202b, p203, p204, and p205 that is also known as D3. The human counterparts are IFI16 (Interferon-inducible protein 16), MNDA (Myeloid Nuclear Differentiation Antigen), AIM2 (Absent in Melanoma 2) and IFIX (Interferon-inducible protein X).

Mouse



Human

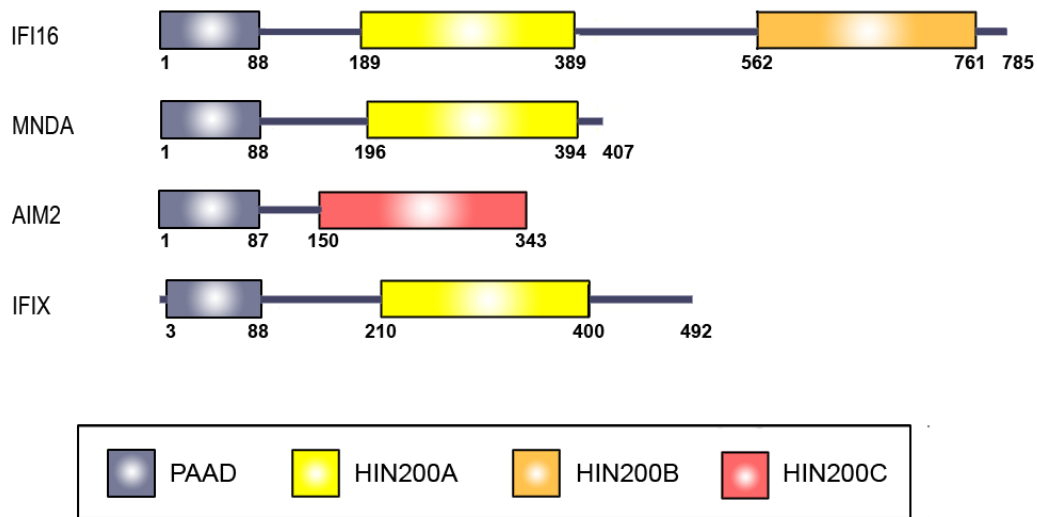


Figure 1-1. Domain structures of murine and human proteins of the HIN200 family.

There are five murine members and four human members identified thus far for the HIN200 family. p202a and p202b are two nearly identical proteins with difference in only 8 amino acids, and are the only HIN200 members that lack an N-terminal PAAD domain in their sequences. All members have at least one HIN200 domain of either A, B, or C subclasses. The subclasses were distinguished based on their amino-acid sequence conservation.

As shown in Figure 1-1, members in the HIN200 family not only contain at least one HIN200 domain, but most of them also contain a PAAD domain at their N-termini except for p202a and p202b. PAAD was named after the protein families from which it was discovered, Pyrin, AIM, ASC [apoptosis-associated

speck-like protein containing a caspase recruitment domain (CARD)], and Death-domain-like. The PAAD domain represent the fourth branch of the death domain superfamily and is commonly found in proteins that are involved in inflammatory and apoptotic signalling pathways (Fairbrother *et al.* 2001; Pawłowski *et al.* 2001). In apoptosis, apoptotic proteins associate together via PAAD-PAAD interactions and consequently activate caspases, which drive the programmed cell death. This has only been demonstrated in other apoptotic proteins but not yet in any HIN200 family members. The HIN200 domains, which can also mediate protein-protein interactions (more detail in section 1.2.1), are divided into three subclasses based on their amino-acid sequence conservation: HIN200A, HIN200B, and HIN200C. The sequence identity of the three HIN200 domain subclasses ranges from 24.4% to 45.8%; it has been suggested that the low sequence identity could impact on the secondary structure of the domains (Ludlow *et al.* 2005).

The murine p202a and p202b are the only HIN200 members that do not possess a PAAD domain. One functional study attempt has been made on this murine member in which knockout of p202a gene was achieved. However, knockout mice displayed no distinct phenotype because it was found that p202b was up-regulated and compensated for the loss of p202a (Wang *et al.* 1999). Besides having a role in interferon biology, HIN200 proteins are commonly implicated in cell proliferation, differentiation, and cell cycle regulation. The murine p202a and p204 are involved in the G₁-S transition in cell cycle, and

resulted in arrest in G1 phase (Yan *et al.* 1999; Gariglio *et al.* 1998; Flati *et al.* 2001).

For the human members, IFI16 was first discovered from a cytotoxic T cell cDNA expression library (Trapani *et al.* 1992). Other human family members (MND1, AIM2, and IFI1) were discovered either by probing cells with specific antibodies, or by querying (BLAST) the human genome with homologous mouse genes (Goldberger *et al.* 1984; DeYoung *et al.* 1997; Ding *et al.* 2004).

This thesis focuses on the human member IFI16, which is a very interesting member as it has been implicated in many aspects of cellular processes such as cell cycle regulation, transcriptional repression, apoptosis, tumor suppression, and DNA repair.

1.2 Interferon-inducible protein 16 (IFI16)

1.2.1 Domain organization of IFI16

Three different IFI16 protein isoforms exist from alternative mRNA splicing: IFI16a, IFI16b, and IFI16c, in the order of increasing length of their amino-acid sequence. The sites of alternative splicing in the mRNA are located at the region corresponding to the spacer between the two HIN200 domains. In the longest isoform IFI16c, the spacer region contains three copies of approximately 56 amino-acid repeat that are rich in serine, threonine, and proline (S/T/P). Shorter isoforms IFI16b and IFI16a have one and two of these repeats spliced out respectively, as indicated in Figure 1-2. IFI16b is the most abundant

among the three isotypic variants according to a quantitative western blot analysis (Johnstone *et al.* 1998). The function of this region is unknown, but S/T/P-rich regions are often unstructured and it may therefore provide a flexible hinge for the protein (Ludlow *et al.* 2005).

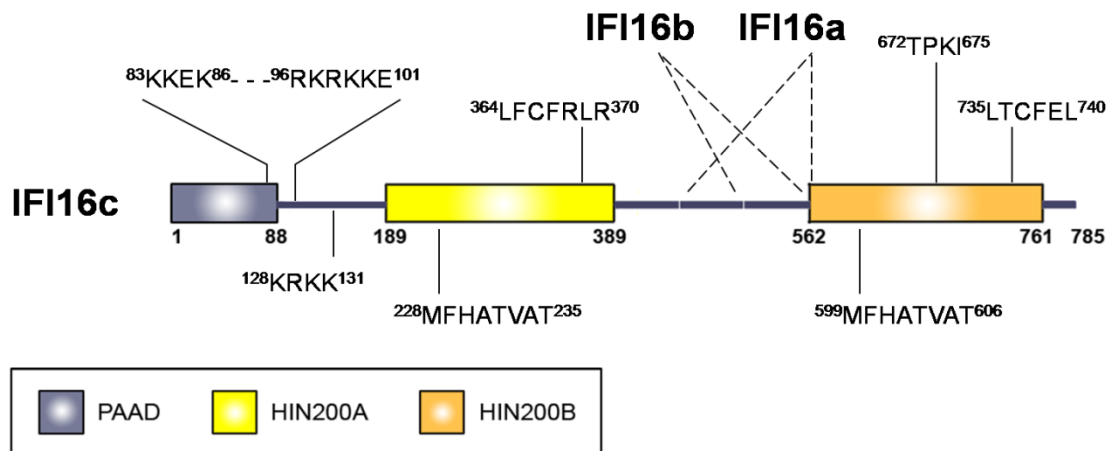


Figure 1-2. A closer look at the domain structure of IFI16.

IFI16 has three alternatively spliced forms. IFI16c is the longest isoform. Dotted lines showed the spliced out regions for IFI16a and b. Also shown in this figure are the positions of nuclear localization signals near the end of PAAD domain, MFHATVAT motif in the beginning of HIN200A and HIN200B, Rb-binding LXCXE motif, and CDK2 phosphorylation site TPKI. The structure of HIN200A domain is solved and is shown on Figure 1-5.

Although some IFI16 isoforms are shorter in length, the main features remain. They contain two nuclear localization signals (NLS) that are found within the region between the PAAD and HIN200A domains. IFI16 contains a mono-partite NLS at the 128 amino-acid position and a bipartite NLS between amino-acids 83 and 101 (Figure 1-2). Mono-partite NLS consists of a short sequence of positively charged residues, usually lysines (K) or arginines (R). Bi-partite NLS is two clusters of positively charged amino acids separated by a mutation-tolerant

region of 10-12 amino acids. Together, recognition of these signals mediates the nuclear import of IFI16.

IFI16 consists of two HIN200 protein domains (HIN200A and HIN200B). Near the beginning of each HIN200 domains is the MFHATVAT (M) motif that is responsible for the self-association of IFI16 (Koul *et al.* 1998). Some studies also suggested that the M motif is necessary for direct interaction with p53 (Datta *et al.* 1996). Another motif that is found in both HIN200 domains is the LFCF(R/H) motif in HIN200A or LXCXE motif in HIN200B, both of which are involved in binding retinoblastoma protein (pRb or Rb) and are referred as Rb binding (L) motifs (Xin *et al.* 2003). Finally, the consensus amino-acid sequence TPKI, which served as a putative cyclin-dependent kinase 2 (CDK2) phosphorylation site (Landolfo *et al.* 1998), is found only in HIN200B. The characterization of the CDK2 phosphorylation site in one domain but not the other indicates that there may be functional differences between HIN200A and HIN200B-containing proteins.

Perhaps the most important feature within both HIN200 domains is the existence of two oligonucleotide/oligosaccharide binding (OB) fold (Albrecht *et al.* 2005). This structural motif comprises a five-stranded β -sheet of approximately 100 amino acids that wraps to form a β -barrel, with an α -helix bridging the 3rd and 4th β -strands of the barrel (Theobald *et al.* 2003). This structure has been described as a closed or partly opened β -barrel with Greek-key motif in the SCOP (Structural Classification Of Proteins) database (Murzin *et al.* 1995). As

hinted by its name, this fold allows binding to different nucleic acid structures. It can also bind to oligosaccharides, proteins, and metal ions (Arcus 2002).

1.2.2 Cell cycle regulation and transcriptional repression by IFI16

IFI16 is involved in cell cycle regulation and it binds to cell cycle regulatory proteins p53, p21, Rb and E2F and regulates their transcriptional activities (Johnstone *et al.* 2000; Xin *et al.* 2003). During cell cycle progression, transcription factor E2F is required to form a complex with another protein called DP1 to promote transcription of their target genes, which are necessary for the G₁ to S transition (Figure 1-3). Rb can repress the E2F/DP1 complex, but not when it is in its hyperphosphorylated state (ppRb). This switch is controlled by cyclin/CDK2 which, when active, drives the phosphorylation of Rb and results in E2F/DP1 complex activation. Cells respond to stimuli such as DNA damage by increasing expression of the CDK2 inhibitor p21 via a p53-dependent pathway, resulting in accumulation of hypophosphorylated Rb and a pause in cell cycle progression. IFI16 has a positive effect on p21 transcription when it interacts with p53, supported by one study where knock-down of IFI16 in human fibroblasts suppressed p53-mediated transcription of p21 (Xin *et al.* 2004). Another study also suggested that IFI16 up-regulates p21 expression, down-regulates cyclin expression, and stimulates cellular senescence (Zhang *et al.* 2007). On the other hand, IFI16 can also bind to hypophosphorylated Rb and E2F and this may promote inactivation of the E2F/DP1 complex. In any case, the consequence is repression of E2F/DP1 target genes, which explains why IFI16 is a cell cycle regulator (Figure 1-3).

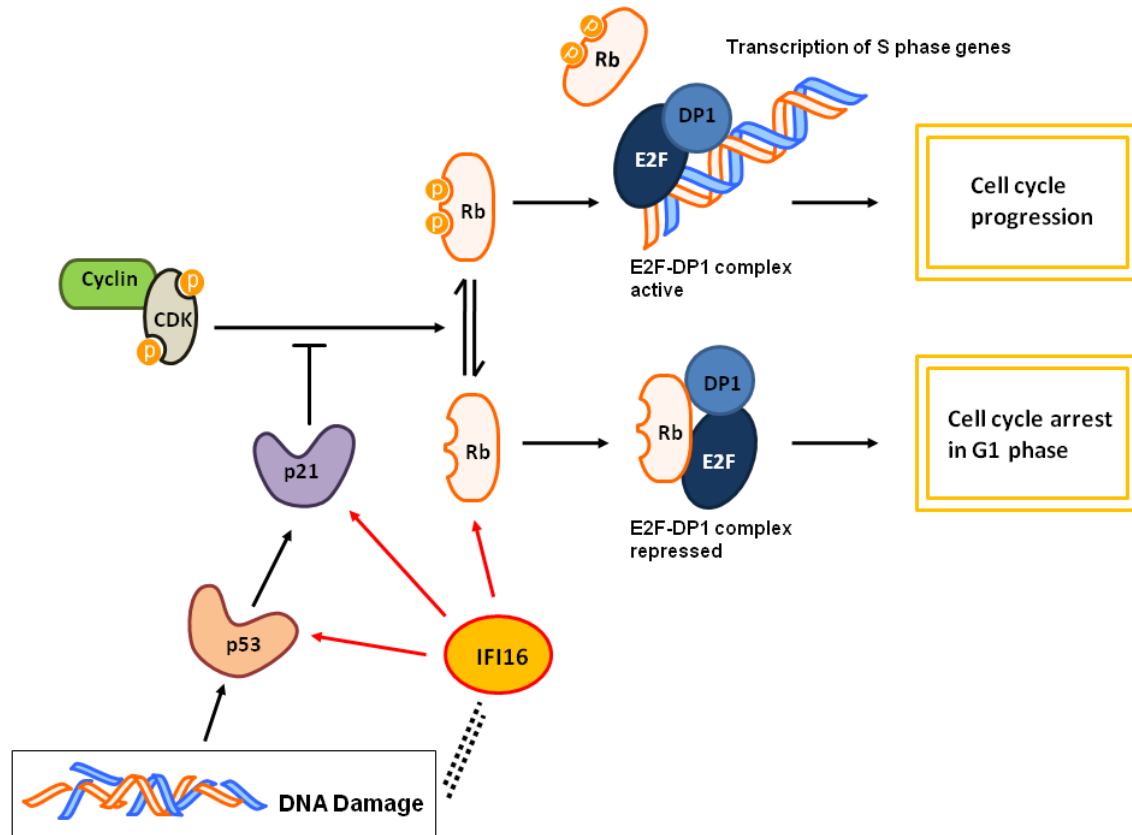


Figure 1-3. Schematic diagram of cell cycle regulation by IFI16.

Cell cycle progression requires active E2F-DP1 complex for transcription of S phase genes. This complex is repressed by hypo-phosphorylated Retinoblastoma protein (Rb). Phosphorylation of Rb protein is controlled by Cyclin/CDK. Interacting with these proteins by other proteins can affect cell fate. p21 and p53 that are over-expressed in response to DNA damage can inhibit phosphorylation of Rb protein and causes cell cycle arrest in G1 phase. IFI16 also halts cell cycle progression, which could be due to its direct interaction with p53, p21, and Rb proteins. This also indicates a possible link between IFI16 and DNA damage signalling.

Besides having a role in cell cycle regulation, IFI16 is also involved in transcriptional regulation. Figure 1-3 illustrates an example of how IFI16 regulates transcription by binding to transcription factors. In another example, IFI16 fused to the GAL4 DNA binding domain can function as a transcriptional repressor by blocking expression of a gene positioned up to 1 kb away from the GAL4 binding site (Johnstone *et al.* 1998). Furthermore, IFI16 forms a stable

complex with a transcriptional activator SP1 at DNA element IR1 and represses activation of CMV DNA polymerase expression by SP1 (Luu *et al.* 1997).

IFI16 is also a transcription factor by itself. Its nuclear localization and its ability to bind DNA are typical features of transcriptional factors. IFI16 is able to bind to guanine-rich sequence within promoters of p53 and c-Myc (a gene that promotes cell growth), and negatively affects their expression level (Egisteili *et al.* 2009). The fact that IFI16 is responsible for these cellular reactions also implies that defects in IFI16 may lead to genome instability and cause various diseases including cancer (Zhang *et al.* 2007). For this reason, IFI16 is also known as a tumor suppressor.

1.2.3 Role of IFI16 in autoimmune disease

Autoimmune diseases such as systemic sclerosis (SSc) and systemic lupus erythematosus (SLE) are believed to be due to a loss of tolerance to nuclear antigens causing anti-nuclear antibody production followed by immune deregulation. The most significant observation in SSc and SLE patients is the expression of a spectrum of IFN stimulated genes (Mondini *et al.* 2007).

Consistent with IFN inducibility of IFI16, increased levels of IFI16 protein in human SLE patients are associated with the diseases (Choubey *et al.* 2008).

One study used immunohistochemical analysis to evaluate the expression of IFI16 in skin biopsy specimens obtained from patients with SSc and SLE (Mondini *et al.* 2006). It was found that expression of IFI16 was greatly increased in all layers of the epidermis of lesional skin from patients of both diseases and this contributes to higher anti-IFI16 IgG antibody levels compared

with normal controls. In another study, high titer anti-IFI16 antibodies were detected in the sera of 28.7% (107 of 374) SLE patients but not in healthy controls (Seelig *et al.* 1994). These studies provide evidence that IFI16 may be involved in the pathophysiologic mechanisms of SSc and SLE, but further experiments are required to establish whether altered expression of IFI16 is necessary or sufficient for initiation and pathogenesis of these autoimmune diseases.

1.2.4 IFI16 is involved in DNA repair and apoptosis

When human cells suffer DNA damage from environmental factors such as ultraviolet (UV) light or ionizing radiation (IR), one DNA repair mechanism is to form a multi-protein complex called BRCA1 Associated Surveillance Protein Complex (BASC) that has the ability to sense and repair damaged DNA. The BASC complex contains numerous proteins involved in different DNA transactions such as recombination and DNA repair. This protein complex is composed of the following proteins: BRCA1, ATM/ATR kinases, MLH1, MSH2, MSH6, Rad50, MRE11, NBS, and DNA helicase BLM, as shown by a mass spectrometry study (Wang *et al.* 2000), and together as a whole, BASC identifies and corrects damaged DNA molecules. BRCA1 (Breast Cancer Associated Gene 1) is a large, multifunctional nuclear phosphoprotein of 1863 amino acids forming the central core of the BASC complex. It was found that a breast cancer cell line HCC1937, which expresses a truncated form of BRCA1, failed to confer resistance to DNA damage-induced cell death but reintroduction of wildtype BRCA1 recovered it (Scully *et al.* 1999). Interestingly, the N-terminus of IFI16

that contains the PAAD domain has been shown to physically interact with the 501-801 amino-acid region of BRCA1 (Aglipay *et al.* 2003). The same study also demonstrated that IFI16 failed to relocalize within the nucleoplasm in response to IR treatment when BRCA1 is mutated. Reintroduction of wildtype BRCA1 restored the nuclear relocalization of IFI16, suggesting that DNA repairing activity of BRCA1 may require IFI16 as an intermediate component.

The direct interaction between IFI16 and BRCA1 through the PAAD domain is also important for p53-mediated apoptosis. Co-expression of BRCA1 and IFI16 in BRCA1(-)/p53(+) mouse embryonic fibroblasts (MEF) led to severe retardation of cell growth and loss of cell viability 36-48 days after IR treatment, whereas co-expression of BRCA1 and IFI16 lacking PAAD domain in the same MEFs resulted in resistance to growth retardation (Aglipay *et al.* 2003). Expression of IFI16 alone only slightly retarded cell growth and no apoptosis was detected, confirming that apoptosis under conditions of DNA damage requires cooperation between IFI16 and BRCA1.

The findings that IFI16 is a functional component of BASC strongly support a role in DNA repair pathway and apoptosis. Furthermore, each BASC member has the ability to recognize different abnormal DNA structures that could possibly be the result of DNA damage. This evidence further supports a DNA-damage sensing role in IFI16 suggested in Figure 1-3.

1.2.5 The PAAD domain of IFI16 is a novel ssDNA binding protein

As mentioned previously, PAAD domains mediate the association of apoptotic proteins. The PAAD domain of IFI16 directly interacts with BRCA1 (Section 1.2.4). In fact, the PAAD domain was originally thought to be exclusively a protein-protein interaction domain (Liepinsh *et al.* 2003; Hiller *et al.* 2003; Liu *et al.* 2003; Pawłowski *et al.* 2001). Recently, the novel ssDNA binding properties of IFI16-PAAD were established (Dalal K, M.Sc thesis, 2006). It was found that IFI16-PAAD can bind to single-stranded nucleic acids, including RNA, and prefers guanine-rich DNA. Also, IFI16-PAAD destabilizes dsDNA structure and its presence effectively lowered the melting temperature of dsDNA by approximately 10 °C. On the other hand, the stability of IFI16-PAAD was greatly improved when bound to single-stranded DNA compared to double-stranded DNA, indicating a stable protein-ssDNA complex (Dalal K, M.Sc thesis, 2006). This nucleic acid binding property would imply a broader range of functions for the Death domain super-family in apoptosis and inflammation. More specifically, a PAAD-DNA binding domain could imply functions in DNA repair/replication/recombination or transcriptional regulation for the full length protein.

1.3 Replication protein A (RPA)

Based on the above information, it is evident that IFI16 can bind to ssDNA via its PAAD domain. As a transcriptional repressor, it can bind to dsDNA structures. The region of dsDNA binding has been mapped to the HIN200

domain (Johnstone *et al.* 1998), which is not surprising because it contains two OB folds. In fact, we will show in this study that the HIN200A domain of IFI16 has similar function to Replication Protein A (RPA), a human single-stranded DNA binding protein (Yan *et al.* 2008; Albrecht *et al.* 2005). Before the crystal structure of IFI16-HIN200A was available, our laboratory built a model for this protein domain by comparative modelling using the 3D jury method. As a result, we identified the OB-fold-containing region of RPA (PDB: 1JMC) as the best template based on structural alignment (Pio, 2003, unpublished data). In humans, RPA is commonly involved in DNA replication and recombination processes. This section covers an overview of RPA structure and function.

1.3.1 Structure of RPA

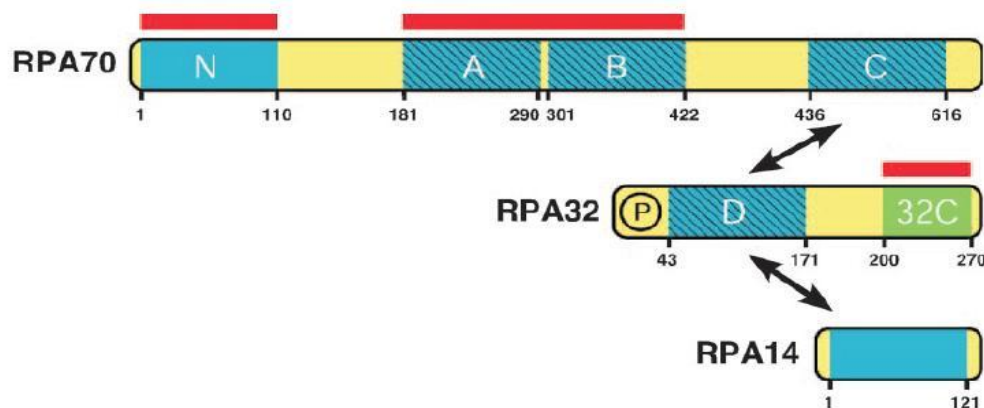


Figure 1-4. Domain structure of RPA (Fanning *et al.* 2006).

RPA is a heterotrimer composed of RPA70, RPA32, and RPA14. Protein-binding domains are denoted by red bars; ssDNA binding domains A, B, C, and D are denoted by hatching. Green box indicates the region of winged-helix-turn-helix motif. Each OB fold is denoted by a blue box. Arrows indicate intersubunit associations. The structure of RPA70AB (aa 181-422 of RPA70) is solved and is shown on Figure 1-5.

RPA is the most abundant single-stranded DNA binding protein (SSB) in humans (Seroussi *et al.* 1993). It is a heterotrimeric protein that consists of three subunits: RPA70, RPA32 and RPA14, designated according to their molecular weights (Wold *et al.* 1988). The domain organization of these subunits is shown in Figure 1-4. RPA contains four ssDNA binding domains, three of which are found in RPA70 (A, B, and C) and one in RPA32 (D). RPA70 therefore plays a major role in recognizing ssDNA (Bochkareva *et al.* 1998; Pfuetzner *et al.* 1997; Bochkareva *et al.* 2001). Besides having ssDNA-binding domains, RPA also contains protein-binding domains. The N-terminus of RPA70 (RPA70N) mediates interactions with a variety of replication proteins and transcription factors (Jacobs *et al.* 1999). Furthermore, ssDNA binding domains RPA70A and B are also capable of protein binding, so is the winged-helix-turn-helix fold at the C-terminus of RPA32. RPA14 associates with RPA70C and RPA32D to form the trimerization core, which links the three RPA subunits together and keeps RPA in its heterotrimeric form. Detailed structural information of RPA is available as most of the RPA domains have been successfully crystallized.

1.3.2 RPA contains OB folds

The OB fold, consisting of a β -barrel with a Greek-key topology, is one of the most common structures in any single-stranded nucleic acid binding protein. In human RPA, there are six of these OB folds and they play an important role in DNA transactions (Kerr *et al.* 2003; de Laat *et al.* 1998; Henricksen *et al.* 1994; Wold 1997). OB folds of RPA bind strongly to ssDNA with preference to guanine-rich sequence and their binding affinity to ssDNA shows a strong

dependence on the length of the oligonucleotides (Bochkareva *et al.* 2001). They can also recognize dsDNA and RNA but binding is weaker by 100-1000 fold (Kim *et al.* 1994). RPA displays different binding modes for interactions with DNA, and this is triggered by the oligomerization of its OB folds (Bochkareva *et al.* 2002). Furthermore, RPA OB folds contribute to the destabilization of dsDNA during the initiation of DNA replication (Treuner *et al.* 1996). Remodelling of OB-fold was previously shown by swapping the OB fold in RPA from archaea (Robbins *et al.* 2005). They showed that the OB folds when interchanged could still perform their functions, one of which is to wrap and stretch ssDNA. In addition to human RPA, most OB-fold proteins contain a prevalent 3'-5' ssDNA binding polarity: the 5' end is oriented towards the C-terminus of the protein and the 3' end extends towards the N-terminus (Theobald *et al.* 2003).

1.3.3 Function of RPA

In its stable complex form, RPA is known to be important for DNA processing pathways such as DNA replication, DNA recombination, DNA repair, transcription, translation, cold shock response and telomere maintenance although the role of RPA in DNA replication has been more extensively studied (Wold 1997). RPA can bind tightly to ssDNA and protects it from nucleolytic damage, prevents hairpin formation, and blocks DNA re-annealing until the processing pathway step is successfully completed (Stenlund 2003). OB folds play an important role during these processes because they can adopt different DNA-binding modes that can be switched from one to the other through self association, protein interactions or phosphorylation. Beside ssDNA binding, RPA

also has DNA unwinding activity (Georgaki *et al.* 1992). It was found that RPA can unwind up to at least 350 bp of DNA, and that the unwinding activity is independent of $MgCl_2$ and ATP. Thus, RPA could initiate DNA transactions by providing a proper substrate for DNA helicase, polymerase or primase for further DNA processing. Indeed, it has been shown that RPA can physically interact with DNA helicase Bloom Syndrome protein (BLM) and stimulate its energy-driven DNA unwinding activity (Brosh *et al.* 2000). RPA has been suggested to bind damaged DNA and function in early stage of DNA repair (Hashimoto *et al.* 2000; Coverley *et al.* 1992). RPA was also shown to bind to negative regulatory sequences found upstream of a number of DNA repair and DNA metabolism genes (Singh *et al.* 1995). Therefore, it is capable of regulating certain gene transcriptions as well.

1.4 Similarities between IFI16 and RPA

1.4.1 Structural similarities of IFI16-HIN200A and RPA70AB

The major common feature between RPA and IFI16 is that they both contain OB folds. As mentioned, the OB-fold region of RPA (RPA70AB) and the OB-fold region of IFI16 (IFI16-HIN200A) are also structurally similar according to comparative modeling (Section 1.3). The similarity is more apparent upon comparison of the crystal structures of IFI16-HIN200A (PDB: 2OQ0) and RPA70AB (PDB: 1JMC). The crystal structure of IFI16-HIN200A has a resolution of 2.0 Å. RPA70AB was crystallized with an octadeoxycytosine oligonucleotide

and diffracted to 2.4 Å (Bochkarev *et al.* 1997) (Figure 1-5). From the surface representations of both structures, IFI16-HIN200A appears to possess a DNA binding groove similar to that of RPA70AB. The specific lysine residues in RPA70, namely K183, K259, K263, and K343 (highlighted in blue in Figure 1-5) were identified using mass spectrometric protein footprinting to determine the ssDNA contacts (Shell *et al.* 2005). In the structure of IFI16-HIN200A, lysine side chains that may interact with ssDNA are exposed within the putative DNA binding groove and this may provide evidence of the groove being the site of DNA binding. According to the protein structure comparison service SSM (secondary-structure matching) at European Bioinformatics Institute (<http://www.ebi.ac.uk/msd-srv/ssm>), authored by E. Krissinel and K. Henrick (Krissinel *et al.* 2004), the OB fold structures of IFI16-HIN200A and RPA70AB were superimposed with a root mean square deviation (RMSD) of 3.99 Å (Appendix C). It is possible that a reorientation of the IFI16-HIN200A OB folds may occur to compensate for the weak structural conservation upon binding to ssDNA.

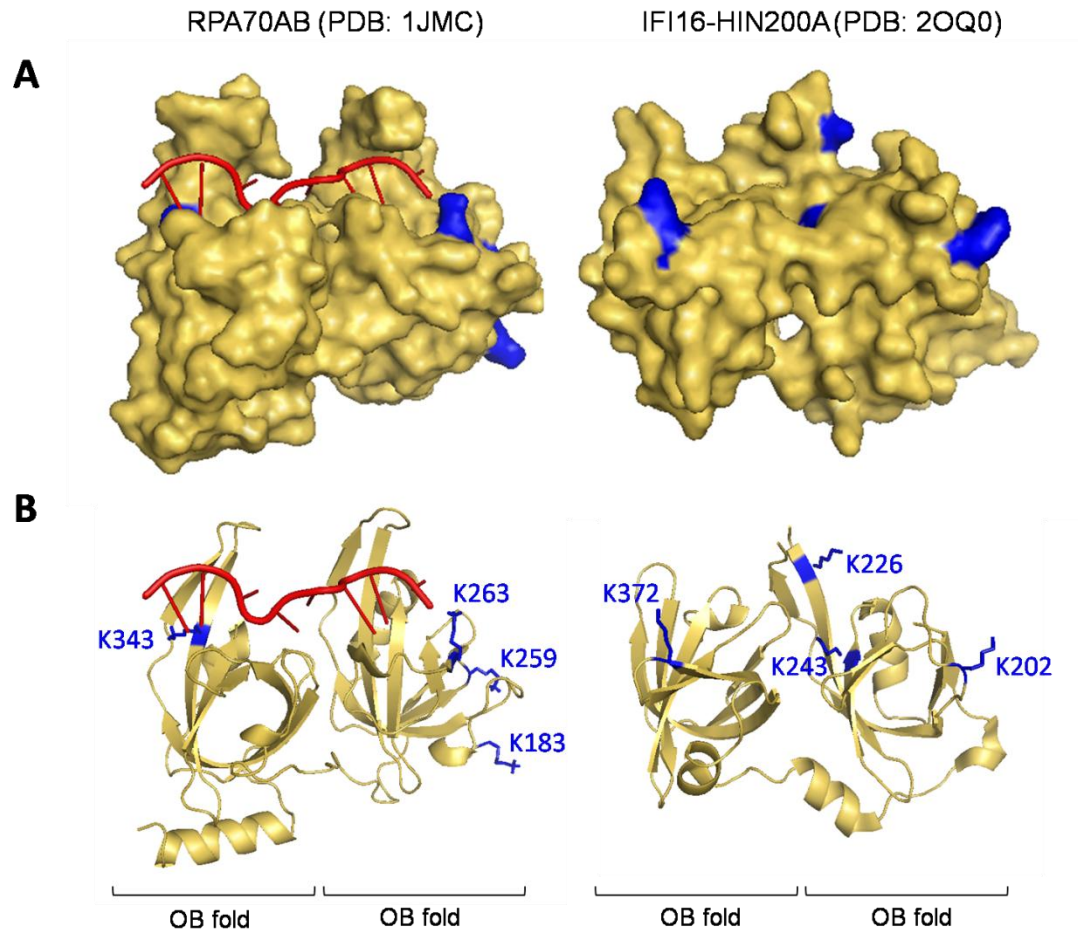


Figure 1-5. Crystal structures of RPA70AB and IFI16-HIN200A.

Crystal structure of RPA70AB (left) in complex with poly(dC)₈ (red) and IFI16-HIN200A (right), each containing two OB folds as indicated. **A)** Surface representation of the structures with lysine side chains highlighted in blue. A potential DNA binding pocket was found in IFI16-HIN200A similar to the one on RPA70AB. **B)** Left: Ribbon representation of RPA70AB showing lysines K343, K263, K259, and K183 that were known to directly interact DNA. Right: lysine residues that may interact with ssDNA were also found around the potential DNA binding pocket in IFI16-HIN200A. Lysine residues were shown as stick representation.

1.4.2 IFI16 and RPA are involved in similar DNA pathways and have similar binding partners

IFI16 and RPA not only have structurally similar OB folds, but they also take part in similar processes such as DNA damage-induced cell cycle regulation, transcriptional repression, and DNA repair. Furthermore, RPA has

also been implicated in autoimmune disease like IFI16. A study showed that antibodies directed against RPA70 and RPA32 were found in SLE patients (Garcia-Lozano *et al.* 1996). Although not much is known about RPA in autoimmune disease, this demonstrated once again that IFI16 and RPA are involved in similar pathways of DNA metabolism.

RPA and IFI16 also share common interacting partners. As a member of BASC complex, IFI16 is functionally associated with its components. RPA can also physically interact with many of these individual BASC components, including p53, BLM, RFC and RAD50-MRE11-NBS1 (Fanning *et al.* 2006). RPA directly interacts with PCNA, a potential BASC component, according to a combination of affinity chromatography and mass spectrometric analyses (Ohta *et al.* 2002; Wang *et al.* 2000). RPA also directly interacts with the BLM, which is a DNA helicase, and enhances its unwinding activity (Brosh *et al.* 2000). BLM belongs to the RecQ helicase family and it has been shown to be an early responder to DNA double-strand breaks (Karmakar *et al.* 2006). It could therefore be an important part of the BASC for its unwinding function, which is critical in DNA transactions.

1.5 An overview of objectives, rationale of my research, and results summary

Although IFI16 has been implicated in cell cycle regulation, transcriptional repression, and DNA repairing pathways, its biochemical mechanism in these pathways has been elusive. Since these are processes that involve interactions

with DNA, understanding OB fold property of IFI16 is crucial for elucidating IFI16 function in these mechanisms. In a recent study where a group of researchers tried to determine the DNA-sensing component in an inflammasome that triggers innate immunity, they specifically searched for proteins in the NCBI database that contain OB fold motifs and a death domain, which is commonly found in inflammatory proteins. They found the cytosolic AIM2, another HIN200 family member, met these two criteria. They discovered that the OB folds of AIM2 were able to sense potentially dangerous cytoplasmic DNA from viral infection. AIM2 can also form a multi-protein complex called the inflammasome with other inflammatory proteins to activate caspase-1, which cleaves pro-IL-1 β to its active form and induce inflammation in 293T-caspase-1-ASC cell line. AIM2 was confirmed to be a cytoplasmic DNA sensor that triggers activation of the innate immune system (Fernandes-Alnemri *et al.* 2009). This finding is consistent with other studies published in Nature during the same period (Bürckstümmer *et al.* 2009; Hornung *et al.* 2009).

It is possible that the HIN200 family of proteins may constitute a new type of DNA sensor proteins involved in DNA repair and inflammation. Like the AIM2 study above, studying how the OB folds of IFI16 recognize nucleic acids will be key to understanding its function, although AIM2 is a cytosolic protein and IFI16 is a nuclear protein. We have evidence from comparative modelling that the OB-fold containing HIN200A domain of IFI16 is structurally similar to the nuclear protein RPA. Since RPA and IFI16 have been implicated in similar pathways and diseases and have common interacting partners (p53, BLM, RFC and RAD50-

MRE11-NBS1), they may also be functionally similar. However, the low sequence similarity between RPA and IFI16-HIN200A (less than 10%) make it difficult to infer a function with confidence to this domain since the OB-fold superfamily members are highly diverged, sharing little sequence conservation and perform a wide variety of functions (see Structural Classification of Proteins, <http://scop.mrc-lmb.cam.ac.uk/scop>) (Watson *et al.* 2005). For this reason, a proposed RPA function of IFI16-HIN200A must be further verified experimentally.

My study was initiated to improve the understanding of the function of IFI16 and how it is involved in the cellular pathways, using RPA as the basis. Karen Yan and collaborators (Yan *et al.* 2008) initially hypothesized that the OB-fold containing HIN200A of IFI16 had similar functional properties as the OB-fold of RPA. We tested six RPA nucleic acid-binding characteristics on IFI16-HIN200A, and our results indicated that IFI16-HIN200A possesses most of the ssDNA-binding and OB fold properties of RPA. Specifically, IFI16-HIN200A binds to GC-rich ssDNA with higher affinity than to dsDNA, oligomerizes upon ssDNA binding, wraps and stretches ssDNA, and recognizes ssDNA in the same orientation as RPA but does not destabilizes dsDNA (Section 3.1). Next, I demonstrated that IFI16-HIN200A unwinds dsDNA like RPA (Section 3.2). To extend the investigation of RPA function in IFI16, I looked at the interaction of IFI16 and BLM, which is known to interact with RPA and increase its unwinding of dsDNA by RPA unwinding activity (Brosh *et al.* 2000). This interaction between IFI16 and BLM was not detected (Section 3.4.1). Interestingly, I found that IFI16 has DNA cleavage activity, a function that is not found in RPA (Section

3.3). These findings may provide new insights on the role of IFI16 in DNA damage/repair signalling pathways, transcriptional repression and apoptosis.

2: MATERIALS AND METHODS

2.1 Gene cloning

All clones were propagated in *Escherichia coli* DH5 α , plasmids were purified with Qiagen miniprep kit, and nucleotide sequences were verified by MacroGen sequencing service in Korea before use.

2.1.1 Cloning for gene expression

Genes encoding IFI16 domains were previously generated as described (Dalal K, M.Sc Thesis, 2006; Yan H, M.Sc Thesis, 2007): HIN200A (aa 194-393), HIN200B (aa 557-767), 2HIN (aa 194-767) were PCR-amplified from the full-length IFI16c cDNA template (Uniprot accession number: Q16666) and cloned into pET100/D-TOPO vector (Invitrogen, CA). The expression plasmid for the PAAD domain of IFI16 was generated by inserting the Human DNA fragments encoding for PAAD protein domains of IFI16 referred to as IFI16-1 (3–88, CAI15085) into a pET28b vectors (Novagen).

IFI16 gene used for full-length gene expression was purchased from PlasmID (ID#:HsCD00043079) and was cloned into yeast expression vector pYES2 (Appendix E) and mammalian expression vector pCEP4 (Appendix E). Primer pairs EXP_IFI16FL_F and pYES2_IFI16FL_R were used to amplified the IFI16 full-length gene for pYES2 and it was cloned downstream of the GAL1 gene promoter between *KpnI* and *XhoI* sites of pYES2. For cloning IFI16 full-

length gene into pCEP4, primer pairs EXP_IFI16FL_F and pCEP4_IFI16FL_R were used with restriction enzymes *KpnI* and *BamHI* (see Table 2-1 for primer sequences).

2.1.2 Cloning for yeast two-hybrid

MATCHMAKER Two-Hybrid system by Clontech was used, in which IFI16 (PlasmID), BRCA1 (Addgene, Plasmid 12341), and BLM (plasmid pJK1 gifted by Ian D. Hickson) full-length genes were cloned into bait vector pAS2-1 and prey vector pACT2 (see section 2.11 for details in yeast two-hybrid method, Appendix E for plasmid maps). To clone full-length IFI16 into both pAS2-1 and pACT2, primers Y2H_IFI16FL_F and Y2H_IFI16FL_R were used to amplify the gene and restriction enzymes *SfiI* and *BamHI* were used to clone it into both vectors. Y2H_BRCA1FL_F and pAS2-1_BRCA1FL_R (*SfiI* and *SaII*) were used in cloning BRCA1 full-length gene into pAS2-1 and primers Y2H_BRCA1FL_F and pACT2_BRCA1FL_R (*SfiI* and *XhoI*) for pACT2. For cloning BLM, Y2H_BLMFL_F and pAS2-1_BLMFL_R (*SfiI* and *PstI*) were used for pAS2-1; Y2H_BLMFL_F and pACT2_BLMFL_R (*SfiI* and *XhoI*) were used for pACT2 (see Table 2-1 for primer sequences).

Table 2-1. Sequences of primers used for cloning

Name of primers	Nucleotide sequence (restriction sites underlined)
Y2H_IFI16FL_F	5'-GACATGGG <u>CCCATGGAGGCC</u> ATGGGAAAAAATACAAGAACATTG-3'
Y2H_IFI16FL_R	5'-CGC <u>GGATCC</u> GAAGAAAAAGTCTGGTGAAGTTTC-3'
Y2H_BRCA1FL_F	5'-GACATGGG <u>CCCATGGAGGCC</u> ATGGATTTATCTGCTCTTCGC-3'
pAS2-1_BRCA1FL_R	5'-ACGCGT <u>CGAC</u> GTAGTGGCTGTGGGGGAT-3'
pACT2_BRCA1FL_R	5'-GCA <u>ACTCGAGG</u> TAGTGGCTGTGGGGGAT-3'
Y2H_BLMFL_F	5'-GACATGGG <u>CCCATGGAGGCC</u> ATGGCTGCTGTTCTCAA-3'
pAS2-1_BLMFL_R	5'-GCA <u>ACTG</u> CAGTGAGAATGCATATGAAGGCTT-3'
pACT2_BLMFL_R	5'-GCA <u>ACTCGAGT</u> GAGAATGCATATGAAGGCTT-3'
EXP_IFI16FL_F	5'-CGGGG <u>TACCAT</u> GGGAAAAAATACAAGAACATTG-3'
pYES2_IFI16FL_R	5'-GCA <u>ACTCGAGT</u> CATCAGTGATGATGATGATGATGGAAGAAAAAGTCT GGTGAAGTTTC-3'
pCEP4_IFI16FL_R	5'-CGC <u>GGATCC</u> TCATCACATGTGATGATGATGATGATGGAAGAAAAAGT CTGGTGAAGTTTC-3'

2.2 Protein Expression

2.2.1 Bacterial expression system

Each protein domain was expressed as a His tag-fusion protein in *Escherichia coli* BL21(DE3) transformed with its respective construct. Cell cultures were grown to 0.5 of O.D₆₀₀ reading before 1 mM of IPTG was added to induce the protein expression. Cell cultures were then kept shaking at 37 °C for additional 4 hours. Cells were pelleted down by centrifuging at 5000 rpm with F10 rotor (Beckman) at 4 °C for 20 minutes. The pellets were resuspended with lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 1 mM imidazole, 500 µM PMSF, pH 8.0) and cell lysis was done by sonication at approximately 20 Volts. After sonication, cell lysate was subjected to centrifugation with a JA17 rotor (Beckman) at 31000 x g for 1 hour at 4 °C to remove cell debris and the

supernatant was filtered using 0.45 µm syringe filter before His-tag purification took place.

2.2.2 Yeast expression system

Yeast strain S288c (genotype: *MATα SUC2 gal2 mal mel flo1 flo8-1 hap1 ho bio1 bio6*, provided by Gabriel Alfaro of Beh Lab) was used for full-length IFI16 gene expression. Cells transformed with pYES2_IFI16 by PLATE transformation method (Section 2.10) were selected by -uracil SC medium (Appendix F) and were grown overnight at 30 °C with shaking. To induce expression of IFI16 from the GAL1 promoter, overnight culture was transferred from glucose-containing medium to galactose-containing medium and adjusted to OD₆₀₀ 0.5. Without glucose, the GAL1 promoter is not repressed and transcription of IFI16 is induced by continuous shaking at 30 °C for additional 8 hours. After induction, cells were centrifuged at 1500 x g for 15 minutes at 4 °C. Finally, cells were lysed by sonication in lysis buffer (50 mM NaH₂PO₄, 300 mM NaCl, 1 mM imidazole, 500 µM PMSF, pH 8.0) and IFI16 expression was detected by western blot (Section 2.8).

2.2.3 Mammalian expression system

Both Cos-7 and HEK293T cell lines were used for full-length IFI16 expression trials. They were maintained in T75 tissue culture flasks with 15 ml of Dulbecco's Modified Eagle Medium (DMEM) supplemented with 5% (v/v) Fetal Bovine Serum (FBS) at 37 °C under 5% CO₂. Once they were grown to 80-90% confluence, cells were detached from the flasks by incubation with 0.25% (1X)

trypsin for 2 minutes at 37 °C, diluted 10-fold with the addition of fresh media, and transferred to new plates. Cells were passaged every 3 days until ready to be transfected.

IFI16 full-length gene was cloned into mammalian expression vector pCEP4 (Invitrogen, Section 2.1.1). Plasmid DNA transfection was carried out with Lipofectamine 2000 (Invitrogen according to product protocol. Briefly, 2×10^5 mammalian cells in 2 ml of growth medium (DMEM, 5% FBS) were plated into 6-well plates one day before transfection. At the time of transfection, cells were 90-95% confluent. DNA plasmids were complexed with lipofectamine 2000 as follows: For each well, 10 μ l of lipofectamine 2000 was diluted in 250 μ l DMEM without serum, and 4 μ g of DNA plasmid in 250 μ l of DMEM without serum was added to lipofectamine and was mixed gently. The mixture was allowed to incubate for 20 minutes at room temperature. The 500 μ l DNA/lipofectamine mixture was then added to each well containing cells and medium. Plates were gently mixed and were incubated at 37 °C for 2-3 days prior to testing for expression of IFI16.

2.3 Protein Purification

2.3.1 Purification of His-tagged proteins

Proteins were purified according to the QIA-expressionist protocol (Qiagen) using gravity spin columns containing Ni-NTA agarose. Targeted proteins from cell lysates that contain a his-tag were trapped into the beads while

the supernatant (see section 2.2.1) was passed through the columns. Columns were then washed with lysis buffer containing 20 mM imidazole to remove protein contaminants that came from non-specific binding to the column. Targeted proteins were eluted in the lysis buffer containing 250 mM imidazole and were then dialysed promptly in 50 mM Tris, 100 mM NaCl, pH 8.0. A final concentration of approximately 15 mM of β -mercaptoethanol was added just before overnight dialysis to avoid protein aggregation. After dialysis, all purified proteins were concentrated to 5-10 mg/ml using an Amicon Ultra centrifugation filter with a 3000 dalton molecular weight cut-off and then stored at -80 °C.

2.3.2 Ion-exchange chromatography

Ion-exchange purification was performed with Mono S HR 5/5 column (GE Healthcare) attached to FPLC purification system ÄKTApurifier (GE Healthcare). IFI16-HIN200A recombinant protein has a theoretical pI value of 9.07 according to ExPASy ProtParam tool (Appendix B). Mono S column was first equilibrated with five column volumes of 50 mM HEPES, pH 7.5, which was attached to pump A of FPLC purifier. At pump B, it attached the buffer of 50 mM HEPES, 1M NaCl, pH 7.5. During purification, IFI16-HIN200A was injected into the sample loop and was pushed into the Mono S column by 100% pump A buffer at a flow rate of 0.5 ml/min. Elution of protein was done in the following gradient setting: 0-100% pump B buffer in 100 minutes at 0.5 ml/min flow rate, 1 ml fraction size. IFI16-HIN200A was eluted out when the gradient at pump B was approximately 25% (~250 mM NaCl). Eluted fractions were immediately dialyzed overnight in 50 mM Tris, 100 mM NaCl, 15 mM β -mercaptoethanol, pH 8.0 buffer.

2.3.3 Gel filtration

Gel filtration column Superose 12 HR 10/30 (GE Healthcare) was used for IFI16-2HIN gel filtration with FPLC purification system ÄKTApurifier (GE Healthcare). First, the column was equilibrated with five column volumes of 50 mM Tris, 100 mM NaCl, 15 mM β -mercaptoethanol, pH 8.0 buffer. Next, 1 ml of concentrated IFI16-2HIN (6-8 mg/ml) was injected into 1-ml sample loop and was slowly pushed into the column at a flow rate of 0.5 ml/min. Fraction collection began after passage of the void volume (~ 7 ml); each fraction was 0.5 ml.

2.4 UV cross-linking of HIN200A-DNA complexes

IFI16-HIN200A (3.7 μ M) was incubated with ssDNAs A₂₅, T₂₅, GC- 5, and GC-3 (Table 2-2) as well as dsDNA (A-T)₂₅ (A₂₅ and T₂₅ annealed together), and dsDNA (GC)₅₋₃ (GC-5 and GC-3 annealed together) at 1:1 DNA:protein molar ratio in 20 mM sodium cacodylate, pH 7.25. Each reaction was placed on the lid of a 1.5 ml Eppendorf tube and exposed to 254 nm UV light at a distance of 20 cm on a UV Stratalinker 2400 (Stratagene, CA) at 1000 J/min for 15 min. Cross-linked products were analysed on 15% SDS-PAGE and visualized by silver staining or western blotting with anti-his tag antibody (Section 2.8) to detect the his-tagged IFI16-HIN200A. For each gel, a calibration curve was plotted from the migration distance (R_f) of each low range molecular weight standard (Bio Rad), which was used to estimate the molecular weights of the protein monomer and protein-DNA complexes.

Table 2-2. Oligonucleotides used in UV cross-linking and tyrosine fluorescence quenching

Name of oligos	Nucleotide sequence
T ₂₅	Poly(dT) ₂₅
A ₂₅	Poly(dA) ₂₅
GC-5	5'-GGAAGAAGGAAGTGGGATCAGGATCCGCTGGCTCC-3'
GC-3	5'-GGAGCCAGCGGATCCTGATCCCACTTCCTTCTTCC-3'

2.5 Chemical cross-linking of protein complexes

IFI16-HIN200A (1.8 μ M) was incubated alone or with T₂₅ at a DNA:protein molar ratio of 0.1:1 for 10 minutes at 20 °C in 400 μ l of 20 mM HEPES, pH 7.5. 16 μ l of the chemical cross-linker, glutaraldehyde (25% initial concentration), was added and incubated for 2 minutes at 20°C. The two reactions were then stopped by adding 20 μ l freshly-made quenching solution (2 M NaBH₄ and 0.1 M NaOH) for 20 minutes at the same temperature. Cross-linked species were precipitated by adding 18 μ l trichloroacetic acid (TCA, 78% initial concentration) and incubating on ice for 5 minutes. The precipitant was pelleted by 10-minute centrifugation at 13,000 rpm at 4 °C and washed once with 600 μ l cold acetone followed by another centrifugation. Air-dried pellet was resuspended in 20 μ l SDS-PAGE loading buffer and analysed on a 15% SDS-PAGE, together with IFI16-HIN200A without T₂₅ and glutaraldehyde as the negative control. The gel was visualized by silver staining.

2.6 Tyrosine fluorescence quenching assay

Wildtype IFI16-HIN200A has six tyrosine residues that absorb light and their emission may be quenched by oligonucleotide binding. To obtain the binding constants for different oligonucleotides A₂₅, T₂₅, GC-5, GC-3, dsDNA (A-T)₂₅ and (GC)₅₋₃ (Table 2-2), IFI16-HIN200A (3.7 μ M) was titrated by each of them in 20 mM sodium cacodylate, pH 7.25. The DNA: protein molar ratio was increased from 0.01 to 0.1 at an increment of 0.01 and from 0.1 to 3.0 at an increment of 0.1. At each titration point, the sample was excited at 275 nm (λ_{ex} of tyrosine) and the emission intensity at 304 nm (λ_{em} of tyrosine) was recorded and plotted against the corresponding oligonucleotide concentration. Spectra of oligonucleotides alone were recorded under the same conditions as the negative control.

The data were analyzed by curve fitting to a one-site ($Y = B_{\text{max}}X/(K_d+X)$) or two-site ($Y=B_{\text{max1}}X/(K_{d1}+X) + B_{\text{max2}}X/(K_{d2}+X)$) binding model using Graphpad Prism 4.03 (Graphpad Software Inc.). The parameters B_{max} (maximal binding) and K_d (dissociation constant) were calculated from a non-linear regression fit. The statistical F test in Graphpad was used to determine the preferential binding models. Scatchard plots were analyzed in the Graphpad software.

2.7 Melting depression of dsDNA

Double-stranded DNA (A-T)₂₅ was diluted to a final concentration of 10 μ M in 100 μ l 20 mM sodium cacodylate, pH 7.25, containing 100 mM NaCl. Its

absorbance at 260 nm (A_{260}) was measured by the “Thermal” program on a Cary 300 Bio UV-Visible spectrophotometer equipped with a temperature regulator from 20°C to 90°C in the absence and presence of IFI16-HIN200 at 1:1 protein:DNA molar ratio. Melting depression of dsDNA for IFI16-PAAD was measured in the same way, except that the reaction was carried out in 20 mM HEPES, pH 7.5, at a protein:DNA molar ratio of 2.5:1.

2.8 Western blotting

Transfer from SDS-PAGE gel to Hybond-ECL nitrocellulose membrane (Amersham Biosciences) was done in 1X transfer buffer (48 mM Tris base, 0.04% SDS, 39 mM glycine, and 20% methanol) for 2 hours at 80 V in 4 °C. Membrane was incubated overnight with blocking solution (5% Bovine Lacto Transfer Technique Optimizer [BLOTTO], 0.1% Tween20). The next day, membrane was washed two times with 1X TBS (50 mM Tris-HCl, 150 mM NaCl, 0.1% Tween20, pH 7.5) for 2 minutes before 1-hour primary antibody incubation (refer to Table 2-3 for the amount and the choices of antibody used in different experiments). Then, the membrane was washed three times with 5% BLOTTO, 0.1% Tween20, 15 minutes each, followed by two 5-minute rinses with 1X TBS, 0.1% Tween20. The same washing procedure was performed after the membrane was incubated with HRP-conjugated secondary antibody at room temperature for 1 hour. Finally, immunodetection on X-ray film was carried out using ECL Plus Western Blotting kit (GE Healthcare).

Table 2-3. Antibodies used in western blots

Experiment	Primary antibody		Secondary antibody	
	Name ([])	Dilution	Name ([])	Dilution
UV cross-linking	Mouse anti-his (0.2 mg/ml)	1:20000	Goat anti-mouse- HRP conjugated (0.4 mg/ml)	1:4000
Full-length IFI16 expression	Rabbit anti- IFI16_PAAD (5 mg/ml)	1:100,000	Goat ant-rabbit- HRP conjugated (0.4 mg/ml)	1:4000

2.9 Fluorescence Resonance Energy Transfer (FRET)

2.9.1 DNA compaction and extension

FRET was used to measure the distance between two chromophores attached at both termini of a dsDNA with a 3' overhang and to assess if increasing the amount of protein had an effect on DNA compaction and extension as was shown for RPA (Robbins *et al.* 2005). In this system, a FRET 18-58 dsDNA was used. This nucleotide was obtained by annealing an 18-mer and 58-mer ssDNA oligonucleotide (Table 2-4) to yield a partial duplex consisting of an 18 bp dsDNA with a single-stranded poly(dT)₄₀ tail at the 3' end. The 3' end of 58-mer was labelled with the donor Cy3 analog Quasar 570 (λ_{ex} =548 nm, λ_{em} =567 nm), while the 5' end of 18-mer was labelled with the acceptor Cy5 analog Quasar 670 (λ_{ex} =648 nm, λ_{em} =667 nm) (Figure 2-1, A). Equimolar concentrations of these two partially complementary oligonucleotides were annealed to form FRET 18-58 in annealing buffer (20 mM Tris, 400 mM NaCl, pH 8.0) by heating up to 95 °C and cooling down to 25 °C at a rate of 1.67 °C/min. The energy transfer between donor and acceptor upon binding to IFI16-HIN200A was measured by exciting the FRET 18-58 (300 nM in 20 mM sodium

cacodylate, pH 7.25) at 548 nm and recording the emission intensity at 568 nm and 667 nm from 1:1 to 20:1 protein:DNA molar ratio. Spectra of ssDNA 18-mer and 58-mer under the same conditions were measured as the negative control. If the poly(dT)₄₀ tail is being wrapped around the protein, the distance between the chromophores becomes shorter when more protein is added. If the oligonucleotide is stretched by the protein, the chromophores become farther away when more IFI16-HIN200A molecules bind to the DNA (Figure 2-1, B and C).

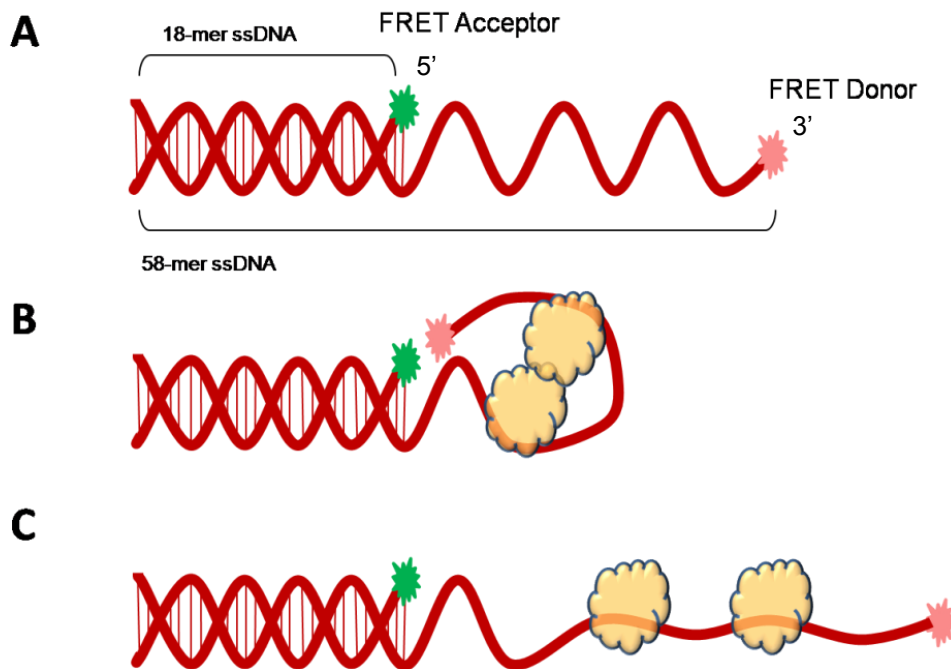


Figure 2-1. DNA compaction and extension by IFI16-HIN200A (Yan *et al.* 2008).

18-mer ssDNA, 5'-labelled with the acceptor Quasar670, is annealed to 58-mer that is 3'-labelled with donor Quasar570. The resulting DNA duplex has a poly(dT)₄₀ tail, which may be wrapped and stretched by IFI16-HIN200A. **B)** During wrapping, the FRET distance between donor and acceptor should become closer and the FRET increases. **C)** If stretching occurs, the distance between chromophores increases and the efficiency of fluorescence transfer is diminished.

Table 2-4. Oligonucleotides used in FRET assays

Name of Oligos	Nucleotide sequence
18-mer	5'- QUASAR670-GCCTCGCTGCCGTCGCCA -3'
58-mer	5'- TGGCGACGGCAGCGAGGC-(T) ₄₀ -QUASAR570-3'

2.9.2 DNA polarity assay

DNA polarity assay was used to determine the orientation of ssDNA upon binding to IFI16-HIN200A. Orientation could be either 5'->3' (5' end directs to the N-terminus of the protein and 3' end extends towards the C-terminus) or 3'->5' (vice versa). To do this, FRET was measured between FRET acceptor Quasar 670 at 5' end of an 18-mer ssDNA and FRET donor DyLight549 maleimide (Pierce) that was labelled on cysteine residue of proteins (Figure 2-2). IFI16-HIN200A has three cysteine residues and they are all located in the C-terminal OB-fold. The FRET measured corresponds to the distance between 5' end of ssDNA and C-terminus of protein. This was eventually compared to FRET between the same piece of ssDNA and a mutant IFI16-HIN200A, which had C-terminal cysteines mutated to serines and had a new cysteine introduced at the N-terminus (Section 2.9.3). The FRET experiment was performed by titrating the labelled protein (0.767 μ M) with the 18-mer ssDNA from 1:1 to 12:1 DNA:protein molar ratio and the emission spectra were recorded from wavelength 540 nm to 700 nm. Energy transfer between the protein and 18-mer were approximated using equations $E=I_A/(I_A+I_{DA})$ and $E=1-(I_{DA}/I_D)$. The difference in FRET could suggest the orientation of ssDNA binding of IFI16-HIN200A (see Figure 2-2 for detail).

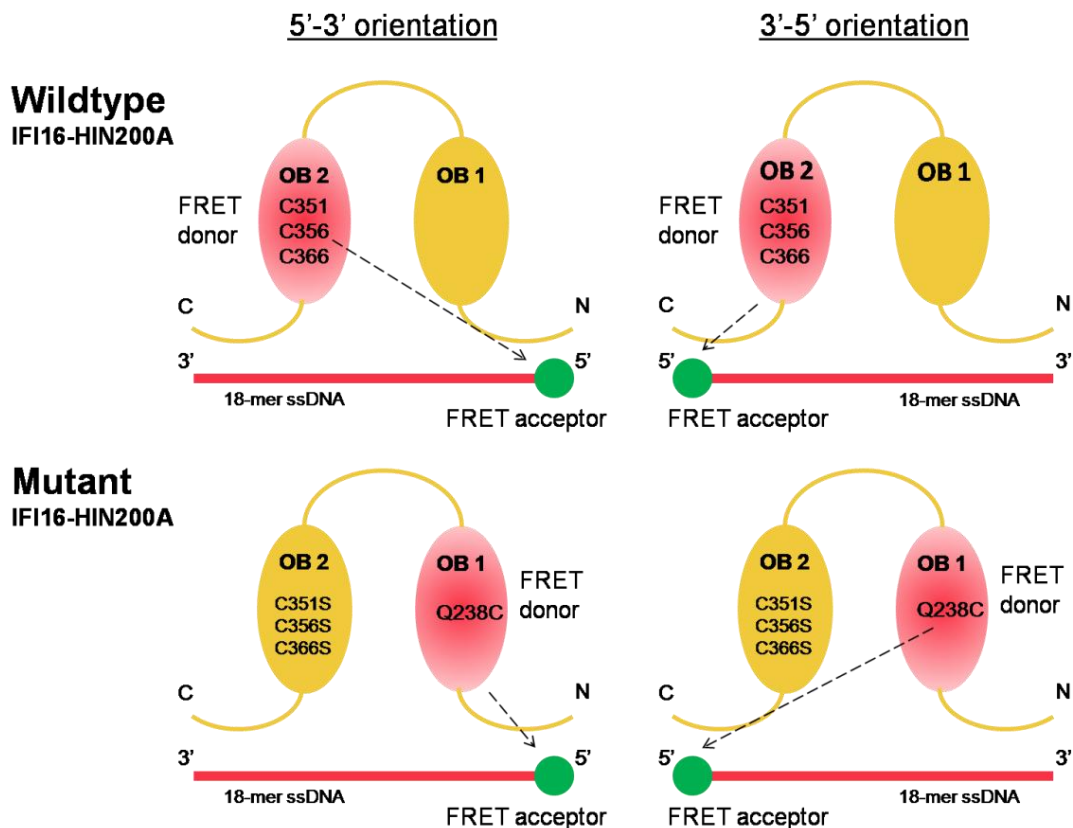


Figure 2-2. DNA polarity assay.

The orientation of ssDNA when bound to IFI16-HIN200A was determined by comparing FRET between wildtype IFI16-HIN200A and ssDNA to that between mutant IFI16-HIN200A and ssDNA. Wildtype IFI16-HIN200A had cysteines at its C-terminus while mutant IFI16-HIN200A contained an N-terminal cysteine, where FRET donor (Dylight549 maleimide) was located. If energy transfer was greater in mutant than wildtype, it indicated the distance between N-terminus of protein and 5'-end of ssDNA is shorter, and the orientation of ssDNA would be 5'->3'. If energy transfer was greater in wildtype than mutant, orientation of ssDNA would be 3'->5'.

2.9.3 Site-directed mutagenesis for DNA polarity assay

Mutant IFI16-HIN200A plasmid was constructed using QuickChange® II Site-Directed Mutagenesis Kit (Stratagene, USA). Three rounds of mutageneses were required. The first round involved two mutations to change cysteines to serines at the C-terminus: C351S and C356S. In the second round, C366S mutation was done on the product from first round, yielding a triple mutant, IFI16-HIN200A^{C351/356/366S}. Finally in the third round, a glutamine in the N-terminus was

mutated to a cysteine, and the construct IFI16-HIN200A^{Q238C+C351/356/366S} was generated. The corresponding forward and reverse primer sequences are summarized in Table 2-5. In each mutagenesis, mutant strands were PCR-amplified and the parental stands were digested by *DpnI* (Stratagene, USA) at 37 °C for 1.5 hours. The mutant plasmid was transformed into *Escherichia coli* DH5α by electroporation. The mutant sequences were verified by sequencing (Macrogen, Korea). All mutants were expressed in *Escherichia coli* BL21(DE3) as described (Section 2.2.1).

Table 2-5. Mutagenic primers for mutant IFI16-HIN200A and their sequences

Name of primers	Nucleotide sequence
IFI16-HIN_C351/356S_F	5'-GGACAGGACAATCTCACAATATCCCCTCTGAAGAAGGAG-3'
IFI16-HIN_C351/356S_R	5'-CTCCTTCTTCAGAGGGGATATTGTGAGATTGTCCTGTCC-3'
IFI16-HIN_C366S_F	5'-GATAAGCTCCAACCTTTTCTCCTTTCGACTTAGAAAAAAGAACC-3'
IFI16-HIN_C366S_R	5'-GGTTCTTTTTTCTAAGTCGAAAGGAGAAAAGTTGGAGCTTATC-3'
IFI16-HIN_Q238C_F	5'-CATGCTACAGTGGCTACACAGACATGCTTCTTCCATGTGAAGG-3'
IFI16-HIN_Q238C_R	5'-CCTTCACATGGAAGAAGCATGTCTGTGTAGCCACTGTAGCATG-3'

2.10 PLATE transformation

In a microcentrifuge tube, 10 µl of 10 mg/ml sonicated salmon sperm DNA (Stratagene, #201190) was first boiled at 100 °C for 5 minutes and immediately chilled on ice for at least 2 minutes to ensure the DNA had become single-stranded. 1.5 µg of each DNA plasmid to be transformed was then added to the salmon sperm DNA and mixed well, followed by addition of 500 µl of a standard yeast transformation solution PLATE (40% PEG3350 (w/v), 100 mM lithium

acetate, 10 mM Tris, pH 7.5, 0.4 mM EDTA). With an autoclaved toothpick, yeast cells that were grown on SC medium or a specific drop-out medium agar plate (Appendix F) were transferred to the PLATE mixture and mixed well. Mixture was allowed to incubate at 30 °C for 15 minutes and then was heat-shocked at 42 °C for another 15 minutes. After that, it was chilled in ice for 2 minutes, followed by 30 °C incubation for 1 - 2 overnight. Cells were then collected by centrifugation at 5000 rpm for 1 minute. Supernatant was aspirated, resuspended gently with 100 µl of ddH₂O and spreaded onto solid media that selects for the plasmid. Transformed yeast cells were re-streaked onto a new selective plate before ready to use.

2.11 Yeast two-hybrid

Yeast strain pJ69-4A (genotype: *MATa trp1-901 leu2-3,112 ura3-52 his3-200 gal4Δ gal80Δ LYS::GAL1-HIS3 GAL2-ADE2 met2::GAL7-lacZ*) was used for yeast two-hybrid experiment. Bait vector (pAS2-1) cloned with one protein (IFI16, BRCA1, or BLM) and prey vector (pACT2) cloned with another protein were introduced into pJ69-4A by two rounds of PLATE transformations. Successful transformants were selected in -leu, -trp SC agar plates, as pAS2-1, which carried a TRP1 gene, and pACT2, which carried a LEU2 gene, would enable the transformed yeast to survive in -leu, -trp media. In addition, pAS2-1 generates a fusion of the GAL4 DNA binding domain (DBD) and the bait ; pACT2 generates a fusion of the GAL4 Activation Domain (AD) and the prey. Direct interaction between the two proteins would drive the expression of reporter

genes HIS3 or ADE2, with ADE2 being more stringent. Here, to investigate the direct interaction between two proteins, yeast carrying the two plasmids were streaked onto -leu, -trp, -ade SC agar plate and were checked for growth. Positive control for this experiment was a pair of proteins that were known to have direct interaction with each other: SNF1 and SNF2 (provided by Gabriel Alfaro of Beh Lab).

2.12 Radioactive labelling DNA with ^{32}P

2.12.1 5' end labelling – kinase assay

Oligonucleotides were 5'- ^{32}P -labeled with [γ ^{32}P]-ATP by forward labelling reaction. A 25 μl -reaction containing 15 μl of 100 μM DNA of interest, 1 μl of T4 Polynucleotide kinase (Invitrogen), 5 μl of forward reaction buffer (provided by kit), 2.5 μl of 10 mCi/ml [γ ^{32}P]-ATP (Perkin-Elmer) and 1.5 μl of ddH₂O was incubated in 37 °C for 1 hour. The reaction was inactivated by 20-min incubation at 65 °C, and 90% unincorporated [γ ^{32}P]-ATP was removed by ethanol precipitation of the oligonucleotides. DNA oligos were purified by gel purification method using a 15% TBE gel with 6 M urea was used. After gel extraction of the DNA, another round of ethanol precipitation was performed to isolate the [γ ^{32}P]-ATP-labelled DNA oligos. Purified DNA oligos were resuspended in 1X TE buffer.

2.12.2 3' end labelling

3' labelled oligonucleotides were prepared by adding [α^{32} -P] cordycepin (3'deoxyadenosine) 5'triphosphate to the 3' end of the oligonucleotide with terminal transferase. A 25 μ l-reaction containing 7.5 μ l of 100 μ M DNA of interest, 1 μ l of terminal transferase (New England BioLab), 2.5 μ l of 2.5 mM CoCl_2 , 2.5 μ l of NEB buffer 4 (provided with kit), 2.5 μ l of Labelled [α^{32} -P] cordycepin, and 9 μ l of ddH₂O was incubated in 37 °C for 1 hour. Radio-labelled oligonucleotides were purified as described (section 2.12.1).

2.13 Determination of labelled oligo concentration

Concentrations for the labelled oligonucleotides were determined by UV spectrometer. The standard conversion factor according to Invitrogen was used: 1 absorbance unit at wavelength 260 nm equals to 33 μ g/ml of ssDNA.

2.14 DNA unwinding assay

DNA substrates used in unwinding assay were generated by annealing radioactive oligonucleotides to Φ X174 virion DNA (New England Biolabs) in 1:1 molar ratio. Annealing was carried out by mixing Φ X174 virion DNA with the respective oligonucleotide (final concentration 200 nM) in TBE buffer (89 mM Tris, 89 mM boric acid, 2 mM EDTA, pH 8.4) containing 50 mM NaCl and 20 mM MgCl_2 . In a PCR machine, the annealing mixture was first incubated at 95 °C for 5 mins and was programmed to slowly cool down to 22 °C at a rate of

approximately 0.6 °C/min. Four DNA substrates were used in the unwinding assays: 1) 30-mer, A DNA substrate with a 30-bp double stranded region. 2) Bubble, with the same double stranded region as 30-mer except there is a 6bp bubble incorporated in the middle. 3) 5'T, prepared by annealing a 45-base oligo, which consists of the same DNA sequence as 30-mer with a poly(dT)₁₅ at its 5' end, to ΦX174 virion DNA. 4) 3'T, prepared by annealing a 45-base oligo, which consists of the same DNA sequence as 30-mer with a poly(dT)₁₅ at its 3' end, to ΦX174 virion DNA. (see Figure 2-3 and Table 2-6)

Unwinding of all DNA substrates were carried out in a final volume of 20 µl containing 20 mM Tris-HCl (pH 7.5), 4% (w/v) sucrose, 8 mM dithiothreitol (dT), 80 µg/ml bovine serum albumin (BSA), 5 nM of DNA substrate and one of the four IFI16 protein domains: PAAD, HIN200A, HIN200B, or 2HIN. For each domain, a range of concentrations obtained by 2-fold serial dilution were tested, starting from a protein concentration of 50 µM or as indicated. Incubation was for 60 minutes at 37 °C unless otherwise mentioned. Reactions were stopped using a final concentration of 1% SDS. After addition of loading dye, samples were electrophoresed in 8 - 12% non-denaturing acrylamide gel containing 20% glycerol at a constant 25 mA/gel for 2-3 hours or until dye front is 2/3 from the bottom of the gel. The running buffer used for electrophoresis was Tris-Glycine pH 8.8 prepared by mixing 3 g Tris and 14.4 g Glycine in 1 L of MilliQ water. Finally, the gel was autoradiographed using Typhoon Imager and quantified using ImageQuant software.

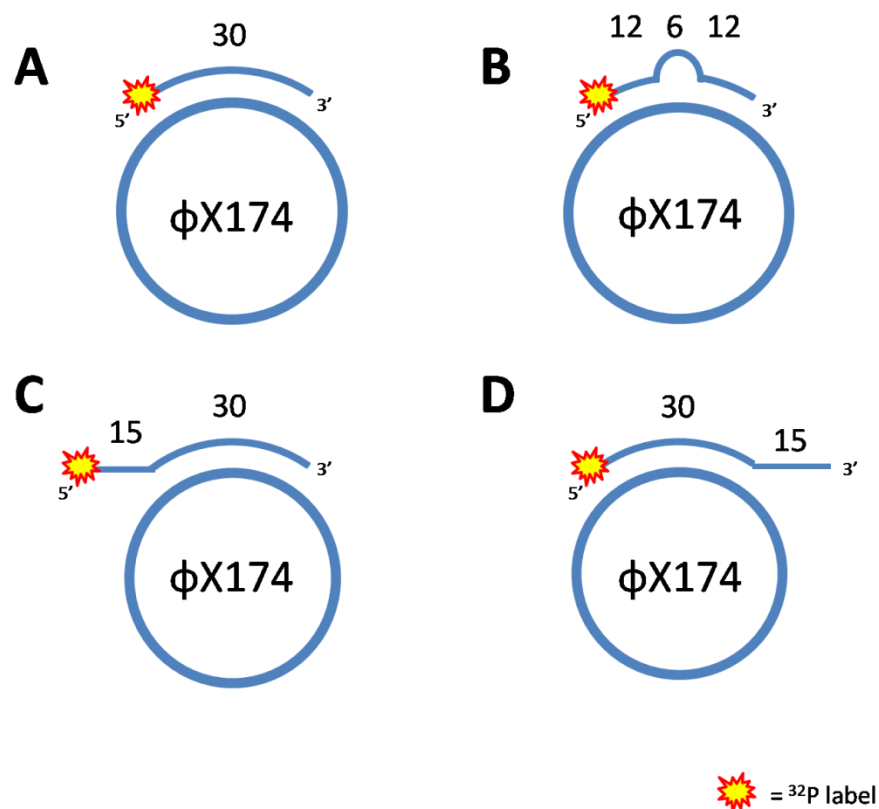


Figure 2-3. DNA substrates used in unwinding assay.

Four partial duplexes generated by annealing different oligonucleotides to Φ X174: **A)** a 30 nt oligonucleotide with sequence complementary to a 30-base region of Φ X174 (named 30-mer). **B)** a 30 nt oligonucleotide with exactly the same sequence as 30-mer except for 6 bases in the middle, creating a 6 nt bubble after it was annealed to Φ X174 (named Bubble). **C)** a 45-nt long oligonucleotide consisting of a 15-nt long poly(dT) tail at the 5' end of the strand followed by 30-mer sequence (named 5'T). **D)** a 45-nt long oligonucleotide similar to 5'T except the 15-nt long poly(dT) tail is at the 3' end of the strand instead (named 3'T).

Table 2-6. Oligonucleotides used in DNA unwinding assay

Name of oligos	Nucleotide sequence
30-mer	5'-AAGTAAGAGCTTCTCGAGCTGCGCAAGGAT-3'
Bubble	5'-AAGCGATAAAACTACATGAGTTGGATACGC-3'
5'T	5'-TTTTTTTTTTTTTTTTTAAGTAAGAGCTTCTCGAGCTGCGCAAGGAT-3'
3'T	5'-AAGTAAGAGCTTCTCGAGCTGCGCAAGGATTTTTTTTTTTTTTTTTT-3'

2.15 DNA cleavage assay

Reaction samples for cleavage assay were prepared in exactly the same way as for unwinding assay except the DNA oligos were not annealed to Φ X174 virion DNA. Cleavages of all DNA oligos (Table 2-7) were carried out in a final volume of 20 μ l in the same reaction buffer as unwinding assay. Incubation was for 60 minutes at 37 °C unless otherwise mentioned. Reactions were stopped with 20 μ l of urea loading dye (8 M Urea, 2 mM Tris, pH 7.5, 20 mM EDTA, 1.68 mg/ml bromophenol blue, xylene cyanol) and heat inactivated at 95 °C for at least 5 minutes followed by a quick transfer to ice incubation for 2 minutes. Samples were electrophoresed using 20% gel containing 6 M urea and 20% glycerol at a constant 20 Watts/gel (equivalent to approximately 750 V) for 25 - 30 minutes. High urea concentration and high voltage electrophoresis were used to ensure all DNA molecules are completely denatured and free of secondary structure. The gel was autoradiographed as described previously and quantified using ImageQuant software.

Table 2-7. Oligonucleotides used in DNA cleavage assay

Name of oligos	Nucleotide sequence
30-mer ^A	5'-AAGTAAGAGCTTCTCGAGCTGCGCAAGGAT-3'
5'T ^A	5'-TTTTTTTTTTTTTTTAAGTAAGAGCTTCTCGAGCTGCGCAAGGAT-3'
T ₇₀ ^A	Poly(dT) ₇₀
GC-5 ^B	5'-GGAAGAAGGAAGTGGGATCAGGATCCGCTGGCTCC-3'
GC-3 ^B	5'-GGAGCCAGCGGATCCTGATCCCACTTCCTTCTTCC-3'

^A both 5' and 3' end labelled oligos were tested

^B only 5' end labelled oligos were tested

3: RESULTS

3.1 The first HIN200 domain of IFI16 has similar DNA binding properties as RPA

This section 3.1 is composed of results from the following publication:

Yan H, Dalal K, **Hon BK**, Youkharibache P, Lau D, Pio F. (2008) RPA nucleic acid-binding properties of IFI16-HIN200. *Biochim Biophys Acta* **1784**(7-8):1087-1097.

Note regarding to authors contribution: HY produced the IFI16-HIN200 construct and optimized the stability condition for the purified protein. DL assisted in secondary and tertiary structure analysis for IFI16-HIN200. HY performed all UV cross-linking experiments and tyrosine fluorescence quenching experiments while **BKH** repeated the experiments for reproducibility and standard deviation data. KD contributed to all EMSA experiments. **BKH** performed multiple site-directed mutageneses to produce the mutant protein required for DNA polarity assay and assisted in collecting data for DNA melting experiments. HY and KD collected data for FRET assay that showed IFI16-HIN200 can wrap and stretch ssDNA. HY and **BKH** collected FRET data for DNA polarity assay. PY contributed to the comparative modelling of IFI16-HIN200 using a 3D jury method (Ginalski *et al.* 2003). FP contributed to the development of hypothesis, experimental design, and writing of the manuscript assisted by mainly HY, and KD.

3.1.1 IFI16-HIN200A binds to single-stranded nucleic acid

IFI16-HIN200A (residues 194-393) recombinant protein of approximately 27 kDa was expressed and purified as described (Section 2.1.1 and 2.3.1). To demonstrate IFI16-HIN200A interaction with ssDNA, UV cross-linking

experiments were performed with the following oligonucleotides: A_{25} , T_{25} , GC-5, and GC-3 (Section 2.4). The molecular weights of A_{25} and T_{25} were approximately 7 – 8 kDa while GC-5 and GC-3 were between 10 – 11 kDa. IFI16-HIN200A and oligonucleotide were cross-linked together after UV treatment and the covalently linked complex was isolated by 15% SDS-PAGE. After silver staining the gel, protein monomer (M), dimer (D), and protein-ssDNA complex (C) were identified (Figure 3-1). The cross-linking of IFI16-HIN200A with A_{25} and T_{25} (Figure 3-1, A) showed a complex band at the position that agreed with the experimental weight of one IFI16-HIN200A monomer plus one A_{25} or T_{25} molecule (Table 3-1). This band was not observed when either oligonucleotides was excluded. IFI16-HIN200A interaction with dsDNA was also investigated by cross-linking with annealed $(A-T)_{25}$. A complex band was also detected but the molecular weight was not different from the complex found for cross-linking with ssDNA. The same result was found for GC-5 and GC-3 and their double-stranded form (Figure 3-1, C). Furthermore, immunoblotting with an antibody against the his-tag on IFI16-HIN200A after UV cross-linking showed the presence of IFI16-HIN200A protein in the protein-nucleic acid complexes (Figure 3-1, B and D).

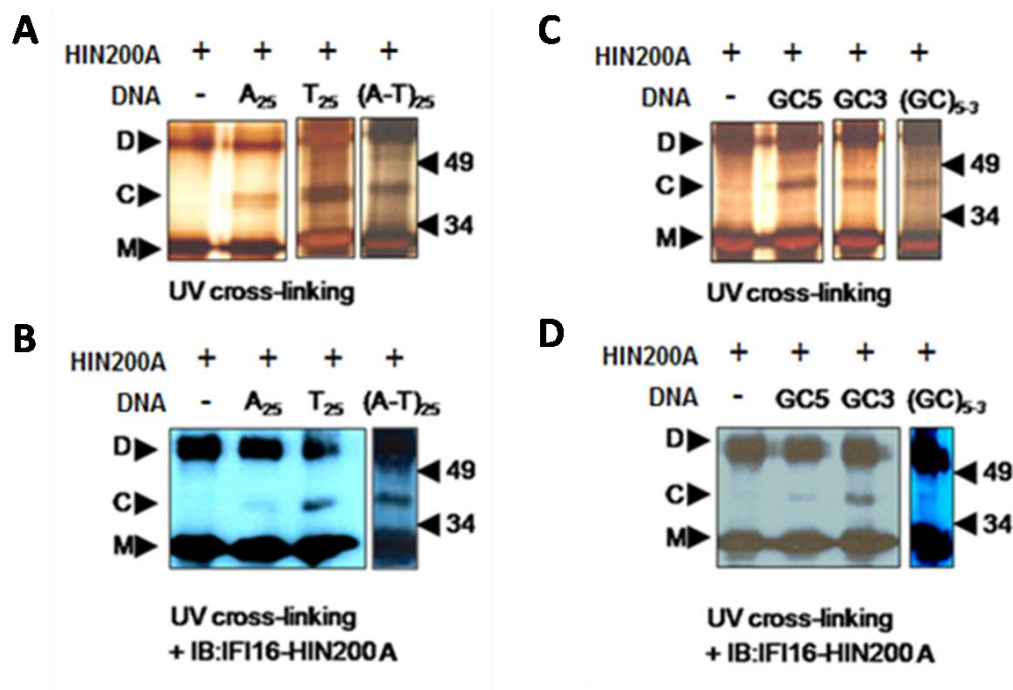


Figure 3-1. Protein(IFI16-HIN200A)-ssDNA complex was detected after UV cross-linking (Yan *et al.* 2008).

IFI16-HIN200A and various ssDNA formed complex after 15 minutes of UV treatment in 20 mM sodium cacodylate, pH 7.25. **A)** SDS-PAGE of HIN200A with poly(dA)₂₅, poly(dT)₂₅, or double-stranded (A-T)₂₅ showed complex band (C) at the molecular weight that correspond well to the sum of the molecular weights of the monomeric protein (M) and a ssDNA oligo. This complex band was not observed when ssDNA oligo was excluded from the reaction. Dimeric HIN200A (D) was found in all lanes. **B)** Immunoblotting of the SDS-PAGE gel with anti-his tag antibodies confirmed that the complex band contained IFI16-HIN200A. **C)** and **D)** By the same approach, IFI16-HIN200A also binds to GC-rich ssDNA and dsDNA oligos (GC5, GC3, (GC)₅₋₃). On the right of all figures are the molecular markers in kDa.

To further confirm that the complex bands observed were correct, a more accurate method of molecular weight (MW) determination, Log MW vs R_f plot, was used. For each gel, a calibration curve was generated by plotting the molecular weights (MW) of each band from the protein ladder, run simultaneously with the sample, against their relative mobility value (R_f). IFI16-HIN200A has a theoretical MW of 27.3 kDa but in all SDS-PAGE gels, the monomer ran slower than its expected mobility and resulted in a higher

experimental MW than the theoretical. Table 3-1 summarizes all molecular weights calculated, with each being the average of results from at least 3 gels. In all cases, experimental MW of protein-DNA complex agreed well with their theoretical MW.

Table 3-1. Experimental molecular weights of IFI16-HIN200A, DNA, and protein-nucleic acid complex from SDS-PAGE (Yan *et al.* 2008)

DNA		Protein monomer	Protein-DNA complex T ^a	Protein-DNA complex E ^b	Difference between protein complex and monomer
Molecular Weight (kDa)					
A ₂₅	7.8	31.2 ± 1.0	39.0 ± 1.0	39.5 ± 1.3	8.3 ± 2.3
T ₂₅	7.5	31.6 ± 1.2	39.1 ± 1.2	39.6 ± 1.3	8.0 ± 2.5
(A-T) ₂₅	15.3	30.1 ± 2.3	37.9 ± 2.3 or 37.6 ± 2.3 ^c	40.7 ± 0.9	10.6 ± 3.2
GC-5	10.9	31.1 ± 1.2	42.0 ± 1.2	41.4 ± 1.3	10.3 ± 2.5
GC-3	10.6	31.2 ± 1.4	41.8 ± 1.4	41.2 ± 1.0	10.0 ± 2.4
(GC) ₅₋₃	21.5	30.7 ± 2.1	41.6 ± 2.1 or 41.3 ± 2.1 ^d	41.7 ± 1.1	11.0 ± 3.2

^a Theoretical molecular weight of protein-DNA complex, the sum of protein monomer and actual molecular weight of DNA

^b Experimental molecular weight of protein-DNA complex, based on position of the complex band

^{c,d} Theoretical molecular weight assuming protein binds only to single-stranded DNA

An Electrophoretic mobility shift assay (EMSA) was also performed with all oligonucleotides. With each radio-labelled oligonucleotide, the addition of IFI16-HIN200A resulted in a slower radio-probe migration in native gel. This indicated that the oligonucleotides formed complexes with the protein, changing its mobility shift (data not shown). In addition, the same assay was used to investigate HIN200A interaction with RNA. A 43-mer RNA library was electrophoresed with IFI16-HIN200A at various protein concentrations (Figure 3-2). Radio-labelled probe mobility was shifted as IFI16-HIN200A concentration increased, signifying

a protein-RNA complex formation. Together with the UV cross-linking data, these data demonstrate that IFI16-HIN200A has single stranded nucleic-acid binding properties.

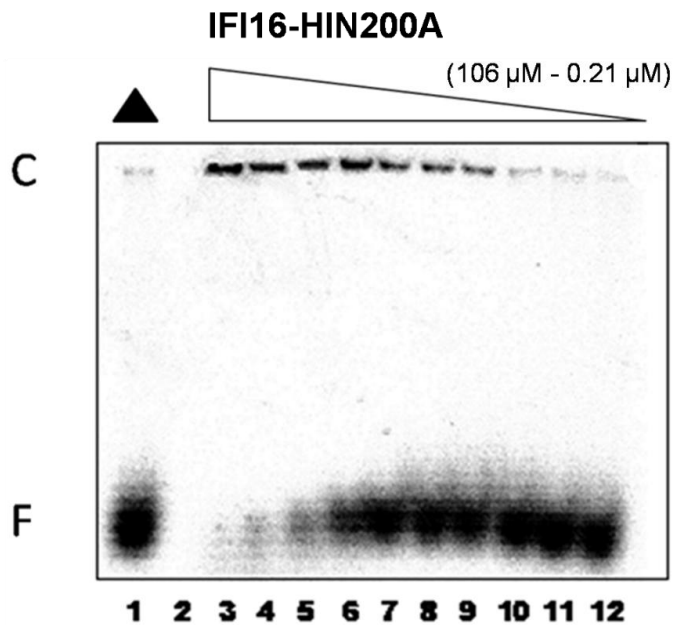


Figure 3-2. IFI16-HIN200A also binds RNA (Yan *et al.* 2008).

Various IFI16-HIN200A concentrations (106 μ M – 0.21 μ M) were assayed with constant 360 nM of radio-labelled 43-mer RNA library in 20 mM sodium cacodylate, pH 7.25 at room temperature for 10 minutes. At high protein concentration, IFI16-HIN200A formed complex with RNA, shown by slow migrating band at the top of the 6% native gel (C). As protein concentration decreased, the mobility of RNA was shifted to the free-probe position (F), which was verified in the negative control lane (\blacktriangle) containing only RNA and no IFI16-HIN200A.

3.1.2 IFI16-HIN200A has higher binding affinity for GC-rich ssDNA than dsDNA

Next, we examined whether there was a preference for IFI16-HIN200A in the type of DNA it binds to. Here, we used tyrosine fluorescence quenching assays to determine the ssDNA binding property and affinities for IFI16-HIN200A on different types and lengths of oligonucleotides, including single-stranded and double-stranded form of T₂₅, A₂₅, and GC-rich 35-mers. In this assay, the

intrinsic tyrosine fluorescence of IFI16-HIN200A at its maximum emission wavelength 304 nm ($\lambda_{em} = 304$ nm) was progressively quenched as more oligonucleotides were added into the reaction to allow interaction with the protein. The percentage of tyrosine fluorescence quenched was plotted against the concentration of each oligonucleotide and the dissociation constants (K_d) of the bindings were estimated by the software GraphPad Prism (Section 2.6). All of the B_{max} values obtained by fitting the fluorescence quenching data to one-site or two-site binding equations are around 100%, even though no constraints were imposed during the non-linear regression, indicating that the intrinsic tyrosine fluorescence was completely quenched over the titration range (Table 3-2). The relative affinity ranking for all oligonucleotides tested from the highest to the lowest was GC-5 > GC-3 > (A-T)₂₅ > T₂₅ > GC₅₋₃ > A₂₅. For A₂₅ and T₂₅ ssDNA sequences (Table 3-2), one-site binding was the best model ($K_{d1} \sim 10^{-6}$ M), while two-site binding was preferred for the dsDNA (A-T)₂₅ ($K_{d2} \sim 10^{-7}$) but the affinity was slightly higher for the dsDNA than the ssDNA. On the contrary, for GC-rich DNA sequences we observed opposite results: ssDNA binding data agreed better to a two-site binding model, while dsDNA preferred one-site binding. Both K_d 's obtained by the two-site binding models of GC-rich ssDNA showed a large difference in binding affinity: $K_{d1} \sim 10^{-6}$ M, for both ssDNA and dsDNA, and $K_{d2} \sim 10^{-8}$ - 10^{-9} M, for the GC-rich ssDNA only. This result indicated that the binding affinity of IFI16-HIN200A to ssDNA is 10^3 times higher compared to dsDNA. For the six different DNA sequences tested, we found that GC-5 and GC-3 have the highest binding affinity to IFI16-HIN200A, which may indicate that this protein

domain has a binding preference to GC-rich ssDNA sequences. These data suggest that IFI16-HIN200A has different binding modes that are sequence dependent as shown for RPA (Blackwell *et al.* 1994; Wold 1997; Kim *et al.* 1994).

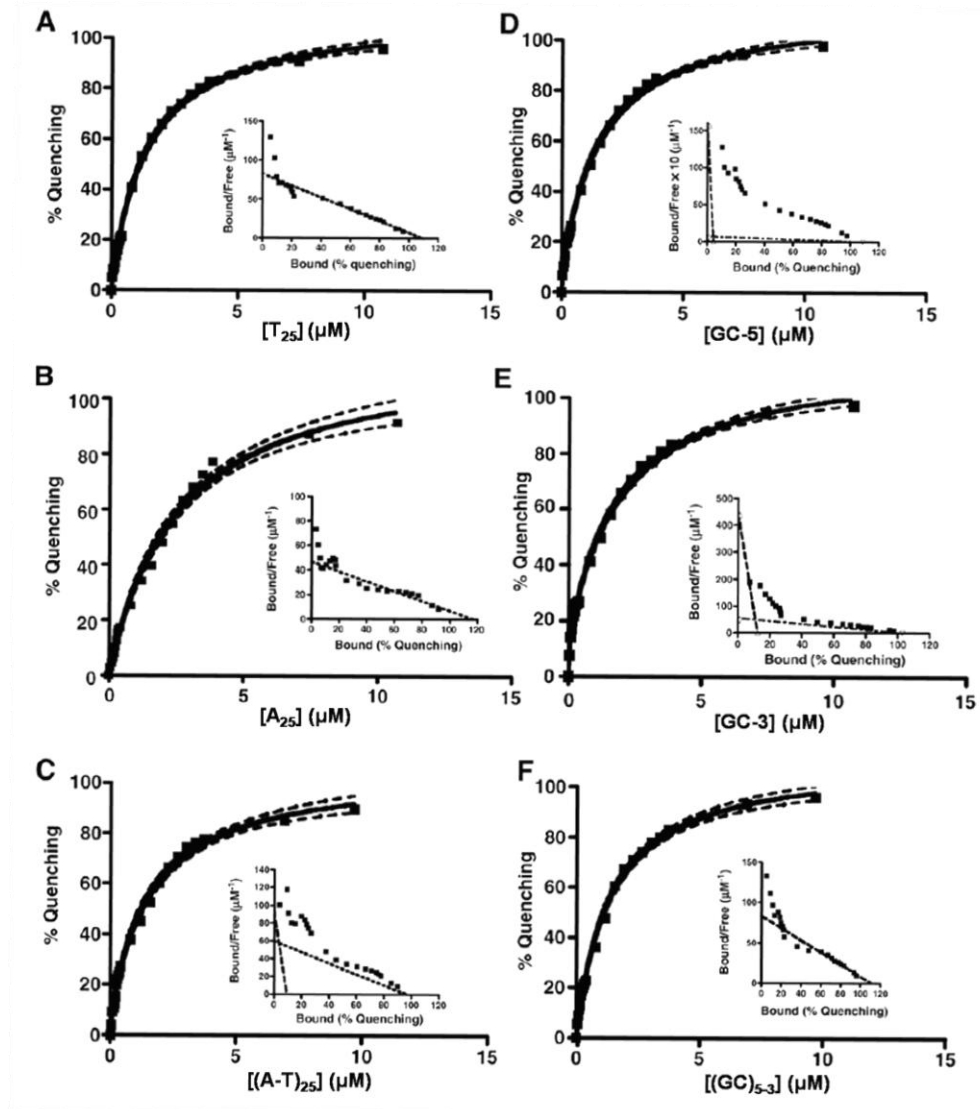


Figure 3-3. IFI16-HIN200A has a higher binding preference to GC-rich DNA (Yan *et al.* 2008).

IFI16-HIN200A (3.7 μM) tyrosines were titrated by T_{25} (A), A_{25} (B), dsDNA $(\text{A-T})_{25}$ (C), GC-5 (D), GC-3 (E) and dsDNA $(\text{GC})_{5-3}$ (F) from DNA:protein molar ratio 0.1 to 3.0 in 20 mM sodium cacodylate, pH 7.25. The percentage of tyrosine fluorescence quenched at 304 nm was plotted against DNA concentration (μM) to estimate the binding affinity of IFI16-HIN200 to different DNA sequences. The inset is the Scatchard plot that indicates the preferential binding mode for each DNA sequence.

Table 3-2. Dissociation constants (K_d) and B_{max} values for IFI16-HIN200A in tyrosine fluorescence quenching (Yan *et al.* 2008)

ssDNA or dsDNA	DNA	K_d1 ($\times 10^{-6}M$)	B_{max1} (%)	K_d2 ($\times 10^{-6}M$)	B_{max2} (%)
ssDNA	T ₂₅	1.331 \pm 0.047	109.5 \pm 1.4	--	--
ssDNA	A ₂₅	2.523 \pm 0.208	117.9 \pm 4.2	--	--
dsDNA	(A-T) ₂₅	1.616 \pm 0.320	96.2 \pm 7.3	0.101 \pm 0.160	9.4 \pm 8.8
ssDNA	GC-5	1.485 \pm 0.098	108.9 \pm 1.5	0.003 \pm 0.023	4.4 \pm 1.5
ssDNA	GC-3	1.856 \pm 0.172	102.9 \pm 2.3	0.028 \pm 0.022	12.5 \pm 2.4
dsDNA	(GC) ₅₋₃	1.338 \pm 0.072	111.3 \pm 2.1	--	--

Note: refer to Table 2-2, Section 2.4 for nucleotide sequences

3.1.3 IFI16-HIN200A oligomerizes, compacts and extends ssDNA after binding

IFI16-HIN200A appeared to have more than one DNA binding mode according to tyrosine fluorescence quenching result. In RPA, different DNA binding modes are triggered by oligomerization of OB folds. To examine if ssDNA binding also triggers oligomerization of IFI16-HIN200A, chemical cross-linking was performed. The protein was incubated with the chemical cross-linker glutaraldehyde, which only cross-links lysines, to probe protein-protein interaction. In this assay the cross-linking was performed in the absence and presence of T₂₅. Samples were loaded on SDS-PAGE and analyzed by silver-staining the gel. Without T₂₅, faint bands for IFI16-HIN200A oligomers were observed. The intensity of these bands was greatly enhanced by the addition of T₂₅ at 0.1:1 DNA:protein molar ratio (Figure 3-4). The molecular weight of these bands correspond to the monomer, dimer, trimer, and tetramer of IFI16-HIN200A. Therefore, IFI16-HIN200A oligomerizes and this is enhanced upon DNA binding.

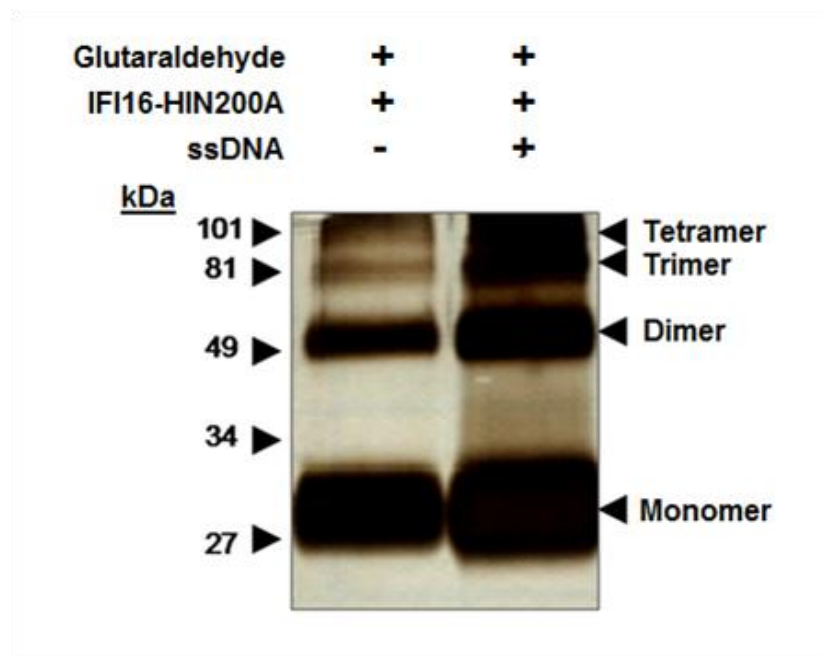


Figure 3-4. IFI16-HIN200A forms different oligomers upon DNA binding (Yan *et al.* 2008).

IFI16-HIN200A was cross-linked by glutaraldehyde in 20 mM HEPES, pH 7.5, with or without T₂₅. On the silver-stained 15% SDS-PAGE, lane 1: protein + glutaraldehyde; lane 2: protein + glutaraldehyde in the presence of T₂₅ at a DNA:protein molar ratio of 0.1:1. Each band corresponds to different oligomers: monomer, dimer, trimer and tetramer, as indicated. Left arrows: molecular weight standards (kDa).

Next, FRET was used to determine if IFI16-HIN200A possesses DNA compaction and extension abilities as other OB-fold proteins (Robbins *et al.* 2005). Double-stranded DNA, named FRET 18-58 (Section 2.9.1), was titrated by IFI16-HIN200A from 1:1 to 20:1 protein:DNA molar ratio and the emissions of the donor (Quasar570 at wavelength 567 nm) and the acceptor (Quasar670 at wavelength 667 nm) were measured when excited at wavelength 548 nm (excitation wavelength of Quasar570). The plot of FRET efficiency versus protein:DNA molar ratio showed a progressive decrease in the emission of the

donor and increase in the emission of the acceptor from 0 to 10 protein:DNA molar ratio (Figure 3-5). This indicated the fluorophores were brought closer together (higher energy transfer) with respect to increasing IFI16-HIN200A concentration until protein:DNA molar ratio reached 10:1. Conversely, above 10:1 protein:DNA molar ratio, the FRET efficiency began to decrease, showing that the fluorophores were more distant from each other (Figure 3-5). These results may suggest that the poly(dT)₄₀ tail first wraps around IFI16-HIN200A molecules and becomes more extended when more protein was added.

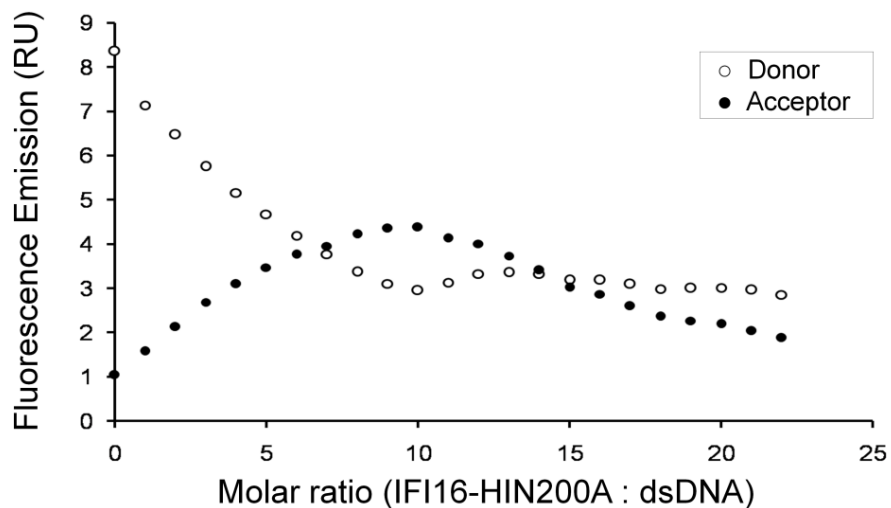


Figure 3-5. IFI16-HIN200A can compact and extend ssDNA (Yan *et al*, 2008).

FRET measurements were performed in 20 mM sodium cacodylate, pH 7.25, at a constant 300 nM of the partial duplex FRET oligonucleotide in response to increasing HIN200A concentration. Maximum emission of donor Quasar 570 (564 nm) and acceptor Quasar 670 (670 nm) were plotted against the molar ratio of IFI16-HIN200A:dsDNA. FRET energy transfer is indicated by decrease in donor emission while acceptor emission is increased.

3.1.4 IFI16-HIN200A does not destabilize dsDNA

An important property of OB fold proteins is the ability to destabilize dsDNA (Treuner *et al.* 1996). If IFI16-HIN200A destabilizes dsDNA, the DNA melting point should be lower than that without the protein. For the (A-T)₂₅ oligonucleotide the melting point of 51 ± 2 °C (n = 3) remained unchanged with and without the protein (Figure 3-6). These data suggest a slight stabilization of the dsDNA by IFI16-HIN200 but remain at the sensitivity threshold of the melting depression assay (3 °C). This observation was different from what was found in RPA and IFI16-PAAD (Figure 3-7), which was previously determined to lower the melting temperature of the double-stranded (A-T)₂₅ by approximately 10 °C (Section 1.2.5).

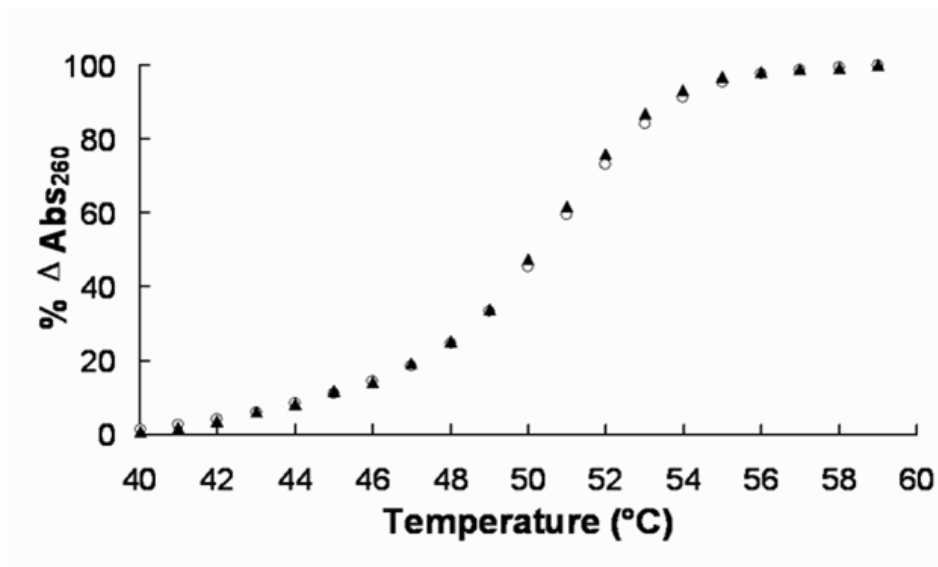


Figure 3-6. IFI16-HIN200A does not destabilize dsDNA (Yan *et al.* 2008).

Melting point depression of 10 μM dsDNA (A-T)₂₅ in 20 mM sodium cacodylate and 100 mM NaCl, pH 7.25. The T_m was around 51.7 °C without protein (▲) and remain unchanged upon the addition of protein at 1:1 protein:DNA molar ratio (○). The T_m was obtained from 3 independent experiments.

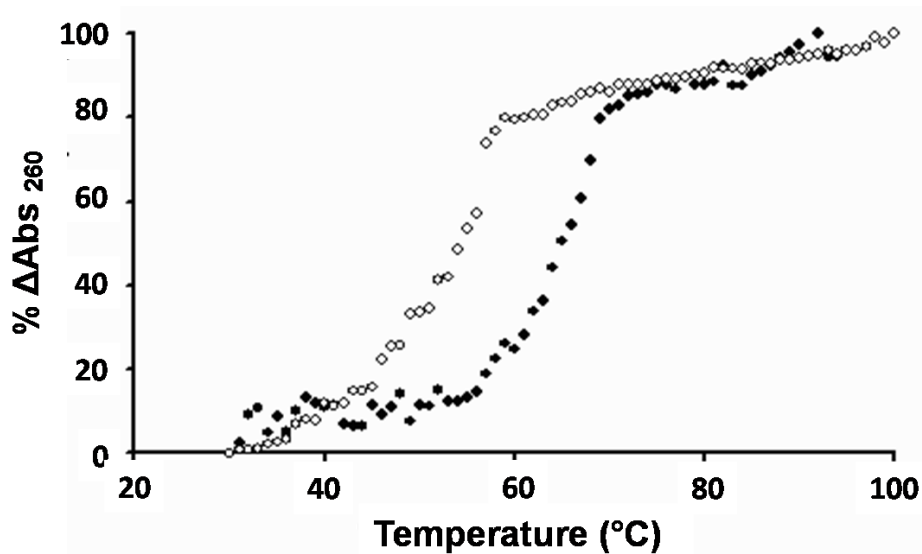


Figure 3-7. Double-stranded DNA destabilization by IFI16-PAAD (Dalal K, M.Sc thesis, 2006 with permission).

Melting point depression of dsDNA with IFI16-PAAD in 20 mM HEPES, pH 7.5. A₂₆₀ of 10 μM dsDNA in presence or absence of 25 μM IFI16-PAAD was monitored as temperature was gradually increased from 20 °C to 100 °C. The dsDNA alone (●) has a melting point of ~65 °C but is lowered by 10 °C in presence of IFI16-PAAD (○).

3.1.5 ssDNA-binding polarity of IFI16-HIN200A

To investigate if IFI16-HIN200A binds ssDNA in the same orientation as RPA, a polarity assay using FRET was carried out (Section 2.9.2 and Figure 2-2). In this experiment, the DyLight549-labeled wildtype and mutant IFI16-HIN200A were titrated by Quasar670-labelled 18-mer ssDNA (Table 2-4) from 1:1 to 12:1 DNA:protein molar ratio, respectively. This assay was repeated 4 times on both wildtype and mutant constructs. The FRET efficiency, calculated based on one of the 4 independent FRET experiments using wildtype and mutant IFI16-HIN200A, was plotted against the corresponding molar ratio (Figure 3-8). From the curve, it can be clearly seen that there was more energy transfer between the wildtype protein and DNA than the mutant, indicating that the

labelled 5' end of DNA is closer to the C-terminal OB fold. This binding orientation was the same as the prevalent 3'-5' binding polarity of RPA with ssDNA as well as most other OB fold proteins (Theobald *et al.* 2003).

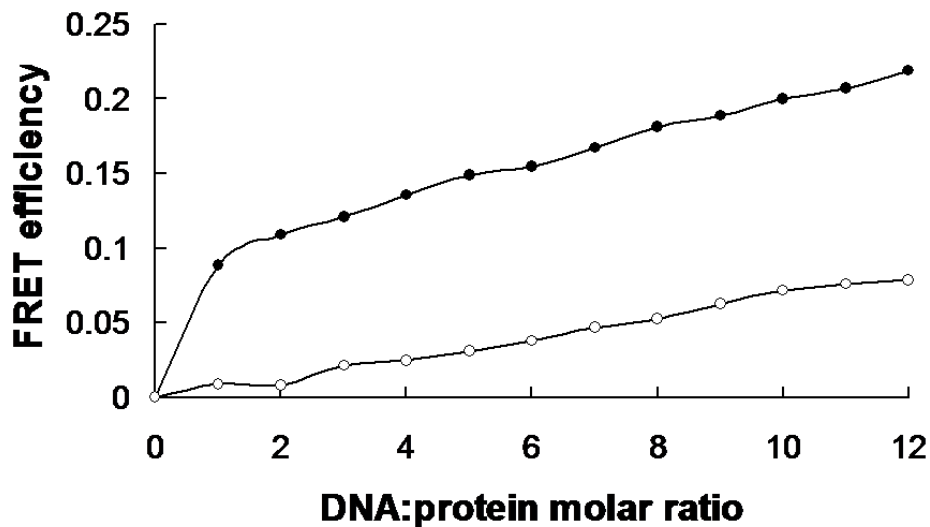


Figure 3-8. IFI16-HIN200A binds the 18-mer ssDNA in 3'-5' orientation (Yan *et al.* 2008).

The energy transfer between wildtype IFI16-HIN200A and the 18-mer ssDNA (●) was greater than the energy transfer between the mutant IFI16-HIN200A and the 18-mer ssDNA (○). FRET efficiency was plotted versus the DNA:protein molar ratio, where HIN200A concentration was fixed at 0.767 μ M. At each titration point, the energy transfer between wildtype protein and DNA is higher than that between mutant protein and DNA. This indicates a 3'-5' orientation of ssDNA when bound to IFI16-HIN200A.

In summary, the first HIN200 domain of IFI16 was determined to have similar nucleic acid binding properties as the eukaryotic ssDNA-binding protein RPA without dsDNA destabilizing properties.

3.2 IFI16 unwinds DNA

3.2.1 IFI16-HIN200A has DNA unwinding activity

RPA has the ability to unwind DNA (Georgaki *et al.* 1992). To determine if IFI16-HIN200A can also unwind DNA like RPA, I developed a DNA unwinding assay. In this assay, a 30-nt long oligonucleotide with a sequence that was completely complementary to Φ X174 was radiolabelled with ^{32}P and hybridized to single-stranded Φ X174 virion DNA as described (Section 2.14). Different concentrations of IFI16-HIN200A (refer to appendix A for purity of protein) were tested for their ability to separate the radiolabelled oligonucleotide from Φ X174. Samples were electrophoresed in native conditions and autoradiography indicated the unwinding of 30-nt oligonucleotide from Φ X174 virion DNA was found only when IFI16-HIN200A was present. (Figure 3-9). When hybridized to Φ X174 virion DNA, migration of the probe was much slower than its free-state. Increasing IFI16-HIN200A concentration from 0 to 80 μM decreased the amount of bound-state probe from maximum to none. In addition, the probe unwound from Φ X174 virion DNA appeared to have a second mobility shift and this observation will be discussed in more detail in section 3.3. IFI16-HIN200A unwinding activity was visible to a protein concentration of approximately 2.5 μM (Figure 3-9, lane 3).

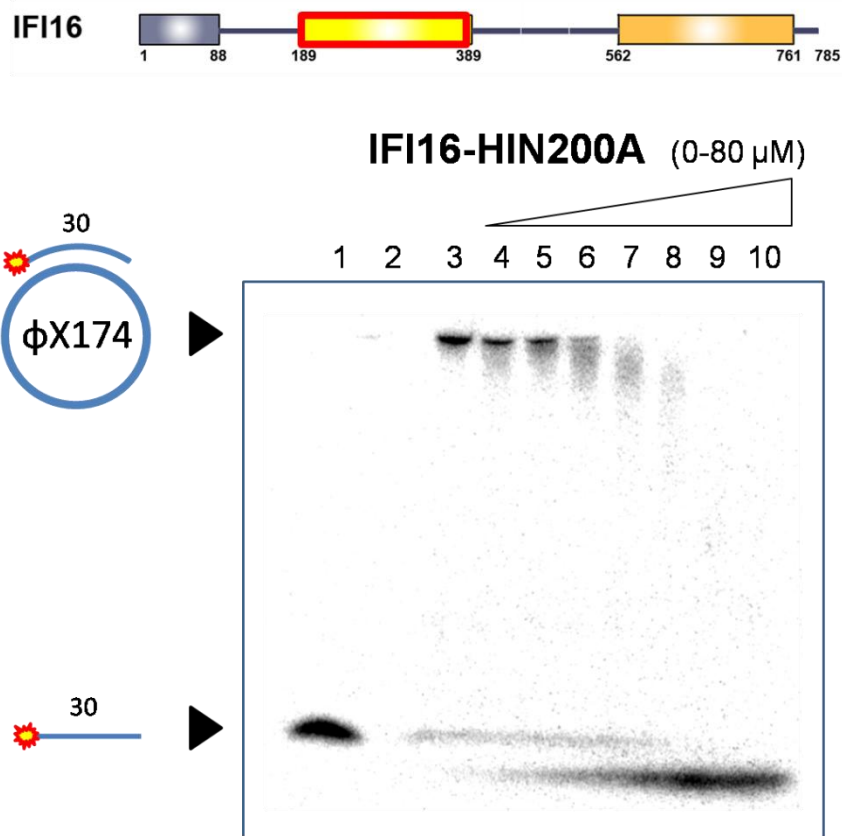


Figure 3-9. IFI16-HIN200A unwinds Φ X174 partial duplex.

Φ X174 partial duplexes (5 nM) were incubated with 2-fold increasing concentrations of IFI16-HIN200A (domain highlighted in red) in 20 mM Tris-HCl, 4% sucrose, 8 mM dTT, 80 μ g/ml BSA, and 2 mM MgCl_2 , pH 7.5, at 37 $^\circ\text{C}$ for 1 hour. Reactions were electrophoresed on 10% native gels containing 20% glycerol and visualized by autoradiography. Lane 1: positive unwinding control, heat-denatured Φ X174 partial duplex. Lane 2: empty. Lane 3: negative protein control, Φ X174 partial duplex without IFI16-HIN200A protein. The radioactive bands migrate to the top of the gel, indicating the DNA was not unwound. Lane 4-10: as IFI16-HIN200A concentration was increased by serial 2-fold increments from 1.25-80 μM , the radioactive 30-mer probe shifted to the bottom of the gel indicating its dissociation from Φ X174 by unwinding. The position of the probe does not correspond to the position of the unwinding control (lane 1), suggesting an additional modification to the probe.

As previously mentioned, IFI16-PAAD has the ability to bind different types of ssDNA as well as to destabilize dsDNA (Section 1.2.5, Figure 3-7). It is therefore interesting to see whether IFI16-PAAD, having these properties, can also unwind duplex DNA. DNA unwinding assay was performed on IFI16-PAAD and surprisingly, it appeared to have no unwinding effect on the DNA substrate

since there was no increase in radioactivity for the free probe observed along with increasing IFI16-PAAD concentration (Figure 3-10). The free probes observed in every lane (Figure 3-10, lower part of lane 1-8) resulted from incomplete annealing during the hybridization step. IFI16-PAAD was confirmed to have no unwinding activity even at a protein concentration up to 200 μM .

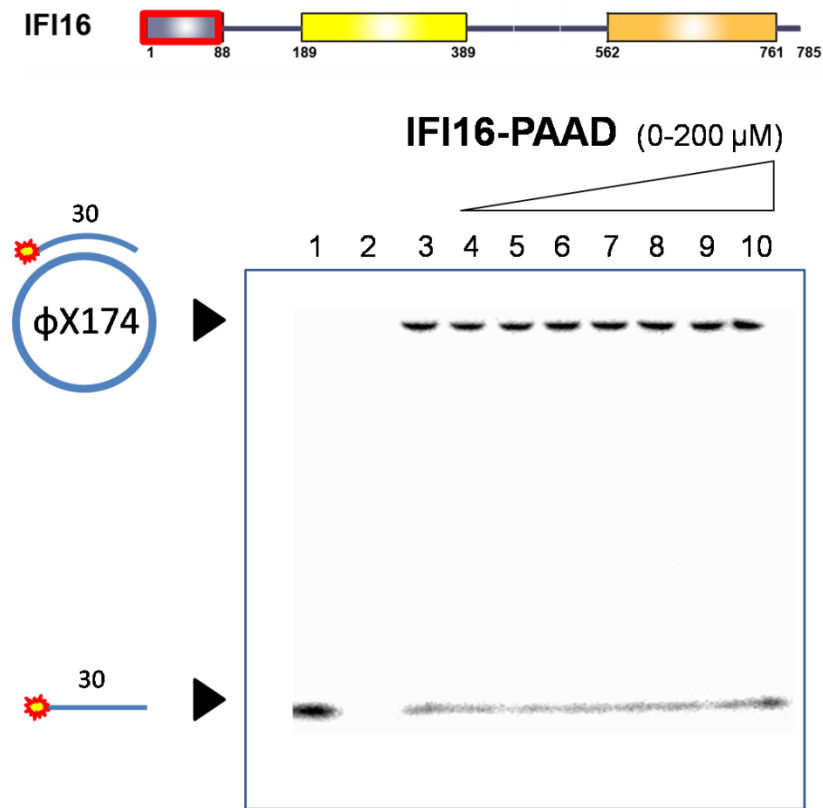


Figure 3-10. IFI16-PAAD has no DNA unwinding activity.

DNA unwinding assay for IFI16-PAAD (domain highlighted in red) was performed as described in materials and methods (Section 2.14). Lane 1: positive unwinding control, heat-denatured ϕX174 partial duplex. Lane 2: empty. Lane 3: negative protein control, ϕX174 partial duplex without IFI16-PAAD protein. DNA substrate is not unwound but excess free probes that did not get annealed can be seen. Lane 4-10: IFI16-PAAD from 3.1-200 μM obtained by 2-fold serial dilutions did not dissociate more free probe from ϕX174 , since the amount of 30-mer free probe is the same from lane 4-10 as in lane 3.

IFI16-HIN200A but not IFI16-PAAD has DNA unwinding activity. This could be due to the lack of OB fold in the PAAD domain, which is composed of mainly α -helix structures. To determine if this is the justification, I looked at another OB-fold domain in IFI16 which differs from HIN200A, and that is the second HIN200 domain (HIN200B). Successful unwinding of 30-nt oligonucleotide from Φ X174 was clearly observed within the IFI16-HIN200B concentration range of 0 μ M to 50 μ M (Figure 3-11). It was believed that the secondary shift in mobility of unwound probe was also present but it was not shown as obviously as in the unwinding of IFI16-HIN200A. This might be due to poor quality in band separation which in turn depends on the quality of the native PAGE used, and deviation between different gel preparations was difficult to avoid. Another experiment at higher gel concentration should be run on samples that contain the 30 nt free probe with increasing concentration of the protein IFI16-HIN200B.

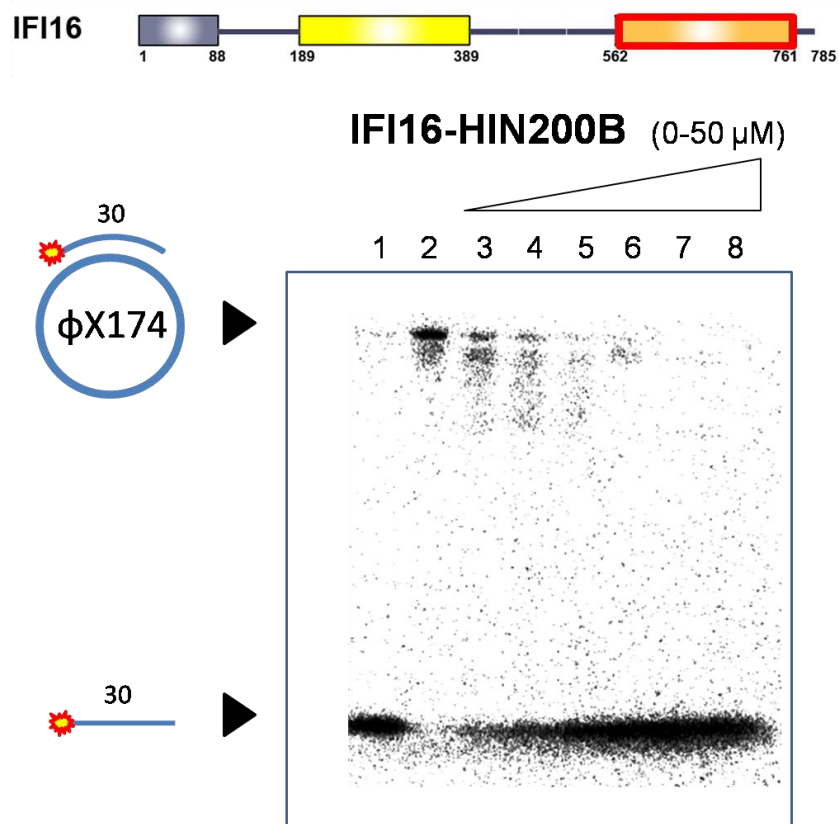


Figure 3-11. HIN200B of IFI16 unwinds Φ X174 partial duplex.

DNA unwinding assay for IFI16-HIN200B (domain highlighted in red) was performed as described in materials and methods (Section 2.14). Lane 1: positive unwinding control, heat-denatured Φ X174 partial duplex. Lane 2: negative protein control, Φ X174 partial duplex without IFI16-HIN200B protein. Lane 3-8: Increasing IFI16-HIN200B concentration from 1.56-50 μ M obtained by 2-fold serial dilutions slowly shifted the radioactivity from the top to the bottom of the gel, indicating an unwinding activity by IFI16-HIN200B.

Finally, the construct containing both HIN200A and HIN200B domains (2HIN) was tested. Over-expression of IFI16-2HIN (194-767 aa, 65.6 kDa) followed by his-tag purification yielded multiple over-expressed products, and gel filtration was used to isolate the products corresponding to the correct molecular weight of IFI16-2HIN (Appendix A). Purified IFI16-2HIN was examined in DNA unwinding assay and interestingly, IFI16-2HIN showed no DNA unwinding activity (Figure 3-12). These data show that the unwinding activity of the

individual HIN200 protein is inhibited. This inhibition maybe alleviated in certain cellular conditions via other proteins and/or factors that may bind to the protein leading to a conformational switch of the protein from an inactive to an active state.

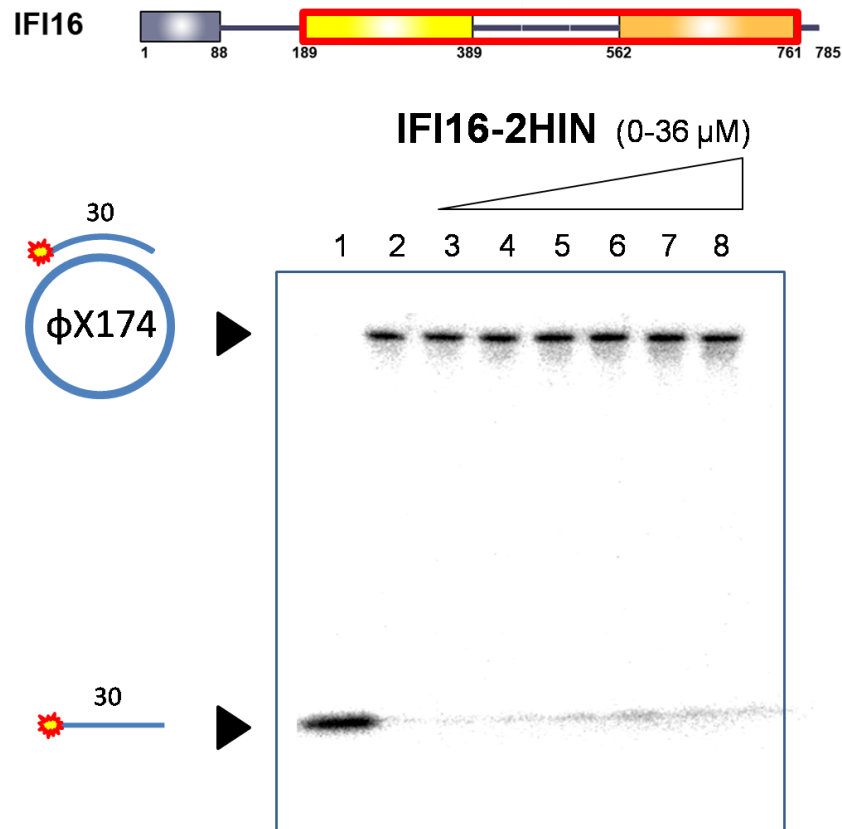


Figure 3-12. DNA unwinding activity was inhibited in IFI16-2HIN, which contains both HIN200A and HIN200B.

IFI16-2HIN (highlighted in red) contains both HIN200A and HIN200B. While each domain individually has DNA unwinding activity, IFI16-2HIN did not unwind the Φ X174 partial duplex substrate. Lane 1: positive unwinding control, heat-denatured Φ X174 partial duplex. Lane 2: negative protein control, Φ X174 partial duplex without IFI16-2HIN protein. Lane 3-8: 2-fold increasing concentrations of IFI16-2HIN from 1.125-36 μ M did not result in increasing radioactivity at the free probe position compared to the negative protein control (lane 2). This indicates the DNA unwinding activity was inhibited.

To summarize, IFI16-PAAD does not have DNA unwinding activity while both OB-fold domains IFI16-HIN200A and IFI16-HIN200B can unwind DNA. However, this activity was inhibited for constructs that contained both HIN200 domains.

3.2.2 DNA unwinding activity of IFI16-HIN200A is ATP independent and requires divalent cation Mg^{2+}

In normal conditions, DNA helicase requires the presence of both $MgCl_2$ and ATP to unwind DNA however, RPA can unwind DNA without ATP and magnesium (Treuner *et al.* 1996; Georgaki *et al.* 1992). DNA unwinding activity of IFI16-HIN200A did not require ATP as ATP was not included in the reactions. To find out whether the DNA unwinding activity requires magnesium ion, unwinding assays with IFI16-HIN200A were carried out in several different conditions: 1) absence of both Mg^{2+} and EDTA, 2) presence of Mg^{2+} without EDTA, 3) presence of EDTA without Mg^{2+} , and 4) both Mg^{2+} and EDTA were included in the reactions. Clear unwinding activity was observed in condition 2 only (Figure 3-13). In conditions 1, 3, and 4, 30-nt oligonucleotides were detected in their bound-state after 1 hour incubation at 37 °C although the intensity of free-state oligonucleotides was slightly stronger than the negative control. Nevertheless, removal of Mg^{2+} by either excluding it from reactions or chelating it with EDTA altered the IFI16-HIN200A unwinding activity. Therefore, DNA unwinding activity of IFI16-HIN200A is ATP independent and requires at least Mg^{2+} or any divalent cation, but other cations remain to be tested.

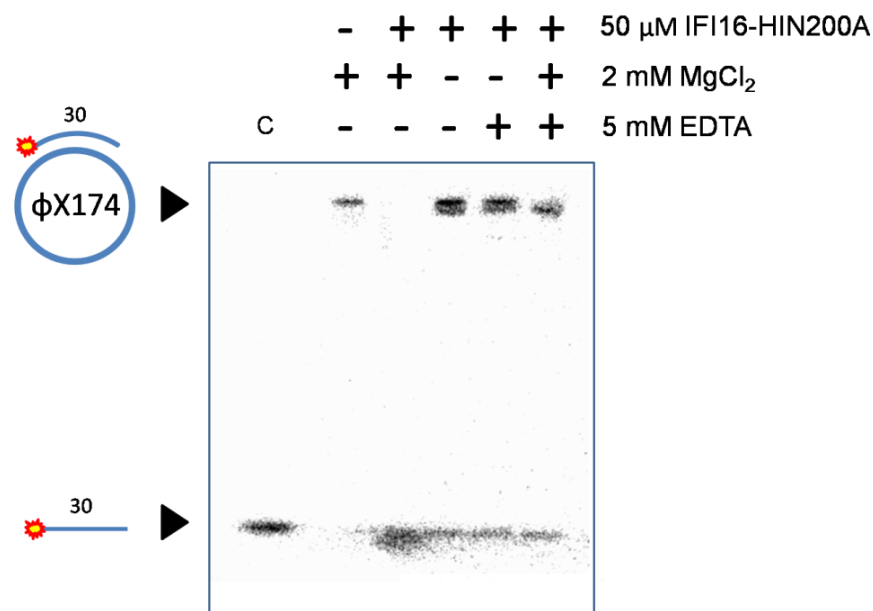


Figure 3-13. IFI16-HIN200A requires magnesium to unwind DNA.

Unwinding assay of IFI16-HIN200A (50 μ M) was performed with 5 nM of Φ X174 partial duplex in 20 mM Tris-HCl, 4% sucrose, 8 mM dTT, 80 μ g/ml BSA (pH 7.5) with or without 2 mM of $MgCl_2$, and was incubated at 37 $^{\circ}$ C for 1 hour before analysis. IFI16-HIN200A unwinds Φ X174 partial duplex in the presence of $MgCl_2$. Without magnesium, IFI16-HIN200A did not unwind the DNA. When 5 mM EDTA was added to chelate the magnesium ions in the buffer, DNA unwinding activity was also diminished. This indicates that DNA unwinding by IFI16-HIN200A was specific to the presence of magnesium ion in the buffer.

3.2.3 Temperature dependence of IFI16-HIN200A unwinding activity

Enzymatic reactions generally operate at an optimal temperature. To confirm that the DNA unwinding activity observed is the result of enzymatic reaction of IFI16-HIN200A and to determine the temperature at which maximum activity is observed, unwinding assays were performed at different temperatures. The DNA unwinding activity of IFI16-HIN200A was performed at 4 $^{\circ}$ C, 25 $^{\circ}$ C, and 37 $^{\circ}$ C. At 4 $^{\circ}$ C, the unwinding measured was minimal (Figure 3-14). More unwinding was detected when assay was done at 25 $^{\circ}$ C. DNA unwinding was

greatly enhanced at 37 °C compared to lower temperatures, indicating that 37 °C was the optimal temperature out of the conditions tested.

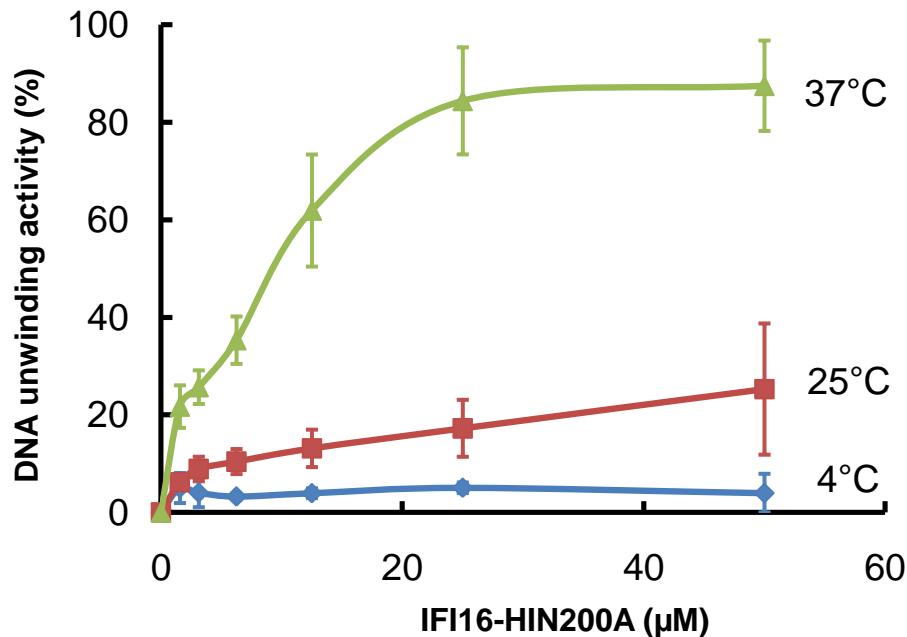


Figure 3-14. DNA unwinding of IFI16-HIN200A is temperature sensitive.

DNA unwinding of IFI16-HIN200A was performed as previously described (Section 2.14) at three different temperatures: 4 °C, 25 °C, and 37 °C. Bands were quantified using Typhoon imager and ImageQuant software. The percentage of DNA unwinding activity was determined by the equation: (radioactivity of free probe)/ (total radioactivity) × 100 for each reaction and was plotted against the concentration of IFI16-HIN200A in μM. DNA unwinding was optimal at 37 °C. Each data point on the graph was an average of values obtained from three independent experiments

3.2.4 A ssDNA segment on the probe does not enhance IFI16-HIN200A unwinding effect

IFI16-HIN200A has a preference to bind ssDNA over dsDNA by approximately 100 fold (Table 3-2). Here I examined whether a ssDNA region on the probe that was not completely complementary to and does not hybridize with the ΦX174 single-stranded plasmid can promote IFI16-HIN200A unwinding

activity. The 30-nt oligonucleotide used in DNA unwinding assay had a sequence that was completely complementary to Φ X174 and ssDNA region of the hybridized substrate was available only from Φ X174, which should be full of supercoils and hairpins. I designed three more oligonucleotides that possessed a ssDNA region within the duplex region. The first oligonucleotide, called bubble, generated a ssDNA region in the middle (Figure 2-3, B). The other two oligonucleotides have a single-stranded poly(dT)₁₅ at the 5' end, and at the 3' end of the duplex (ie. 5' and 3' flap sequence named 5'T and 3'T respectively, Figure 2-3, C and D) (see Table 2-6 for their sequences). Using these oligos for the unwinding assay, it was found that IFI16-HIN200A was able to unwind all of the different DNA substrates with no measurable differences at the level of unwinding compared to the original DNA substrate with the 30-nt oligonucleotide (Figure 3-15). This finding suggested that a 3' or 5' overhang ssDNA end or bubble on the probe does not enhance DNA unwinding activity by IFI16-HIN200A in the complementary region of the probe to Φ X174.

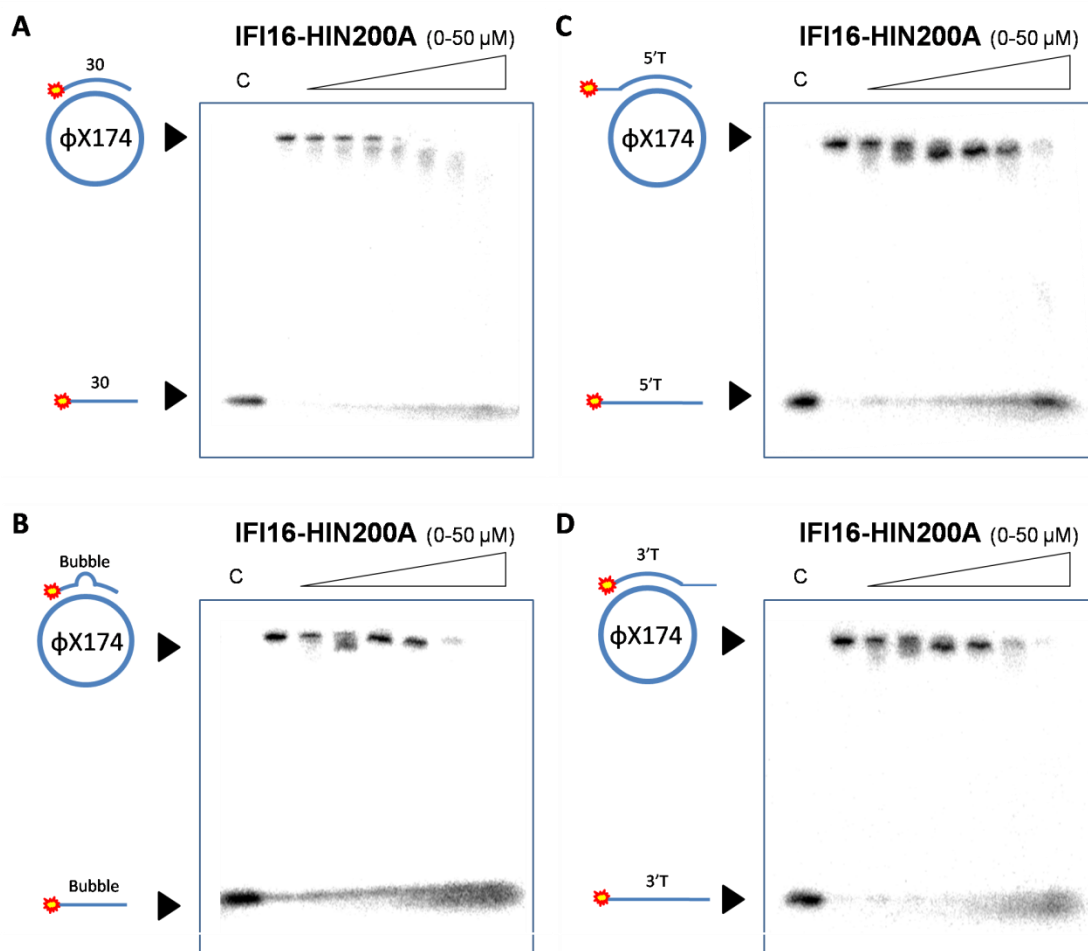


Figure 3-15. IFI16-HIN200A unwinds bubble, 5'T and 3'T like 30-mer.

Unwinding assays for all Φ X174 partial duplexes (5 nM) were performed as described in materials and method (Section 2.14). **A)** 30-mer probe was unwound from Φ X174 by increasing IFI16-HIN200A concentration from 1.56-50 μ M obtained by 2-fold serial dilutions. Other DNA substrates **B)** bubble, **C)** 5'T, and **D)** 3'T were not fully complementary to Φ X174 as shown in the cartoon picture at the top left-hand side of each figures. IFI16-HIN200A was able to unwind all of them. In terms of the concentration required to completely unwind the ssDNA probe from Φ X174, there were no significant difference between different DNA substrates. Therefore, the ssDNA region on the probes did not enhance DNA unwinding activity.

Taken together, IFI16-HIN200A has DNA unwinding activity, like RPA.

However, RPA unwinds DNA such that it does not require Mg^{2+} during the process but IFI16 unwinding activity seemed to be affected tremendously by the

absence of Mg^{2+} . Temperature was also a factor as the activity was the most evident at 37 °C compared to lower temperatures 4 °C and 25 °C.

3.3 IFI16 has endonuclease activity on ssDNA

3.3.1 IFI16-HIN200A cleaves ssDNA

The mobility of free-probe in the DNA unwinding assay was different from expected after it was unwound from Φ X174 by IFI16-HIN200A (Figure 3-9). Two reasons might account for the altered migration of DNA into the gel: 1) The unwound DNA has a different secondary structure which differs in the charge and the overall shape of the molecule causing a difference in migration, or 2) The DNA molecule is cut or modified, causing a change in its charge, shape, or molecular weight. To determine which explanation is correct, an assay was done where increasing concentrations of IFI16-HIN200A were incubated with the 5'- ^{32}P -labelled 30-mer oligonucleotide alone without Φ X174. Reactions were incubated at 37 °C for 1 hour, as for the DNA unwinding assay, and then urea loading buffer was added to denature both protein and DNA. All reactions were boiled at 95 °C for 5 minutes prior to gel electrophoresis. Without protein, the ssDNA probe remained 30 nt in length according to 10 bp DNA ladder. However, the increased IFI16-HIN200A concentration reduced probe length to below 10 nt (Figure 3-16). Since the assay was carried out in denaturing conditions, DNA strands were all linearized and therefore, the mobility shift that was still observable was not due to formation or deformation of intramolecular secondary structures. According to the 10 bp DNA ladder, IFI16-HIN200A reduced the

molecular weight or the length of the 30-mer ssDNA oligonucleotide, suggesting that it has ssDNA cleavage activity.

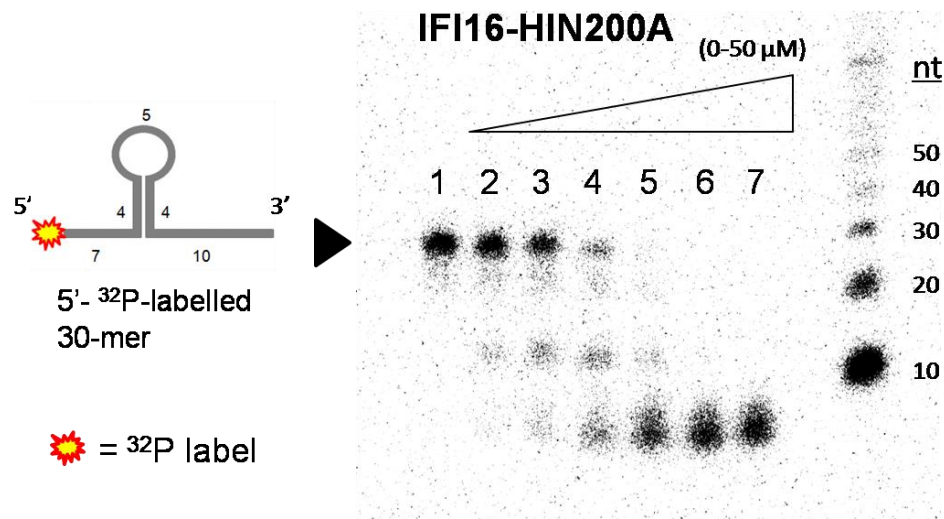


Figure 3-16. IFI16-HIN200A cleaves DNA.

5' end labelled 30-mer (5 nM) were incubated with various IFI16-HIN200A concentrations in 20 mM Tris-HCl, 4% sucrose, 8 mM dTT, 80 $\mu\text{g}/\text{ml}$ BSA, and 2 mM MgCl_2 , pH 7.5, at 37 °C for 1 hour. Reactions were analyzed in a 20% denaturing polyacrylamide gel that contains 20% glycerol. A heat-denatured 10 bp ladder was used as the ssDNA length standard. Lane 1: 30-mer probe without IFI16-HIN200A protein, probe remained uncleaved. Lane 2-7: Increasing IFI16-HIN200A concentration from 1.56-50 μM obtained by 2-fold serial dilutions progressively shortened the 30-mer probe to ~10 nt first and to smaller fragments at higher IFI16-HIN200A concentrations.

3.3.2 Non-sequence specific endonuclease activity of IFI16-HIN200A

According to Figure 3-16, the 30-mer ssDNA was cleaved such that a product of approximately 10 nt was first produced and eventually became less than 10 nt in length at high concentration of IFI16-HIN200A. This two-stage cleavage suggests there are site preferences within the ssDNA that are susceptible to cleavage by IFI16-HIN200A. It also suggests that the cleavage

was not a 5'→3' exonuclease activity because the DNA is 5'-³²P-labelled.

However, it is still possible that the cleavage took place within the phosphate backbone of ssDNA, or at the 3' end of the ssDNA. To distinguish between these possibilities, I carried out cleavage assay with both 5' end-labelled oligonucleotides and 3' end-labelled oligonucleotides: 30-mer ssDNA, 5'T (an oligonucleotide consisting the same sequence as 30-mer with an addition of 15 thymine bases at its 5' end), T₇₀, a 70 nt long poly(dT), and GC-rich oligonucleotides called GC-5 and GC-3 (Section 2.15, Table 2-7). These oligos differ in the base composition and length, which may account for the different patterns of cleavage by IFI16-HIN200A.

First, the 30-mer oligonucleotide was 3' end-radiolabelled and the same cleavage assay was repeated. Similar to the case of 5' end-labelled 30-mer, two products were produced from nucleic-acid cleavage. This was against the explanation of 3'→5' exonuclease activity. Furthermore, the intermediate cleaved product was approximately 15 nt, slightly longer than the one observed for 5' end-labelled oligonucleotide (Figure 3-17). Based on this gel alone, it was difficult to determine the site preference to DNA cleavage. However, I have demonstrated that IFI16-HIN200A does not have either 5'→3' or 3'→5' exonuclease activity on ssDNA, and this strongly indicated that the ssDNA cleavage was an endonuclease activity.

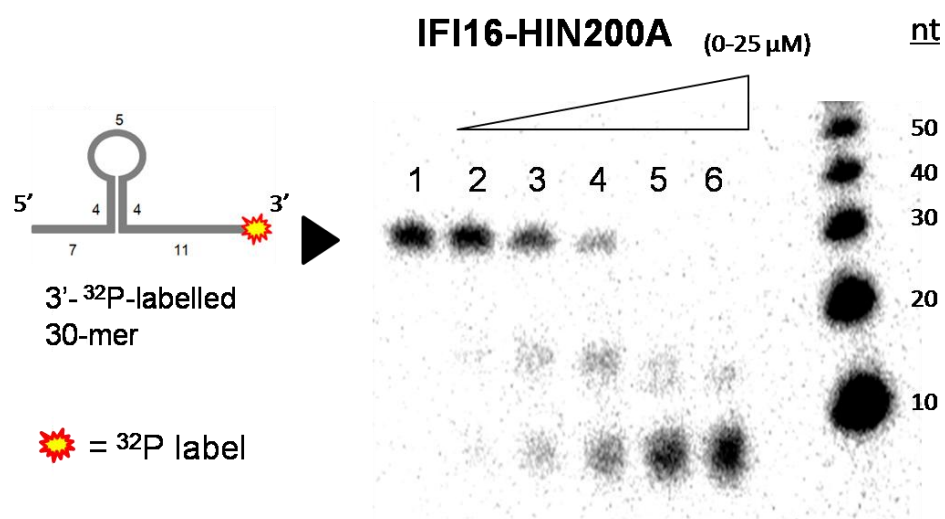


Figure 3-17. Endonuclease activity of IFI16-HIN200A.

Cleavage assay was carried out with IFI16-HIN200A as described previously (Section 2.15) using 3' end labelled 30-mer ssDNA to identify the site of cleavage on the ssDNA. Lane 1: 30-mer probe without IFI16-HIN200A protein, probe remained uncleaved. Lane 2-6: Similar to what was observed for 5' end labelled 30-mer ssDNA, two-step cleaved products was obtained as IFI16-HIN200A concentration was increased by serial 2-fold increments from 1.56-25 μM . The molecular weight of the intermediate products detected here is slightly higher than the 5' end-labelled intermediate products, suggesting that secondary structure on the probe may influence cleavage by IFI16-HIN200A. Since the 3' and 5' end-labelled 30-mer ssDNA produce cleaved fragments of intermediate size during cleavage, we can make the conclusion that IFI16-HIN200A has endonuclease activity

Next, the cleavage pattern on 5'T, T₇₀, and GC-rich oligonucleotides by IFI16-HIN200A were investigated to help identify the position of cleavage on ssDNA. Cleavage of the 5'-labelled 5'T produced an intermediate fragment that was 10 nt long whereas cleavage of the 3'-labelled 5'T, produced an intermediate fragment of approximately 15 nt and 30 nt (Figure 3-18). The extra 30 nt product observed only in 3'-labelled 5'T was interesting. Since 5'T was 45 nt long and has a poly(dT)₁₅ at its 5' end, this suggested that cleavage activity of IFI16-HIN200A has a higher preference at thymine bases. In addition, cleavage on

both 5'-labelled and 3'-labelled T₇₀ showed no intermediate bands produced. The products were one-sized, slightly less than 10 nt long (Figure 3-19). These data indicated that IFI16-HIN200A cleavage activity was not specific for a particular DNA sequence, but had preference either on thymine bases or any polynucleotides without secondary structure.

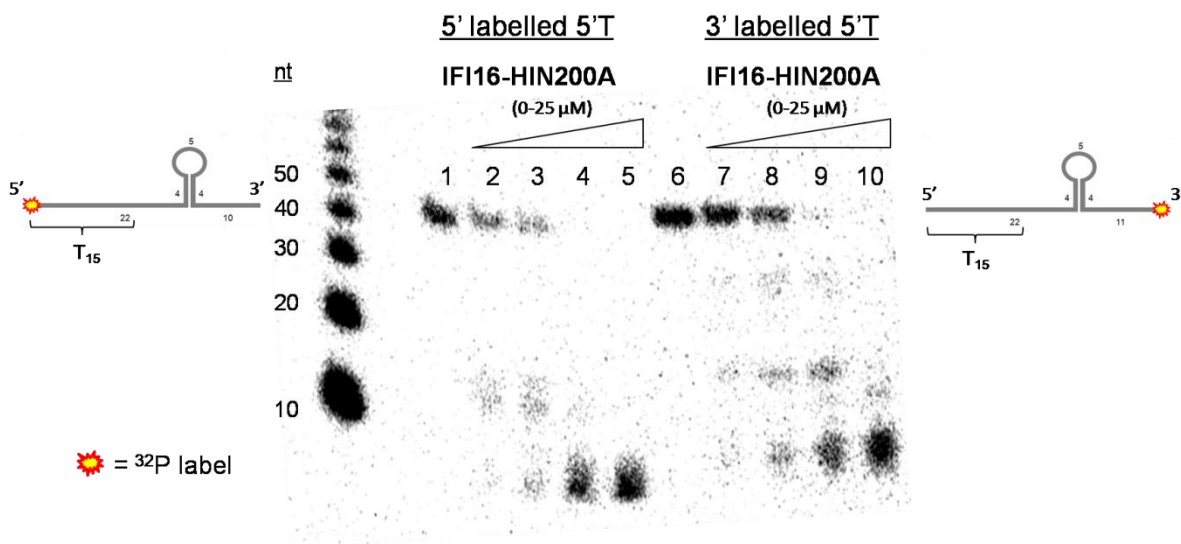


Figure 3-18. Cleavage assay of IFI16-HIN200A on 5' end and 3' end-labelled 5'T oligonucleotide.

IFI16-HIN200A was assayed as previously described (Section 2.15) to cleave another longer oligonucleotide, 5'T (Table 2-7). Reactions were incubated at 37 °C for 1 hour. Lane 1 and 6: 5'T probe without IFI16-HIN200A protein. Lane 2-5: As IFI16-HIN200A concentration was 2-fold increased from 3.1-25 μM, a cleaved product of ~10 nt from 5' end-labelled 5'T was first observed until complete cleavage into lower molecular weight species. Lane 7-10: For 3' end-labelled 5'T, another product of approximately 30 nt long was also produced, indicating the probe were occasionally protected from endonuclease activity of IFI16-HIN200A possibly due to secondary structures on the probe. Secondary structures of 5' end-labelled 5'T and 3' end-labelled 5'T predicted by mfold are shown schematically on the sides (left = 5' labelled, right = 3' labelled, see Appendix D for detail).

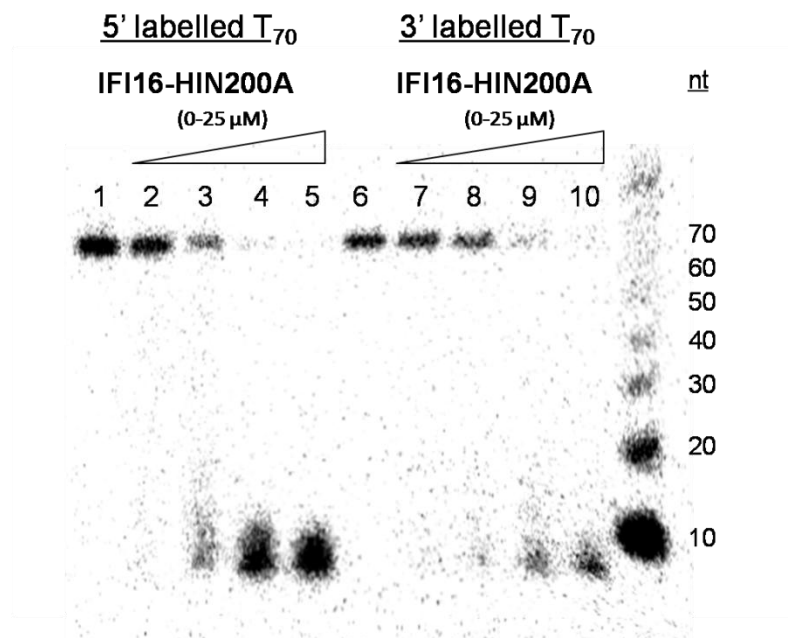


Figure 3-19. Cleavage assay of IFI16-HIN200A on 5' end and 3' end-labelled T_{70} .

Cleavage activity of IFI16-HIN200A on 5 nM poly(dT)₇₀ was performed as described in materials and methods (Section 2.15). Lane 1 and 6: T_{70} probe without IFI16-HIN200A protein. Lane 2-5 and 7-10: Intermediate products were not produced by cleavage for both 5' end-labelled and 3' end-labelled poly(dT)₇₀. This suggests that the cleaved products of two different size obtained with other oligonucleotides (Figures 3-16 to 3-18) are due to the secondary structure of the ssDNA that partially protect ssDNA from cleavage. Poly(dT)₇₀ does not contain secondary structure.

Finally, 5'-labelled GC-5 and GC-3 were investigated. GC-5, which contained very little thymine content, was also cleaved by IFI16-HIN200A in the same manner as that of T_{70} , that is, products less than 10 nt long were generated with intermediate products of slightly longer than 10 nt (Figure 3-20, left). Interestingly, cleavage of GC-3 showed a distinctive pattern. The cleaved product of GC-3 remained at approximately 20 nt in length even at high IFI16-HIN200A concentrations and smaller fragments were present at the highest concentration of 25 μ M of IFI16-HIN200A (Figure 3-20, right). GC-5 and GC-3 were both rich in GC content but they may have different secondary structure,

which may account for the difference in cleavage pattern. Based on their secondary structures (Appendix D) predicted from a web server called mfold (Zuker 2003), the GC-3 oligonucleotide contains a bulky hairpin structure that spans 16 nt at the beginning of 5' end. These data strongly indicate that IFI16-HIN200A cleaves DNA preferably at linearized regions and less efficiently at regions with secondary structures.

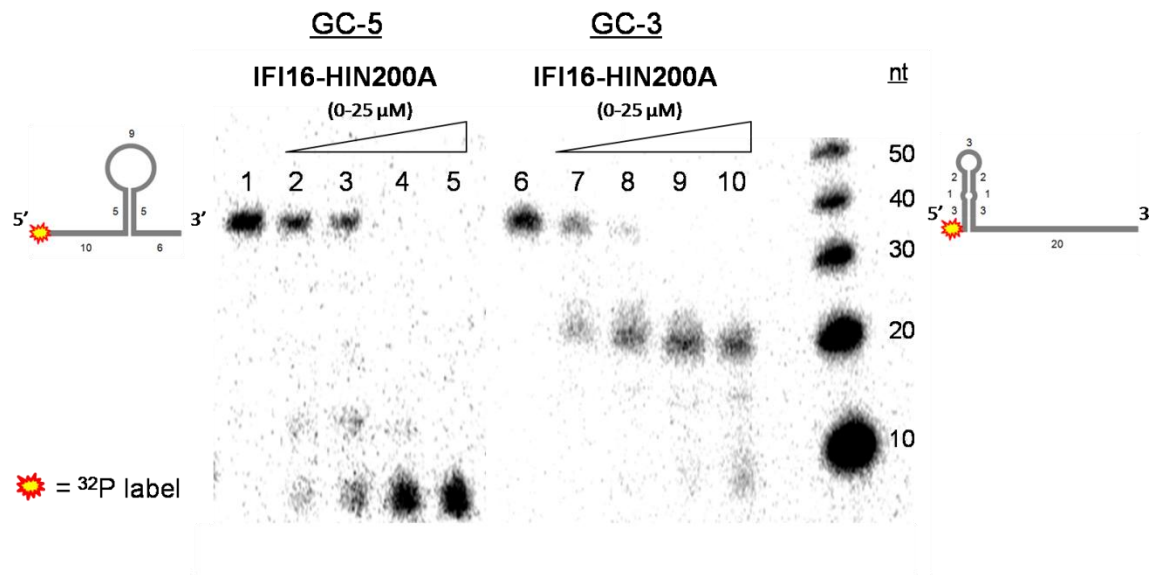


Figure 3-20. Cleavage on GC-5, GC-3 by IFI16-HIN200A showed different pattern of cleaved product.

Cleavage activity of IFI16-HIN200A on 5 nM GC-5 and GC-3 was performed as previously described (Section 2.15). Both GC-5 and GC-3 oligonucleotides were 5'-end-labelled with ^{32}P . Lane 1: GC-5 probe without IFI16-HIN200A protein. Lane 2-5: Increasing IFI16-HIN200A concentration from 1.56-25 μM obtained by 2-fold serial dilutions quickly resulted in cleaved product of 10 nt or less (left). Lane 6: GC-5 probe without IFI16-HIN200A protein. Lane 7-10: In the assay of GC-3, cleaved product remained at around 20 nt long even at higher IFI16-HIN200A concentration until 25 μM where shorter products were observed. Secondary structures of GC-5 and GC-3 predicted by mfold are shown schematically on the sides (left = GC-5, right = GC-3, see Appendix D for details). GC-3 can form a bulky hairpin at the beginning of its 5' end, which may explain how GC-3 was protected from IFI16-HIN200A endonuclease activity.

3.3.3 The DNA endonuclease activity of IFI16-HIN200A is Mg^{2+} and temperature dependent but ATP independent

It was found that DNA unwinding activity of IFI16-HIN200A required Mg^{2+} and was ATP independent. To determine whether the same is true for its DNA cleavage activity, I repeated the same experiment with cleavage assay. The 30-mer oligonucleotides that were found to be cleaved after 1 hour incubation at 37 °C in the presence of IFI16-HIN200A and 2 mM $MgCl_2$ were not cleaved when $MgCl_2$ was excluded or when 5 mM EDTA was added to the system (Figure 3-21). The cleavage activity observed was caused by IFI16-HIN200A, as there was no mobility shift for 30-nt oligonucleotide after 1 hour incubation in the absence of the protein. Therefore, IFI16-HIN200A required magnesium as a cofactor for its endonuclease activity to take place on ssDNA.

50 μ M IFI16-HIN200A	-	-	+	+	+	+	+
1 hr @ 37°C	-	+	-	+	+	+	+
2 mM MgCl ₂	+	+	+	+	-	-	+
5 mM EDTA	-	-	-	-	-	+	+

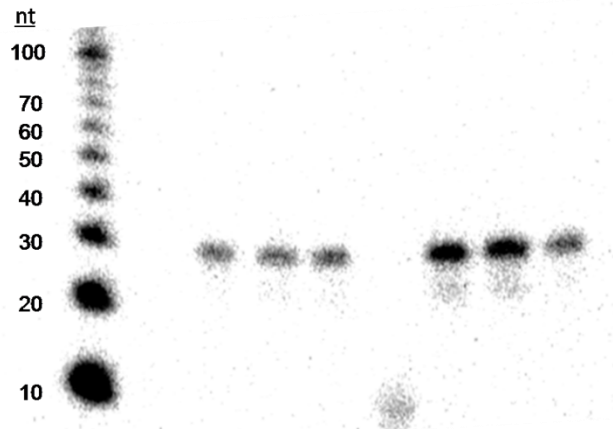


Figure 3-21. Endonuclease activity of IFI16-HIN200A is magnesium dependent.

5 nM 30-mer oligonucleotide probe (Table 2-7) was incubated with 50 μ M of IFI16-HIN200A in 20 mM Tris-HCl, 4% sucrose, 8 mM dTT, 80 μ g/ml BSA, pH 7.5 with or without 2 mM MgCl₂. IFI16-HIN200A cleaved the ssDNA only when the reaction had been incubated at 37 °C and contained 2 mM MgCl₂. Addition of EDTA, which chelated the magnesium ions, inhibited the endonuclease activity. Without protein, 30-mer ssDNA probe remained the same length before and after 1 hr incubation at 37 °C even in the presence of magnesium, showing that the ssDNA was not degraded by itself and the reaction magnesium dependent

To determine if the endonuclease activity of IFI16-HIN200A was temperature sensitive, cleavage assay on 30-mer was performed as before at 4 °C, 25 °C, and 37 °C. At 4 °C, slight cleavage activity of IFI16-HIN200A was observable although it was a lot slower compared to 25 °C (Figure 3-22). Although the activity at 25 °C was quite high, the optimal temperature for IFI16-HIN200A endonuclease activity among the temperature tested was still 37 °C.

In summary, IFI16-HIN200A has endonuclease activity on ssDNA. The activity seemed to be non-sequence specific but has a preference to linearized

region of ssDNA. Similar to its unwinding activity, the endonuclease activity of IFI16-HIN200A required magnesium and was highest at 37 °C compared to 4 °C and 25 °C.

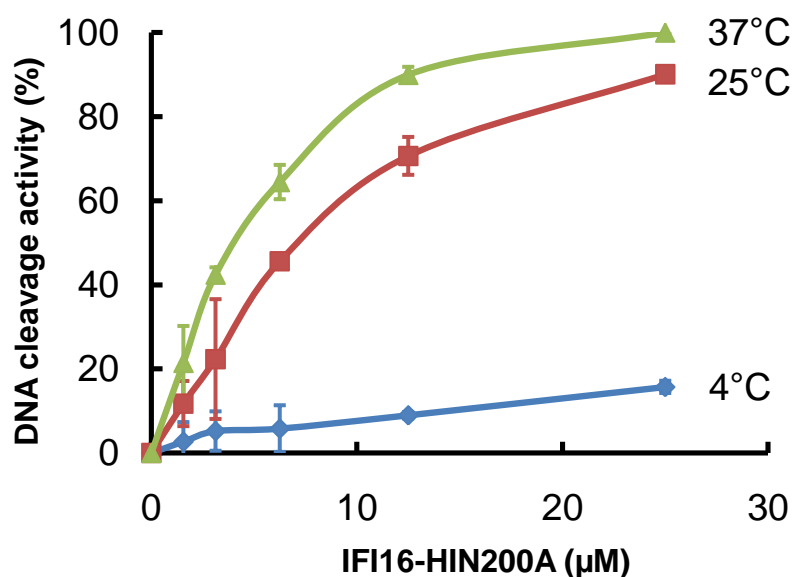
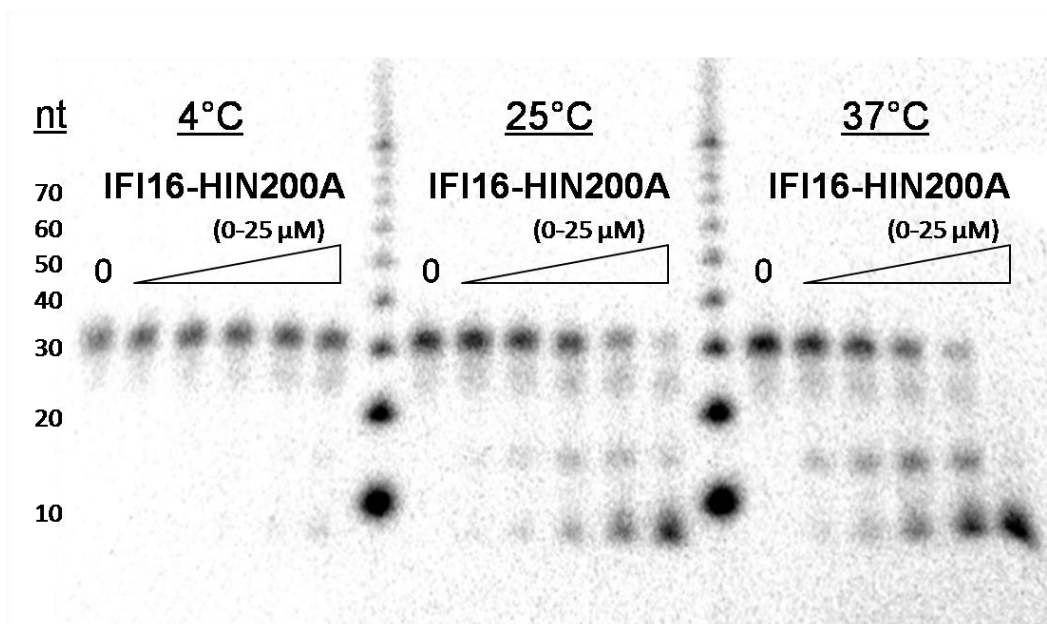


Figure 3-22. Temperature effect on IFI16-HIN200A cleavage activity.

DNA cleavage assay of IFI16-HIN200A was performed as previously described (Section 2.15) at three different temperatures: 4 °C, 25 °C, and 37 °C (Top). The percentage of DNA endonuclease activity was determined by the equation: (radioactivity of free probe)/ (total radioactivity) × 100 for each reaction and was plotted against the concentration of IFI16-HIN200A in μM (Bottom). Like DNA unwinding activity of IFI16-HIN200A, endonuclease activity was optimal at 37 °C. Each data point on the graph is an average of values obtained from three independent experiments

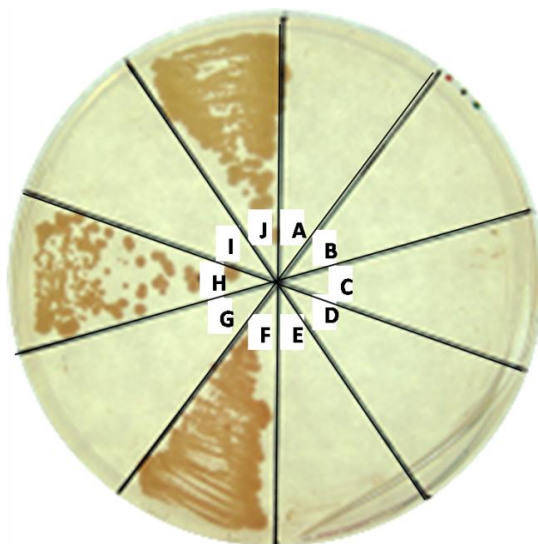
3.4 Studies on full-length IFI16

3.4.1 Yeast 2-hybrid indicates no interaction between IFI16 and BLM

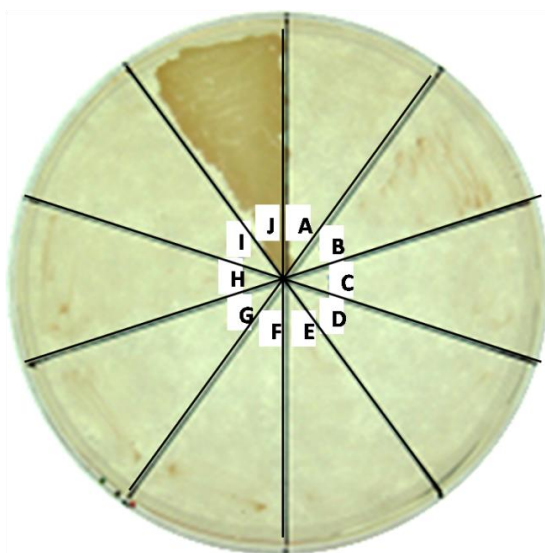
As mentioned in the introduction, RPA physically interacts with BLM and stimulates its DNA unwinding activity (Brosh *et al.* 2000). If IFI16 functions like RPA, it might also have interaction with BLM. In addition, both IFI16 and BLM are parts of the DNA repairing multi-protein complex BASC (Wang *et al.* 2000; Aglipay *et al.* 2003), Section 1.2.4). To investigate the interaction between full-length IFI16 and BLM, *in vivo* yeast 2-hybrid (Y2H) method was performed as described (Section 2.11). Briefly, the full-length sequences of IFI16 and BLM were cloned into both bait (pAS2-1) and prey (pACT2) vectors in frame to their respective fusion proteins DBD (DNA binding domain) and AD (Activation domain). In addition, full-length BRCA1 which has been shown to interact IFI16 (Aglipay *et al.* 2003) was also cloned into both Y2H vectors to serve as a control. These vectors were introduced into yeast strain pJ69-4A by two rounds of PLATE transformation (Section 2.10).

My results showed that yeast transformed with DBD-BLM and AD-IFI16 were able to grow on -leu, -trp, -ade media, indicating the expression of reporter genes (Figure 3-23, top H). However, yeast transformed with DBD-BLM and empty AD also grew (Figure 3-23, top F) suggesting that the earlier observation might be a false positive due to direct interaction between BLM and AD instead of IFI16. Therefore, a reciprocal experiment was done with DBD-IFI16 and AD-BLM. My result indicated no yeast growth on -leu, -trp, -ade media (Figure 3-23, bottom) implying that there was no interaction between IFI16 and BLM.

However, this experiment is inconclusive since one of the positive controls, interaction between BRCA1 and IFI16, showed negative result (Figure 3-23, I)



A	pJ69-4A
B	IFI16FL(AD)
C	BLMFL(BD)
D	BRCA1FL(BD)
E	Empty bait:IFI16FL
F	BLMFL: empty prey
G	BRCA1FL: empty prey
H	BLMFL:IFI16FL
I	BRCA1FL: IFI16FL
J	+control



A	pJ69-4A
B	IFI16FL(BD)
C	BLMFL(AD)
D	BRCA1FL(AD)
E	IFI16FL: Empty prey
F	Empty bait: BLMFL
G	Empty bait: BRCA1FL
H	IFI16FL:BLMFL
I	IFI16FL:BRCA1FL
J	+control

Figure 3-23. Yeast 2-hybrid showed no direct interaction between IFI16 and BLM.

Yeasts were transformed as indicated (right table) and were streaked onto synthetic complete (SC) media lacking leucine, tryptophan, and adenine (-leu, -trp, -ade). Plates were incubated in 30 °C for 2 days before observation. Top: full-length IFI16 gene was fused to activation domain in prey vector whereas BLM and BRCA1 were fused to DNA binding domain of bait vectors. Transformant containing IFI16 and BLM vectors (H) grew but it was a false positive since transformant of BLM and empty prey vector grew. Bottom: full-length IFI16 now fused to DBD BRCA1 and BLM fused to activation domain. No false positives were observed however, transformant containing both IFI16 and BLM did not grow either. Therefore, IFI16 does not interact directly to BLM. + control, SNF1:SNF2 provided by Gabriel Alfaro from Beh Lab.

3.4.2 IFI16 full-length protein can be over-expressed in HEK293T cells

Expression of full-length IFI16 has been tested by Karen Yan, a previous lab member. Specifically, she attempted to express the C isoform of IFI16, the longest IFI16 isoform, using baculovirus system (Yan H, M.Sc Thesis, 2007). She found IFI16c was expressed after the first amplification in the soluble fraction. However, the expression level decreased during the second amplification and no protein could be detected in either supernatant or pellet. It was believed to be due to instability of the protein. Here, I attempted to express the B isoform of IFI16 with yeast, Cos-7, and HEK293T cell expression system. IFI16b is shorter in gene length, smaller in molecular weight (83 kDa compared to 92 kDa), is the most abundant isoform in the cell (Johnstone *et al.* 1998). Yeast, Cos-7, and HEK293T cells were transformed or transfected with the expression plasmid containing full-length IFI16 gene (Section 2.1.1) as described (Section 2.2.2 and 2.2.3). Cells without the expression plasmid served as negative control for IFI16 expression, which was detected by western blotting with polyclonal anti-IFI16-PAAD antibody. It was found that over-expression of full-length IFI16 was successful in HEK293T cells (Figure 3-24). IFI16b gene in pCEP4 was constructed to encode a C-terminal 6xhis tag. Purification of full-length IFI16b gene product will be carried out in the future. Having a full-length IFI16 protein should benefit tremendously in further characterization of this protein.

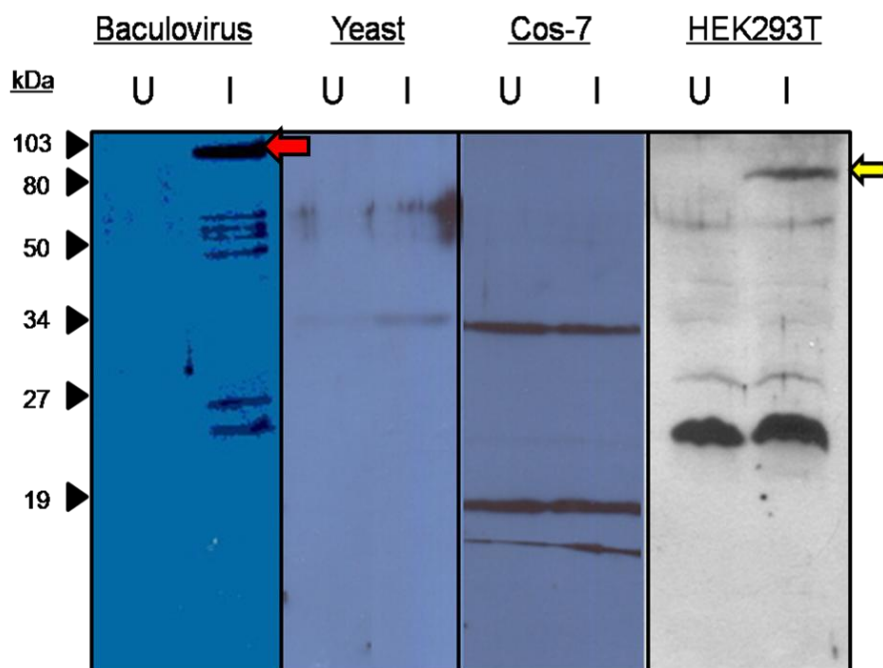


Figure 3-24. Full-length IFI16 expression by different expression systems.

In western blot analyses, Baculovirus expression of IFI16c in T.ni insect cells (taken from Karen Yan's Thesis) was detected using anti-His antibody. Molecular weight of IFI16c is approximately 92 kDa indicated by red arrow. IFI16 detection in Yeast and mammalian expressions of IFI16b was done with anti-IFI16_PAAD antibody. IFI16 was not expressed in yeast strain S288C or mammalian Cos-7. Overexpression of IFI16b was detected in HEK293T. Molecular weight of IFI16b is approximately 83 kDa indicated by yellow arrow. For baculovirus, Cos-7, and HEK293T, U = untransfected, I = transfected. For yeast, U = uninduced, I = induced.

4: DISCUSSION

4.1 IFI16-HIN200 is an RPA-like ssDNA binding protein with unwinding and endonuclease activity

Finding a structural match for a target protein within the Protein Data Bank (PDB) does not always provide a reliable prediction of the protein's function, but may predict structure and suggest possible functional clues (Watson *et al.* 2005). IFI16 and RPA were not only found to have matching OB folds, but they also take part in the same type of cellular pathways: they are both nuclear proteins that interact with nucleic acids, and are involved in cell cycle regulation, transcriptional repression, autoimmune diseases, and DNA repair. Studying the DNA-interacting OB-fold domain of IFI16, in comparison with that of RPA should provide insights to its involvement in these processes. In this thesis, I verified that the HIN200A domain of IFI16 shares many functional properties with the OB fold of RPA but with some distinctions.

The EMSA experiments showed that IFI16-HIN200A binds ssDNA, dsDNA and RNA. We also found that GC-5 and GC-3 (single-stranded oligonucleotides with high GC content) have the highest binding affinity to IFI16-HIN200A, which may indicate that IFI16-HIN200A has a binding preference to GC-rich ssDNA. This is consistent with previous studies showing that the full-length IFI16 has a binding preference to GC-rich dsDNA (Luu *et al.* 1997) and RPA to G-quadruplex (Salas *et al.* 2006).

The ratio of ssDNA/dsDNA binding affinity for IFI16-HIN200A, determined from the fluorescence quenching assay, was $10^2 \sim 10^3$ for the GC rich oligonucleotides but not AT rich oligonucleotides. This ratio is similar to eukaryotic RPA (10^2 - 10^3). For other SSBs it has been reported to be at least 10^6 for gp32, 10^8 for *E.coli* SSB (Gasior *et al.* 2001; Mou *et al.* 2003; Wold 1997). This result was further confirmed by UV cross-linking that detected protein-ssDNA complexes, but never protein-dsDNA complexes, demonstrating further the binding preference of IFI16-HIN200 for ssDNA.

We also showed by melting experiments that IFI16-HIN200A does not destabilize dsDNA since none of the oligo tested lowered the melting temperature. This result is not surprising, since OB-fold proteins have been shown to have various behaviours upon interactions with DNA. One hypothesis is that dsDNA destabilization by SSBs is kinetically regulated (Pant *et al.* 2003). Since in our experimental setting we can only perform steady-state measurements such intermediate kinetic species cannot be dissected at that stage. Further studies are needed on the full-length protein and other domains of IFI16.

Moreover, we have demonstrated through chemical cross-linking and UV cross-linking experiments that IFI16-HIN200A is able to form different oligomers (dimer, trimer and tetramer) in the presence of ssDNA. This property has been shown for the ssDNA-binding OB folds from *Sulfolobus solfataricus* and RPA (Blackwell *et al.* 1994; Bochkarev *et al.* 2004; Lavrik *et al.* 1999). This oligomerization behaviour can further explain the different binding sites observed

in fluorescence quenching assay. These data are also consistent with the crystal structure of the first HIN200 domain of IFI16 (PDB: 2OQ0) which shows that this domain forms homotetramer consisting of 8 OB folds.

Finally, the ssDNA compaction and extension property shown for the OB folds of archaea RPA (Robbins *et al.* 2005) and the prevalent 3'->5' DNA binding polarity observed from 11 of 13 solved OB fold protein co-crystal structures (Theobald *et al.* 2003) were also seen in IFI16 assuming that the ssDNA is positioned the same way as the OB folds of RPA (PDB:1JMC) in the x-ray structure of IFI16-HIN200A (Figure 1-5). The ssDNA recognition by IFI16-HIN200A should imply a reorientation of the two OB folds of one monomer since the structure of the protein-ssDNA of RPA (PDB: 1JMC) does not superimpose with a low root mean square deviation (RMSD) onto the structure of one of the monomer of IFI16-HIN200 (result not shown). This is not surprising since such a reorientation of OB fold is necessary for RPA when binding to ssDNA (Bochkareva *et al.* 2001).

After we confirmed that IFI16-HIN200 has RPA-like ssDNA binding properties, I investigated whether they have similar molecular functions as well. I specifically examined whether IFI16-HIN200A can unwind dsDNA since RPA has been found to have DNA unwinding activity (Georgaki *et al.* 1992). Consistent with RPA, IFI16-HIN200A was found to unwind partial duplex DNA, shown by the adapted DNA unwinding assay. This activity did not require ATP, but was strictly dependent on the presence of the divalent ion magnesium. This strongly indicates that magnesium is required as a cofactor for IFI16-HIN200 to carry out

its DNA unwinding function. Furthermore, DNA unwinding by IFI16-HIN200A is temperature sensitive. Among the temperatures tested, the optimal temperature for the activity is at 37 °C, which is the normal human body temperature.

RPA directly interacts with helicase BLM and the unwinding activity of RPA is important for BLM to unwind longer DNA substrates (Brosh *et al.* 2000). BLM is also one of the components in the multi-protein BASC complex, which contains IFI16 as well (Section 1.2.4). For this reason, direct interaction between IFI16 and BLM was investigated to determine if IFI16 is RPA-like in this aspect and also to verify if IFI16 and BLM have the potential to interact directly with each other within the BASC complex. The yeast two-hybrid method showed that IFI16 and BLM did not interact between each other. A surprising finding for IFI16-HIN200A from the DNA unwinding assay was the ssDNA cleavage activity, which was verified by the DNA cleavage assay. It is also a property that is not found in RPA. DNA cleavage of IFI16-HIN200A was ATP independent and magnesium dependent. With 5'-radiolabelled and 3'-radiolabelled ssDNA, I concluded the activity resembled an endonuclease behaviour and that it cleaves linearized ssDNA better than those with secondary structure. However, the exact position of cleavage could not be determined.

Taken together, IFI16-HIN200 has most of the RPA properties in nucleic-acid binding and unwinding except for dsDNA destabilization and ssDNA cleavage activity. However, since IFI16-PAAD has the ability to destabilize dsDNA, the full-length IFI16 therefore has all of the RPA properties. It is worthwhile to note that other members of the HIN200 family do not share the

RPA-like properties of IFI16-HIN200A (Table 4-1). Effectively, IFI16 is the only member of the family that is involved in DNA damage signalling, DNA repair, and apoptosis (Aglipay *et al.* 2003) suggesting RPA properties along with unwinding and endonuclease activity could be important for IFI16 to function in these processes.

Table 4-1. Summary of DNA-interacting properties of IFI16, MND A, and RPA

	IFI16-PAAD ^A	IFI16-HIN200A	MND A-PAAD ^B	MND A-HIN200 ^B	RPA
Bind	ssDNA, dsDNA, RNA	ssDNA, dsDNA, RNA	ssDNA, dsDNA, RNA	ssDNA, dsDNA	ssDNA, dsDNA, RNA
ssDNA/dsDNA ratio	10 ¹ ~10 ²	10 ² ~10 ³	~1	~1	10 ² ~10 ³
Nucleotide types preference	G,C	G,C	None	None	G,C
Binding mode	2	2	1	1	>2
Oligomerize	Yes	Yes	Yes	Yes	Yes
Compact/extend ssDNA	Compact	Compact extend	No	Compact extend	Compact extend
Destabilize dsDNA	Yes	No	Yes	Yes	Yes
Unwind partial duplex DNA	No*	Yes	No*	No*	Yes
Cleave ssDNA	N/A	Yes	N/A	N/A	N/A

^A taken from M.Sc thesis of Kush Dalal with permission

^B taken from M.Sc thesis of Desmond Lau with permission

* experiments on domains other than IFI16-HIN200 that were done by me

4.2 DNA unwinding and endonuclease activities: new functions to IFI16

The DNA unwinding and cleavage activity of IFI16 is novel and has never been cited in the primary literature. In this study, my data suggested that the cleavage activity of IFI16 was an endonuclease activity that cuts phosphodiester

bonds within sugar backbone of ssDNA. I have also taken into consideration the possibility of it being the result of DNA glycosidic cleavage but that was later ruled out. DNA glycosylase activity removes nitrogenous bases while leaving the sugar phosphate backbone intact. If IFI16-HIN200A has DNA glycosylase activity, the cleavage assay result should be the same for any ssDNA when it is 5'-radiolabelled compared to when it is 3'-radiolabelled. This was not true for the case of the 5'T oligonucleotide. Therefore, endonuclease activity is the most plausible explanation regarding to the mechanism of cleavage. In addition, it was found that IFI16-HIN200A cleaved preferably on linearized DNA and less efficiently at regions with secondary structures. This may be an indication that IFI16 exerts DNA unwinding activity to remove secondary structure prior to DNA endonuclease activity on the linearized DNA.

One point to note is that the IFI16-HIN200A protein used in the unwinding assay was not the most pure as it was only his-tag purified. While the activity observed might have been a result of contamination, IFI16-PAAD (which was also his-tag purified) did not show any signs of DNA unwinding. Since the same cell line was used to over-express both recombinant proteins, this may further support the notion that the unwinding activity was specific to HIN200 domain. In the cleavage assay, IFI16-HIN200A was further purified by cation exchange chromatography and DNA cleavage was still detected. It is very unlikely that the activity is the result of endonucleases of *E.coli* because all *E.coli* endonucleases (I - VII) require different cofactors and process different substrates. They should also be separated from IFI16-HIN200A by cation exchange chromatography as

their pI values are not very close to the 9.07 pI value of IFI16-HIN200A (Appendix B).

In the protein data bank, there is one class of protein like IFI16 that also possesses OB-fold, binds DNA, and cleaves DNA. This protein is called Staphylococcal nuclease (SNase), also known as Micrococcal nuclease (MNase). This enzyme is the extracellular nuclease of *Staphylococcus aureus* that relies on active site residues Glu43, Arg35, and Arg87 for its nuclease activity (Isono *et al.* 2000). It is a non-sequence specific endo-exonuclease that digests single-stranded and double-stranded nucleic acids (Alexander *et al.* 1961), but is more active on single-stranded substrates. Cleavage of DNA or RNA occurs preferentially at AT or AU-rich regions yielding mononucleotides and oligonucleotides with terminal 3'-phosphates. The enzyme activity is strictly dependent on divalent ion Ca^{2+} (Heins *et al.* 1967). These properties are very similar to those found for IFI16-HIN200A.

On the other hand, the endonuclease activity of IFI16-HIN200A was quite different from that of SNases. According to a kinetic study, SNase endonuclease activity was measured to have a k_{cat} value of $7.65 \times 10^2 \text{ s}^{-1}$ and k_{cat}/K_m of $1.66 \times 10^7 \text{ M}^{-1}\text{s}^{-1}$ (Meiss *et al.* 1998). These high values of kinetic parameters correlate to high enzyme efficiency: the higher the k_{cat} (turnover number), the shorter time it is required for one enzyme molecule to convert one substrate molecule to product; the higher the k_{cat}/K_m for an enzyme, the higher its enzyme efficiency. IFI16 seems to behave in the opposite way, where high concentration of protein (μM range) was required to cleave low amount of ssDNA

(nM range), suggesting a low turnover number for its activity. In this study, only magnesium was tested as a cofactor. Since SNase strictly requires calcium as cofactor (Aqvist *et al.* 1989), it is possible that replacing magnesium with calcium as cofactor may enhance enzyme efficiency of IFI16-HIN200A. Furthermore, activity found *in vitro* may not always be the same as it is *in vivo*. Further experiments will be needed to obtain reliable kinetic parameters for the optimal endonuclease activity IFI16-HIN200A.

4.3 IFI16 in transcriptional repression and apoptosis

In general, transcriptional repression can be achieved in several ways. It could be accomplished by transcriptional repressors, which bind either directly to the target gene or indirectly by interacting with other DNA-binding proteins to inhibit gene expression. Transcriptional repressors can also inhibit gene expression by blocking the interaction of the transcriptional activator with other components of the transcription machinery, or preventing the activator from binding to the targeted gene. Moreover, chromatin structure plays a fundamental role in regulation of gene expression (Felsenfeld *et al.* 2003). DNA methylation, on the other hand, can induce a change in chromatin conformation and interferes with transcription by directly blocking transcription activators from binding to the gene (Varriale *et al.* 2010; Ashraf *et al.* 1998). While it has been shown that large GC-rich regions of DNA within the chromatin tend to contain more genes and are characterized by a more expanded chromatin configuration (Bernardi 2007; Saccone *et al.* 2002), the finding of IFI16 having higher affinity to GC-rich

DNA may suggest a method for transcriptional repression similar to DNA methylation at the chromosomal level.

Transcriptional repression of IFI16 has been studied by many groups of researchers (Johnstone *et al.* 1998; Egistelli *et al.* 2009; Courey *et al.* 2001). Specifically, Trapani JA and collaborators have determined that the individual HIN200 domains of IFI16 can actively repress transcription when fused to the GAL4 DNA binding domain and assayed for repression activity using the GAL4-tk-CAT and tk-CAT reporter plasmids. However, expression of the construct containing both HIN200A and HIN200B fused to the GAL4 DNA binding domain did not result in transcriptional repression (Johnstone *et al.* 1998). This observation could not be explained at that time. Interestingly in my study, I have found that the individual IFI16-HIN200 domains, but not the 2HIN, can unwind and cleave DNA, consistent to previous observations. This suggests that IFI16's transcriptional repression function may require DNA unwinding or endonuclease activity from the individual HIN200 domain. This raises the question: how does transcription be regulated via DNA unwinding or endonuclease activities? An example of this could be found in the transcriptional regulation of the oncogene c-Myc, a gene that promotes cell growth. Within the promoter region of c-Myc is a nuclease- hypersensitive element referred as the Nuclease Hypersensitive Element (NHEIII₁) site. This site is characterized by a purine-rich top strand and a pyrimidine-rich bottom strand (Figure 4-1, A). This feature allows the formation of G-quadruplex at the top strand and i-motif at the bottom strand, which is how c-Myc transcription is inhibited (Figure 4-1, B). The formation of G-quadruplex is

induced by negative supercoiling (Brooks *et al.* 2009). C-Myc transcription is activated by a protein called NM23-H2, which can bind to the purine-rich and pyrimidine-rich regions in the NHEIII₁ site and alters the structure of G-quadruplex and i-motif (Brooks *et al.* 2009). This regulation by NM23-H2 involves DNA cleavage activity as well (Postel 1999). Interestingly, a recent study demonstrated that IFI16 and NM23-H2 can simultaneously bind to c-Myc promoter element (Egistelli *et al.* 2009). Another study showed that IFI16 can down-regulate c-Myc mediated telomerase reverse transcriptase transcription (Song *et al.* 2010). These studies suggests that IFI16 may directly repress c-Myc transcription. IFI16 may unwind and cleave DNA within the NHEIII₁ site of c-Myc, induce negative supercoiling and open up the strands. Furthermore, the ability of IFI16 to compact and extend ssDNA may enhance G-quadruplex formation and maintain c-Myc repression. IFI16 promotes cell cycle arrest (Section 1.2.2), and it would make sense that it can repress transcription of c-Myc, which mediates cell growth.

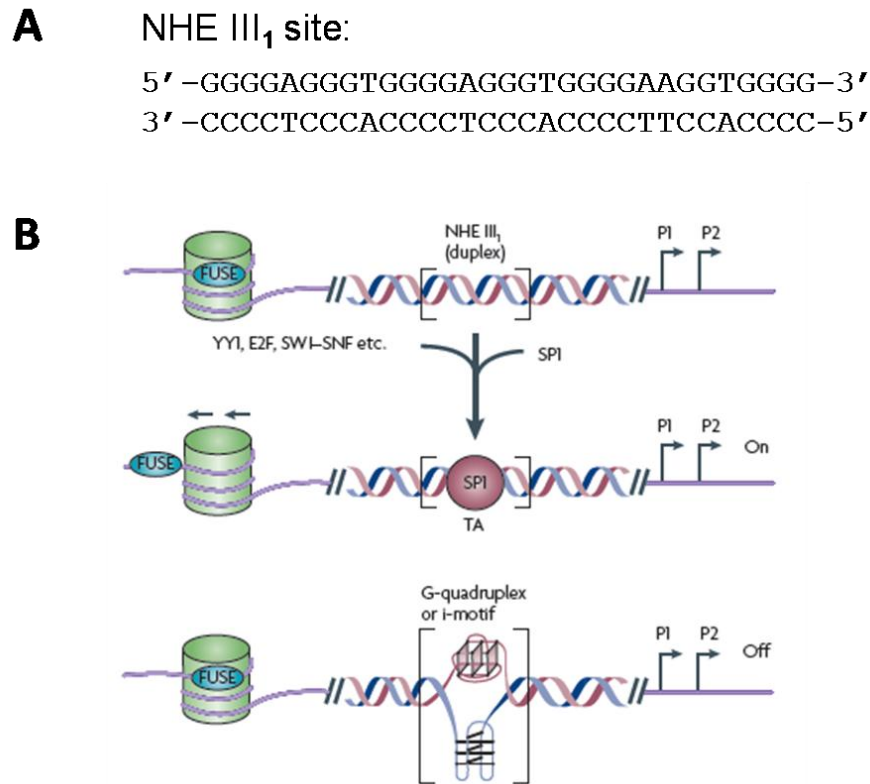


Figure 4-1. Transcriptional regulation of c-Myc (Brooks et al. 2009).

A) Nucleotide sequence of NHEIII₁ site within the c-Myc promoter. Purine-rich sequence on the top strand allows formation of G-quadruplex structure and pyrimidine-rich sequence on the bottom strand favours i-motif formation. **B)** Regulation of c-Myc transcription at the NHEIII₁ site. It can be activated by binding of transcriptional activators such as SP1 at the site. Transcription is turned off when G-quadruplex or i-motif formed within the NHEIII₁ site.

Repressing transcription via unwinding and endonuclease activity of IFI16 will need to be further verified along with how it is regulated. The DNA interacting sites in HIN200 domains could be hindered in the form of 2HIN. Full-length IFI16 that also contains 2HIN may require a mechanism to induce a conformational change that exposes its DNA binding sites for its function, and this could be achieved by binding of the PAAD domain to other proteins.

Nonetheless, IFI16 contains different isoforms of different length that occur in the spacer region between the two HIN200 domains. These different isoforms when produced in the cell may have different functions in regards to unwinding and nuclease activity and this may contribute to the regulation of transcriptional repression.

OB fold-containing micrococcal nuclease has been shown to initiate apoptotic DNA fragmentation (Yin *et al.* 1997), which is one of the best known biochemical hallmarks of programmed cell death besides chromatin condensation. The discovery of endonuclease activity in IFI16-HIN200A is very intriguing because it may suggest that IFI16 may act as an apoptotic endonuclease for degrading chromosomal DNA. There are two levels of DNA fragmentation (Di Filippo *et al.* 2009); one level breaks DNA into high molecular weight (HMW) fragments that range from 50 to 300 kb in length and the second level is further degradation of HMW fragments to low molecular weight (LMW) oligonucleosome-sized fragments. As discussed, GC-rich regions in chromatin are more expanded and therefore, IFI16 may initiate HMW fragmentation by attacking these regions since it has a preference to interact with GC-rich DNA according to this study. The ability of IFI16 to cleave DNA into small fragments was shown in this thesis and suggests that it has a role in LMW fragmentation. This may explain how IFI16 induced p53-mediated apoptosis in MEF cells (Aglipay *et al.* 2003). It is interesting to see in the same study that expression of 2HIN (IFI16 Δ N) lacking endonuclease activity failed to induce apoptosis in p53(+)

MEF cells. Nonetheless, DNA fragmentation by IFI16 will need to be established *in vivo* to confirm its biological relevance.

5: CONCLUSION AND FUTURE DIRECTIONS

In this study, the OB-fold domain of IFI16 was characterized to enhance our understanding of how IFI16 interacts with nucleic acids, which may shed light on determining how IFI16 is involved in transcriptional repression, DNA repair and apoptosis. IFI16 was originally hypothesized to have similar structure and function as the OB-fold domain of RPA. Indeed, it was found that they both have similar DNA binding properties; the ability to bind ssDNA, dsDNA and RNA, the binding preference to GC-rich ssDNA, the binding modes to different DNA sequences, oligomerization upon ssDNA binding, the directionality or the orientation of the bound DNA, and the ability to unwind DNA. IFI16-HIN200A does not destabilize dsDNA like RPA. Interestingly, the PAAD domain of IFI16 has the ability to destabilize dsDNA. In the full-length IFI16, this destabilization domain may compensate for the lack of function in the HIN200.

The DNA unwinding activity of IFI16-HIN200A was confirmed when it successfully unwound the Φ X174 partial duplex DNA. This activity requires magnesium as a cofactor for IFI16-HIN200A, and it is ATP independent, unlike DNA helicases. The activity was the highest at 37 °C compared to 4 °C and 25 °C. Furthermore, DNA unwinding activity was only found in OB-fold containing HIN200A and HIN200B of IFI16 but not the PAAD domain. RPA, which contains OB fold, also has the ability to unwind DNA.

The major contribution of this thesis is the discovery of DNA cleavage activity in IFI16-HIN200A. This new function makes IFI16 distinct from RPA. I have shown that IFI16-HIN200A cuts possibly within the sugar-phosphate backbone of DNA and concluded it has an endonuclease activity. It is magnesium dependent and temperature sensitive. IFI16-HIN200A has a preference to cut linearized ssDNA over those with secondary structures. The activity seemed to be length independent because it can cut longer oligonucleotides up to 70-nt in length.

To explain these results on a biological level, I proposed the possible importance of unwinding and cleavage activities in transcriptional repression, DNA repair and apoptosis. It has been shown that knocking out IFI16 in human cells effectively suppresses transcriptional repression and DNA repair (Aglipay *et al.* 2003; Johnstone *et al.* 1998). The putative role of c-Myc transcriptional repression by IFI16 is interesting to study further. Future experiments could be focused on investigating the ability of IFI16 to induce supercoiling within NHEIII₁ site of c-Myc promoter, and the ability to stabilize G-quadruplex structures. In DNA repair, IFI16 may first bind to sites of DNA damage guided by BASC complex, unwinds dsDNA, and cleaves damaged DNA. I have also proposed that IFI16 may be involved in DNA fragmentation due to its endonuclease activity. It will be worthwhile to show in the future that IFI16 is able to cut chromosomal DNA as well. Apoptotic DNA fragmentation is a crucial biological event to ensure minimal impact to neighbouring tissues *in vivo*. Malfunctions in apoptosis seem to be directly linked to cancer, autoimmunity and other diseases

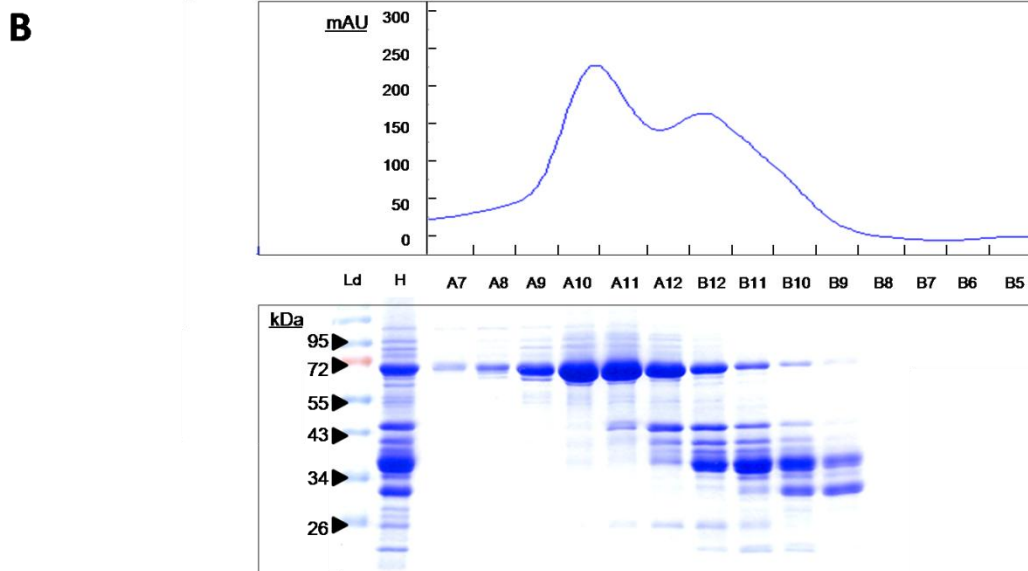
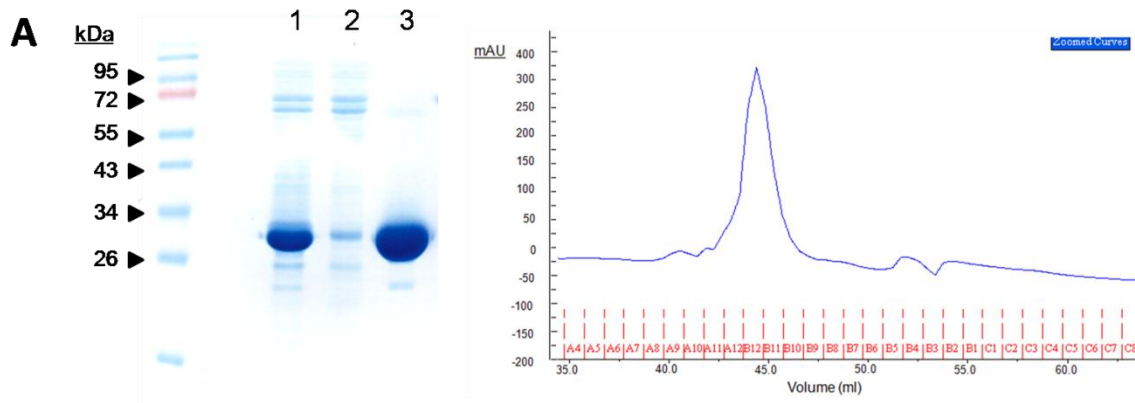
caused by genomic instability. It is therefore no surprise that multiple pathways exist to ensure complete DNA degradation during apoptosis, and IFI16 could be one of these alternative methods. While these aspects require further work, this study opens up a new area of exploration in the role of IFI16 in these processes.

Finally, I have successfully over-expressed full-length IFI16 in human embryonic kidney cells HEK293T. This will be a very useful tool for *in vivo* assays and can be utilized to probe the effect of IFI16 in transcriptional repression, DNA repair, and apoptosis.

APPENDICES

Appendix A: Proteins used in DNA unwinding and DNA cleavage assay

A) Purification of IFI16-HIN200A. IFI16-HIN200A that was used for DNA unwinding assay was His-tag purified (lane 1). In later part of the research, IFI16-HIN200A was further purified by cation exchange purification (lane 3, B12 of right) (right: elution of IFI16-HIN200A by NaCl concentration gradient). Lane 2 is the flowthrough of cation exchange column. **B)** IFI16-2HIN that was used in DNA unwinding assay was originally his-tag purified but it was not pure (lane H). Gel filtration was performed to purify the 65.6 kDa protein. Proteins eluted out from the gel filtration column were collected in fractions (Upper) and fractions A7 to B9 were subjected to SDS-PAGE (Lower). Fraction A10 was used in DNA unwinding assay for IFI16-2HIN.



Appendix B: ProtParam of IFI16-HIN200A

[ExpASY Home page](#) [Site Map](#) [Search ExpASY](#) [Contact us](#) [Proteomics tools](#) [Swiss-Prot](#)
[Search](#) [Swiss-Prot/TrEMBL](#) [for](#) [Go](#) [Clear](#)

ProtParam

User-provided sequence:

```
      10      20      30      40      50      60
MRGSHHHHHH GMASMTGGQQ MGRDLYDDDD KDHPFTTERR NVLQKRPVIV KVLSTTKPFE
      70      80      90     100     110     120
YETPEMEKKI MFHATVATQT QFFHVKVLNT SLKEKFNGKK IIIISDYLEY DSLELVNEES
     130     140     150     160     170     180
TVSEAGPNQT FEVPNKIINR AKETLKIDIL HKQASGNIVY GVFM LHKKT V NQKTTIYEIQ
     190     200     210     220     230
DDRGKMDVVG TGQCHNIPCE EGCKLQLFCF RLRKKNQMSK LISEMHSFIQ IKKKTN
```

References and documentation are available.

^{new} Please note the modified algorithm for extinction coefficient.

Number of amino acids: 236

Molecular weight: 27273.3

Theoretical pI: 9.07

Amino acid composition: [CSV format](#)

Ala (A)	6	2.5%
Arg (R)	9	3.8%
Asn (N)	12	5.1%
Asp (D)	13	5.5%
Cys (C)	3	1.3%
Gln (Q)	13	5.5%
Glu (E)	16	6.8%
Gly (G)	13	5.5%
His (H)	13	5.5%
Ile (I)	17	7.2%
Leu (L)	15	6.4%
Lys (K)	26	11.0%
Met (M)	10	4.2%
Phe (F)	11	4.7%
Pro (P)	8	3.4%
Ser (S)	12	5.1%
Thr (T)	18	7.6%
Trp (W)	0	0.0%
Tyr (Y)	6	2.5%
Val (V)	15	6.4%
Asx (B)	0	0.0%

Glx (Z) 0 0.0%
Xaa (X) 0 0.0%

Total number of negatively charged residues (Asp + Glu): 29
Total number of positively charged residues (Arg + Lys): 35

Atomic composition:

Carbon	C	1204
Hydrogen	H	1922
Nitrogen	N	340
Oxygen	O	356
Sulfur	S	13

Formula: C₁₂₀₄H₁₉₂₂N₃₄₀O₃₅₆S₁₃
Total number of atoms: 3835

Extinction coefficients:

This protein does not contain any Trp residues. Experience shows that this could result in more than 10% error in the computed extinction coefficient.

Extinction coefficients are in units of $M^{-1} \text{ cm}^{-1}$, at 280 nm measured in water.

Ext. coefficient 9065
Abs 0.1% (=1 g/l) 0.332, assuming ALL Cys residues appear as half cystines

Ext. coefficient 8940
Abs 0.1% (=1 g/l) 0.328, assuming NO Cys residues appear as half cystines

Estimated half-life:

The N-terminal of the sequence considered is M (Met).


The estimated half-life is: 30 hours (mammalian reticulocytes, in vitro).
>20 hours (yeast, in vivo).
>10 hours (Escherichia coli, in vivo).

Instability index:

The instability index (II) is computed to be 34.87
This classifies the protein as stable.

Aliphatic index: 73.86

Grand average of hydropathicity (GRAVY): -0.661

 [ExPASy Home page](#) [Site Map](#) [Search ExPASy](#) [Contact us](#) [Proteomics tools](#) [Swiss-Prot](#)
Search [Swiss-Prot/TrEMBL](#) for

Appendix C: Structural Alignment of RPA70AB (PDB: 1JMC) and IFI16-HIN200A (PDB:2OQ0) by EBI-SSM

Query PDB 1jmc:A				Alignment				Target PDB 2oq0:A			
Nres	%res	NSSE	%SSE	Q	P	RMSD	Nalgn	Nres	%res	NSSE	%SSE
238	64	20	45	0.185	-0.00	3.992	152	189	80	17	53
SINGLE STRANDED DNA-BINDING DOMAIN OF HUMAN REPLICATION PROTEIN A BOUND TO SINGLE STRANDED DNA, RPA70 SUBUNIT, RESIDUES 183-420				%seq	Z	NSSE	Ngaps	CRYSTAL STRUCTURE OF THE FIRST HIN-200 DOMAIN OF INTERFERON- INDUCIBLE PROTEIN 16			
				12.5	0.99	9	16				

[view](#)
[download](#)
[view superposed](#)
[view](#)
[download](#)
☐ superpose whole entries

 Viewer: [Jmol](#)

Secondary Structure Alignment

```

1jmc:A  h8ss8h88h88hs--ss8h88h
2oq0:A  -8ss8h88-88hhhss8-88-

```

Query PDB 1jmc:A

```

2|SD  9|A|THR 198 |LYS 206 |  <->
5|SD  7|A|GLU 232 |PHE 238 |  <->
7|SD  5|A|VAL 254 |SER 258 |  <->
8|SD  4|A|THR 261 |ILE 264 |  <->
10|SD 4|A|TYR 276 |THR 279 |  <->
11|SD 4|A|SER 285 |PRO 288 |  <->
16|SD 7|A|VAL 355 |TRP 361 |  <->
18|SD 10|A|VAL 375 |SER 384 |  <->
19|SD 3|A|SER 390 |SER 392 |  <->

```

 SCOP 1: domain [25300](#), family [b.40.4.3](#)

 SCOP 2: domain [25301](#), family [b.40.4.3](#)
[PDB Atlas](#) | [PDB Motif](#) | [OCA](#)
[GeneCensus](#) | [FSSP](#) | [3Dee](#) | [CATH](#) | [PDBsum](#)
[view](#)
[download sequence](#)
[view superposed](#)

Target PDB 2oq0:A

```

1|SD  8|A|VAL 18 |THR 25 |
4|SD  6|A|PHE 52 |VAL 57 |
6|SD  4|A|ILE 72 |SER 75 |
7|SD  3|A|LEU 78 |TYR 80 |
8|SD  4|A|LEU 83 |VAL 86 |
9|SD  4|A|THR 91 |GLU 94 |
15|SD 8|A|GLY 154 |THR 161 |
16|SD 13|A|LYS 174 |LYS 185 |
17|SD 5|A|MSE 188 |ILE 192 |

```

[PDB Atlas](#) | [PDB Motif](#) | [OCA](#)
[GeneCensus](#) | [FSSP](#) | [3Dee](#) | [CATH](#) | [PDBsum](#)
[view](#)
[download sequence](#)

Rotation-translation matrix

(to be applied to the target)

Download the page content

[>>](#) in plain text

[>>](#) in XML

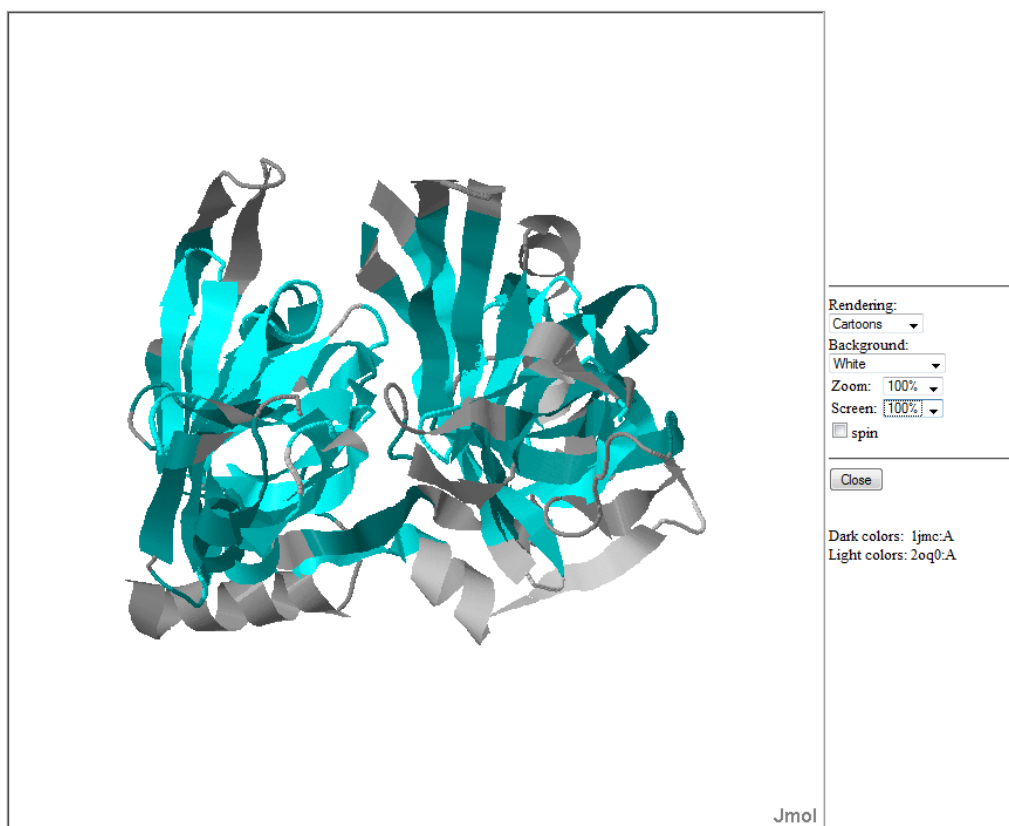
-0.023	-0.192	0.981	×	$\begin{bmatrix} X \\ Y \\ Z \end{bmatrix}$	+	-86.158
-0.895	-0.434	-0.106				72.062
0.446	-0.880	-0.162				99.870

3D Structural alignment

notations

PDB 1jmc:A	SI	Dist. (A)	PDB 2oq0:A	PDB 1jmc:A	SI	Dist. (A)	PDB 2oq0:A	PDB 1jmc:A	SI	Dist. (A)	PDB 2oq0:A
+ A:LYS 193				H+ A:GLN 241				+ A:HIS 294			H- A:ILE 107
- A:VAL 184				H- A:VAL 242				- A:LEU 295			H- A:ILE 108
- A:VAL 185				H+ A:ASP 243				+ A:PRO 296			H+ A:ASN 109
H+ A:PRO 186				H+ A:LYS 244				- A:THR 297			H+ A:ARG 110
H- A:ILE 187				H- A:PHE 245	4.56		H+ A:ASN 59	- A:VAL 298	5.85		H- A:ALA 111
H- A:ALA 188				H- A:PHE 246	1.82		H- A:THR 60	+ A:GLN 299	2.59		H+ A:LYS 112
H- A:SER 189				+ A:PRO 247	4.01		H- A:SER 61	- A:PHE 300	3.84		+ A:GLU 113
H- A:LEU 190				- A:LEU 248			H- A:LEU 62	+ A:ASP 301	4.78		- A:THR 114
- A:THR 191			- A:VAL 12	- A:ILE 249			H+ A:LYS 63	- A:PHE 302	1.04		- A:LEU 115
+ A:PRO 192			- A:LEU 13				H+ A:GLU 64	- A:THR 303			
- A:TYR 193			+ A:GLN 14	+ A:GLU 250			H+ A:LYS 65	H- A:GLY 304	2.96		H+ A:LYS 116
+ A:GLN 194	6.58		+ A:LYS 15	- A:VAL 251	7.89		H- A:PHE 66	H- A:ILE 305	2.31		H- A:ILE 117
- A:SER 195	5.42		+ A:ARG 16	+ A:ASN 252	6.83		H+ A:ASN 67	H+ A:ASP 306	2.47		H+ A:ASP 118
+ A:LYS 196	6.18		+ A:PRO 17	+ A:LYS 253	5.26		- A:GLY 68	H+ A:ASP 307	3.44		H- A:ILE 119
- A:TRP 197	6.02		S- A:VAL 18	S- A:VAL 254	5.00		+ A:LYS 69	H- A:LEU 308	3.96		H- A:LEU 120
S- A:THR 198	4.72		S- A:ILE 19	S- A:TYR 255	5.59		+ A:LYS 70	H+ A:GLU 309	3.74		H+ A:HIS 121
S- A:ILE 199	3.40		S- A:VAL 20				- A:ILE 71	H+ A:ASN 310	2.39		+ A:LYS 122
S- A:CYS 200	3.47		S+ A:LYS 21	S- A:TYR 256	3.11		S- A:ILE 72	H+ A:LYS 311	3.15		+ A:GLN 123
S- A:ALA 201	1.85		S- A:VAL 22	S- A:PHE 257	4.32		S- A:ILE 73				- A:ALA 124
S+ A:ARG 202	4.37		S- A:LEU 23	S- A:SER 258	4.73		S- A:ILE 74	- A:SER 312	2.32		- A:SER 125
S- A:VAL 203	4.70		S- A:SER 24	+ A:LYS 259				+ A:LYS 313			
S- A:THR 204	2.80		S- A:THR 25	- A:GLY 260	2.31		S- A:SER 75	+ A:ASP 314			
S+ A:ASN 205	4.61		- A:THR 26	S- A:THR 261	0.75		+ A:ASP 76	- A:SER 315	1.94		- A:GLY 126
S+ A:LYS 206	6.63		+ A:LYS 27				- A:TYR 77				+ A:ASN 127
- A:SER 207	6.51		+ A:PRO 28	S- A:LEU 262	5.73		S- A:LEU 78	S- A:LEU 316	2.76		S- A:ILE 128
+ A:GLN 208	4.92		S- A:PHE 29	S+ A:LYS 263	5.28		S- A:GLU 79	S- A:VAL 317	3.23		S- A:VAL 129
- A:ILE 209	6.54		S+ A:GLU 30	S- A:ILE 264	6.03		S- A:TYR 80	S+ A:ASP 318	4.32		S- A:TYR 130
			S- A:TYR 31	- A:ALA 265				S- A:ILE 319	5.10		S- A:GLY 131
			S+ A:GLU 32	H+ A:ASN 266				S- A:ILE 320	4.64		S- A:VAL 132
S+ A:ARG 210	4.55		- A:THR 33	H+ A:LYS 267				S- A:GLY 321	3.61		S- A:PHE 133
S- A:THR 211	7.19		+ A:PRO 34	H+ A:GLN 268				S- A:ILE 322	2.84		S- A:MSE 134
S- A:TRP 212			+ A:GLU 37	H- A:PHE 269				S- A:CYS 323	1.54		S- A:LEU 135
S- A:SER 213				H- A:THR 270	3.13		+ A:ASP 81	S+ A:LYS 324	2.05		S+ A:HIS 136
+ A:ASN 214				- A:ALA 271	5.62		- A:SER 82	S- A:SER 325	2.35		S+ A:LYS 137
- A:SER 215				- A:VAL 272				S- A:TYR 326	2.26		S+ A:LYS 138
+ A:ARG 216				+ A:LYS 273				+ A:GLU 327	3.24		S- A:THR 139
- A:GLY 217				+ A:ASN 274				+ A:ASP 328			
S+ A:GLU 218				+ A:ASP 275				- A:ALA 329	2.47		S- A:VAL 140
S- A:GLY 219			S+ A:LYS 38	S- A:TYR 276	6.80		S- A:LEU 83	S- A:THR 330	2.46		+ A:ASN 141
S+ A:LYS 220	7.46		S+ A:LYS 39	S+ A:GLU 277	3.15		S- A:LEU 84	S+ A:LYS 331	1.23		+ A:GLN 142
S- A:LEU 221	8.04		S- A:ILE 40	S- A:MET 278	4.85		S+ A:GLU 85	S- A:ILE 332	2.63		+ A:LYS 143
S- A:PHE 222	4.45		S- A:MSE 41	S- A:THR 279	3.98		S- A:VAL 86	S- A:THR 333			
S- A:SER 223	6.22		S- A:PHE 42	- A:PHE 280	4.64		+ A:ASN 87	S- A:VAL 334			
S- A:LEU 224	3.97		S+ A:HIS 43	+ A:ASN 281	5.68		+ A:GLU 88	+ A:ARG 335			
S+ A:GLU 225	3.12		S- A:ALA 44	+ A:ASN 282	6.33		+ A:GLU 89	- A:SER 336			
S- A:LEU 226	3.09		S- A:THR 45	+ A:GLU 283	7.17		- A:SER 90	+ A:ASN 337			
S- A:VAL 227	1.61		S- A:VAL 46	- A:THR 284	2.13		S- A:THR 91	+ A:ASN 338			
+ A:ASP 228	2.80		S- A:ALA 47	S- A:SER 285	3.88		S- A:VAL 92	S+ A:ARG 339			
+ A:GLU 229	5.56		- A:THR 48	S- A:VAL 286	4.92		S- A:SER 93	S+ A:GLU 340			
- A:SER 230			+ A:GLN 49	S- A:MET 287	5.70		S+ A:GLU 94	S- A:VAL 341			
			- A:THR 50	S+ A:PRO 288			- A:ALA 95	S- A:ALA 342			
			+ A:GLN 51	- A:CYS 289			H- A:GLY 96	S+ A:LYS 343	2.38		S- A:THR 144
- A:GLY 231	1.39		S- A:PHE 52	+ A:GLU 290			H+ A:PRO 97	S+ A:ARG 344	2.21		S- A:THR 145
S+ A:GLU 232	1.55		S- A:PHE 53	+ A:ASP 291			H+ A:ASN 98	S+ A:ASN 345	1.98		S- A:ILE 146
S- A:ILE 233	4.03		S+ A:HIS 54	+ A:ASP 292			H+ A:GLN 99	S- A:ILE 346	2.24		S- A:TYR 147
S+ A:ARG 234	2.40		S- A:VAL 55				H- A:THR 100	S- A:TYR 347	1.82		S+ A:GLU 148
S- A:ALA 235	4.77		S+ A:LYS 56				- A:PHE 101	S- A:LEU 348	1.87		S- A:ILE 149
S- A:THR 236	4.83		S- A:VAL 57				+ A:GLU 102	S- A:MET 349	1.88		S+ A:GLN 150
S- A:ALA 237	4.52		- A:LEU 58				- A:VAL 103	+ A:ASP 350	1.44		S+ A:ASP 151
S- A:PHE 238							H+ A:PRO 104	- A:THR 351	0.95		+ A:ASP 152
H+ A:ASN 239							H+ A:ASN 105	- A:SER 352	2.33		+ A:ARG 153
H+ A:GLU 240				+ A:HIS 293			H+ A:LYS 106	- A:GLY 353			

PDB 1jmc:A	SI	Dist. (A)	PDB 2oq0:A	PDB 1jmc:A	SI	Dist. (A)	PDB 2oq0:A	PDB 1jmc:A	SI	Dist. (A)	PDB 2oq0:A
+ A:LYS 354		2.82	S- A:GLY 154	S- A:ILE 378	::	2.99	S- A:LEU 177	- A:ILE 399	:::	3.84	- A:ILE 199
S- A:VAL 355		2.78	S+ A:LYS 155	S+ A:LYS 379		4.35	S- A:PHE 178	- A:ILE 400		5.30	+ A:GLN 200
S- A:VAL 356		2.06	S A:MSE 156	S- A:GLY 380		3.39	S- A:CYS 179	- A:ALA 401	:	3.03	- A:ILE 201
S A:THR 357		2.23	S+ A:ASP 157	S- A:ALA 381	:	2.63	S- A:PHE 180	+ A:ASN 402	:	4.08	+ A:LYS 202
S- A:ALA 358	:	1.98	S- A:VAL 158	S+ A:ARG 382	:::	2.17	S+ A:ARG 181	+ A:PRO 403			
S A:THR 359		2.33	S- A:VAL 159	S- A:VAL 383	::	2.60	S- A:LEU 182	+ A:ASP 404			
S- A:LEU 360	*	2.54	S- A:GLY 160	S- A:SER 384		3.48	S+ A:ARG 183	H- A:ILE 405			
S A:TRP 361		2.56	S- A:THR 161				S+ A:LYS 184	H+ A:PRO 406			
H- A:GLY 362	:::	3.16	- A:GLY 162	+ A:ASP 385	:	2.77	S+ A:LYS 185	H+ A:GLU 407			
H+ A:GLU 363	::	3.80	+ A:GLN 163	- A:PHE 386		5.84	+ A:ASN 186	H- A:ALA 408			
H+ A:ASP 364		3.50	- A:CYS 164	- A:GLY 387				H A:TYR 409			
H- A:ALA 365		2.73	+ A:HIS 165	- A:GLY 388		4.33	+ A:GLN 187	H+ A:LYS 410			
H+ A:ASP 366	::	2.51	+ A:ASN 166	+ A:ARG 389		2.51	S A:MSE 188	H- A:LEU 411			
H+ A:LYS 367		2.77	- A:ILE 167				S- A:SER 189	H+ A:ARG 412			
H- A:PHE 368		2.38	+ A:PRO 168	S A:SER 390		2.80	S+ A:LYS 190	H- A:GLY 413			
+ A:ASP 369		1.30	- A:CYS 169	S- A:LEU 391	:::	2.58	S- A:LEU 191	H- A:TRP 414			
- A:GLY 370		1.61	+ A:GLU 170	S A:SER 392		2.25	S- A:ILE 192	H- A:PHE 415			
- A:SER 371		3.89	+ A:GLU 171	- A:VAL 393				H+ A:ASP 416			
+ A:ARG 372				- A:LEU 394		1.78	- A:SER 193	H- A:ALA 417			
+ A:GLN 373		1.01	- A:GLY 172				+ A:GLU 194	+ A:GLU 418			
+ A:PRO 374		2.44	+ A:ASP 173	- A:SER 395		5.18	A:MSE 195	- A:GLY 419			
S- A:VAL 375		3.27	S+ A:LYS 174	- A:SER 396	*	3.20	+ A:HIS 196	+ A:GLN 420			
S- A:LEU 376	:::	3.19	S- A:LEU 175	- A:SER 397	:::	3.59	- A:SER 197				
S- A:ALA 377		4.47	S+ A:GLN 176	- A:THR 398		4.67	- A:PHE 198				



Appendix D: Secondary structure prediction on oligonucleotides using mfold

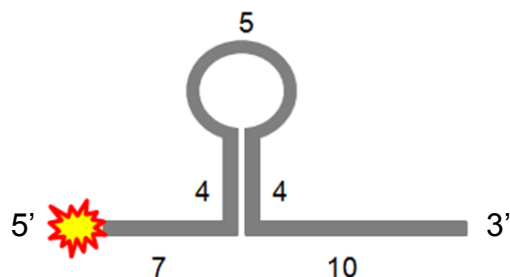
5'-end-radiolabelled 30-mer:

Sequence: *5'-AAGTAAGAGCTTCTCGAGCTGCGCAAGGAT-3'

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 1 Folding bases 1 to 30

dG = -2.45 dH = -29.60 dS = -87.54 T_m = 65.0 °C



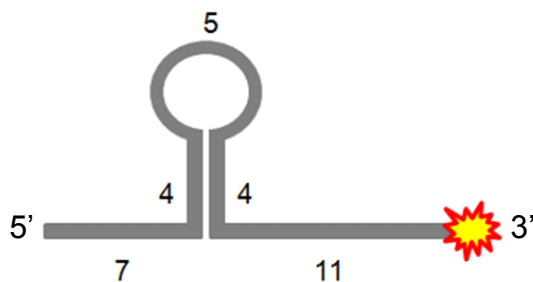
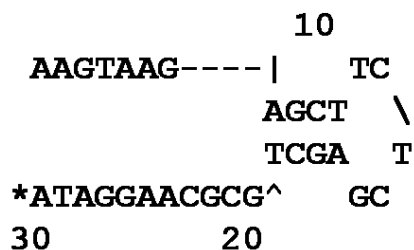
3'-end-radiolabelled 30-mer:

Sequence: 5'-AAGTAAGAGCTTCTCGAGCTGCGCAAGGATA-3'*

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 1 Folding bases 1 to 31

dG = -2.45 dH = -29.60 dS = -87.54 T_m = 65.0 °C



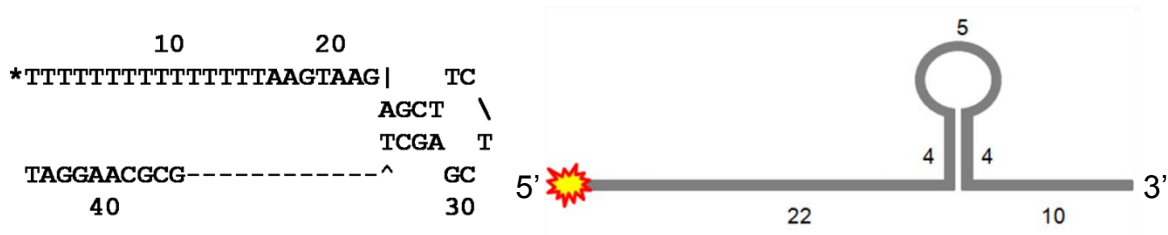
5'end-radiolabelled 5'T:

Sequence: *5'-TTTTTTTTTTTTTTTAAAGTAAGAGCTTCTCGAGCTGCGCAAGGAT-3'

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 1 Folding bases 1 to 45

dG = -2.45 dH = -29.60 dS = -87.54 T_m = 65.0 °C



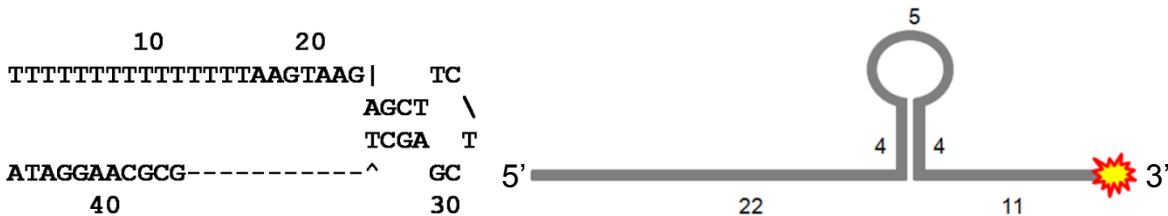
3'end-radiolabelled 5'T:

Sequence: 5'-TTTTTTTTTTTTTTTAAAGTAAGAGCTTCTCGAGCTGCGCAAGGATA-3'*

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 1 Folding bases 1 to 46

dG = -2.45 dH = -29.60 dS = -87.54 T_m = 65.0 °C



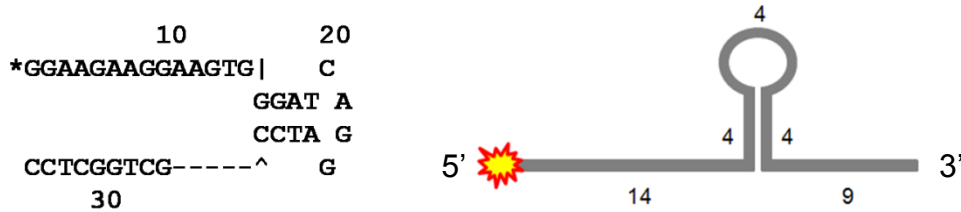
5'end-radiolabelled GC-5 (2 possible structures):

Sequence: *5'- GGAAGAAGGAAGTGGGATCAGGATCCGCTGGCTCC-3'

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 1 Folding bases 1 to 35

dG = -1.45 dH = -36.50 dS = -113.01 T_m = 49.8 °C

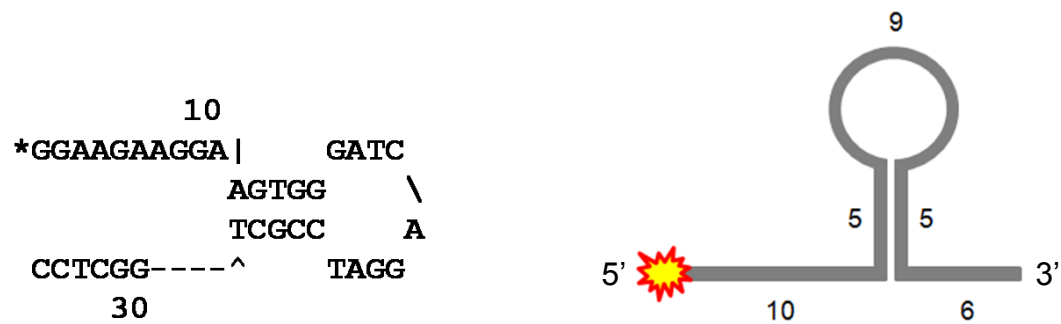


OR

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 2 Folding bases 1 to 35

dG = -0.74 dH = -27.40 dS = -85.96 T_m = 45.6 °C



5'end-radiolabelled GC-3 (2 possible structures):

Sequence: *5'-GGAGCCAGCGGATCCTGATCCCACTTCCTTCTTCC-3'

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 1 Folding bases 1 to 35

dG = -1.45 dH = -36.00 dS = -111.40 T_m = 50.0 °C

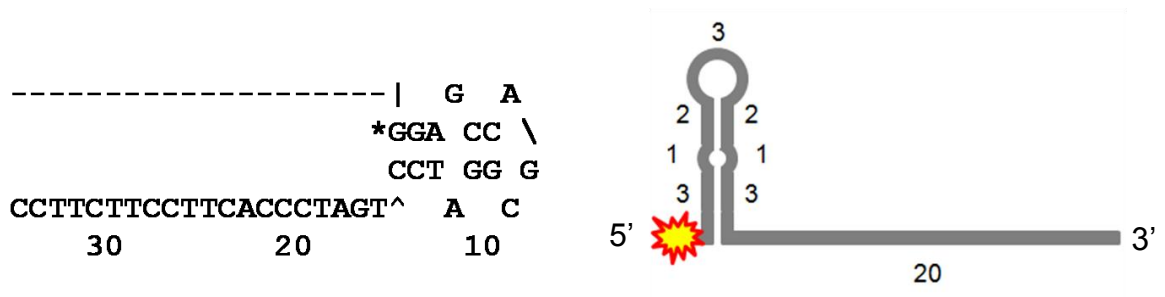


OR

Linear DNA folding at 37 °C. [Na⁺] = 50 mM, [Mg⁺⁺] = 2 mM.

Structure 2 Folding bases 1 to 35 of 10Mar15-04-55-55

dG = -1.24 dH = -30.60 dS = -94.66 T_m = 50.1 °C



Appendix E: Vector plasmids

Yeast 2-hybrid bait vector pAS2-1(Clontech):

BD Biosciences Clontech - Technical Info: pAS2-1 DNA-BD Vector

5/6/05 11:59 AM



BD Biosciences
Clontech
Discovery Labware
Immunocytometry Systems
Pharmingen

Page 1 of 2

Go

ADVANCED SEARCH

SITEMAP

Home About Products Careers How to Order Events

BD Biosciences > Clontech

Privacy | Terms & Conditions

Clontech

Products

Online Catalog
Product Families
Custom Services
Technologies
New Products
Ordering
Sales Reps
Special Offers

Support

User Manuals
Vector Information
Bioinformatics
Citations
Application Notes
FAQs
Email Tech Support

Literature

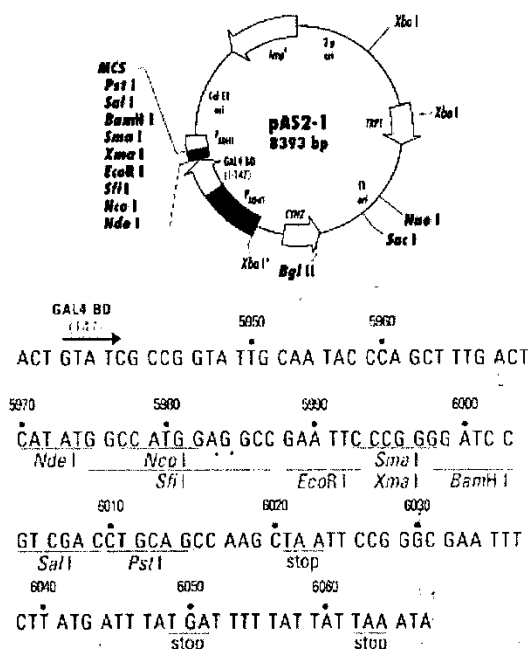
Printed Literature
Brochures
Clontechiques
Order Forms
Product Lists

About

About Us
Contact Us
Careers
Press Releases

pAS2-1 DNA-Binding Domain

This product has been discontinued. This information is being provided as a courtesy. To the best of our knowledge, it was correct at the time the product was discontinued, but has NOT been updated since then.



Annotated Map/MCS (pdf)
Sequence (text)
Sequence (pdf)
Restriction Digest (pdf)
Description
Feature Locations
Primer Sites
Propagation in *E. coli*
Propagation in *S. cerevisiae*
References
License Statement

< Product Quick Links:

Vector Category Links: >

Restriction Map and Multiple Cloning Site of pAS2-1. (Unique restriction sites are in color or bold.) The Xba I site at 4763 bp is methylated in the DNA provided by CLONTECH. If you wish to digest the vector with Xba I at this site, you will need to transform the vector into a *dam*⁻ host and make fresh DNA. The Sfi I and Sma I sites in the MCS tend to compress during sequencing.

Note: The vector sequence file has been compiled from information in the sequence database, published literature, and other sources, together with partial sequences obtained by CLONTECH. This vector has not been completely sequenced.

Description

pAS2-1 generates a fusion of the GAL4 DNA-BD (amino acids 1-147) and a protein of interest cloned into the MCS in the correct orientation and reading frame. pAS2-1 was created from pAS2 (pAS1^{CYH2} in 1) by removing the HA epitope tag and converting a.a. 149 from Glu to Val. This completely eliminates the autonomous activation activity of pAS2 (assayed in Y187 using the *lacZ* reporter); 3). pAS2-1 carries the wild-type yeast *CYH2*, which confers sensitivity to cycloheximide in transformed cells. The hybrid protein is expressed at high levels in yeast host cells from the constitutive *ADH1* promoter (P); transcription is terminated at the *ADH1* transcription termination signal (T) and targeted to the yeast nucleus by nuclear localization sequences (2). pAS2-1 is a shuttle vector that replicates autonomously in both *E. coli* and *S. cerevisiae*, and carries the *bla* gene, which confers ampicillin resistance in *E. coli*. pAS2-1 also contains the *TRP1* nutritional gene that

Yeast 2-hybrid prey vector pACT2 (Clontech):

BD Biosciences Clontech - Technical Info: pACT2 Vector

5/6/05 11:58 AM



BD Biosciences
Clontech
Discovery Labware
Immunocytometry Systems
Pharmingen

[Home](#) [About](#) [Products](#) [Careers](#) [How to Order](#) [Events](#)

[Privacy](#) | [Terms & Conditions](#)

BD Biosciences > Clontech

Clontech

Products

[Online Catalog](#)
[Product Families](#)
[Custom Services](#)
[Technologies](#)
[New Products](#)
[Ordering](#)
[Sales Reps](#)
[Special Offers](#)

Support

[User Manuals](#)
[Vector Information](#)
[Bioinformatics](#)
[Citations](#)
[Application Notes](#)
[FAQs](#)
[Email Tech Support](#)

Literature

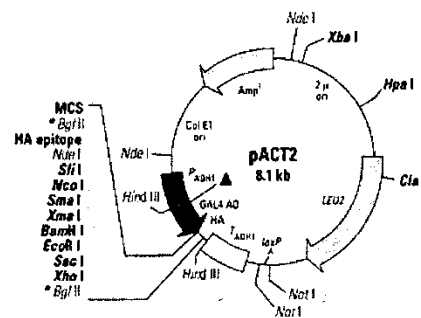
[Printed Literature](#)
[Brochures](#)
[Clontechnotes](#)
[Order Forms](#)
[Product Lists](#)

About

[About Us](#)
[Contact Us](#)
[Careers](#)
[Press Releases](#)

pACT2 AD

This product has been discontinued. This information is being provided as a courtesy. To the best of our knowledge, it was correct at the time the product was discontinued, but has NOT been updated since then.



5085 5070 5060 5050
GAL4 AD (881)
AAA AAA GAG ATC TGT ATG GCT TAC CCA TAC GAT GTT CCA GAT TAC GCT
5040 5030 5020 5010 5000
AGC TTG GGT GGT CAT ATG GCC ATG GAG GCC CCG GGG ATC CGA ATT C
4990 4980 4970 4960
GA GCT CGA GAG ATC TAT GAA TCG TAG ATA CTG AAA AAC
SacI BglII stop stop stop
XhoI

Restriction Map and Multiple Cloning Site of pACT2. (Unique restriction sites are in color or bold.) The BglII sites can be used as a unique site; however, the HA epitope will be deleted.

Note: The vector sequence file has been compiled from information in the sequence database, published literature, and other sources, together with partial sequences obtained by CLONTECH. This vector has not been completely sequenced.

Description

pACT2 generates a fusion of the GAL4 AD (amino acids 788-881), a HA epitope tag, and a protein of interest (or protein encoded by a cDNA in a fusion library) cloned into the MCS in the correct orientation and reading frame. pACT2, which is derived from pACT (1), contains a unique EcoRI site in the MCS. The hybrid protein is expressed at high levels in yeast host cells from the constitutive ADHI promoter (P); transcription is terminated at the ADHI transcription termination signal (T). The protein is targeted to the yeast nucleus by the nuclear localization sequence from SV40 T-antigen which has been cloned into the 5' end of the GAL4 AD sequence (2). pACT2 is a shuttle vector that replicates autonomously in both *E. coli* and *S. cerevisiae* and carries the *bla* gene, which confers ampicillin resistance in *E. coli*. pACT2 also contains the *LEU2* nutritional gene that allows yeast auxotrophs to grow on limiting synthetic media. Transformants with AD/library plasmids can be selected by complementation by the *LEU2* gene by using an *E. coli* strain that carries a *leuB* mutation (e.g., HB101).

Note: The SfiI and SmaI sites in the MCS tend to compress during sequencing.

Product Quick Links:

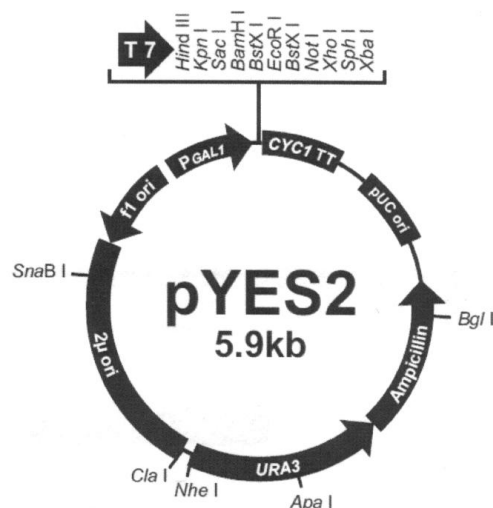
Vector Category Links:

[Annotated Map/MCS \(pdf\)](#)
[Sequence \(text\)](#)
[Sequence \(pdf\)](#)
[Restriction Digest \(pdf\)](#)
[Description](#)
[Feature Locations](#)
[Primer Sites](#)
[Propagation](#)
[License Statement](#)

Yeast expression vector pYES2 (Invitrogen):

Comments for pYES2:
5856 nucleotides

GAL1 promoter: bases 1-451
T7 promoter/priming site: bases 475-494
Multiple cloning site: bases 501-600
CYC1 transcription terminator: bases 608-856
pUC origin: bases 1038-1711
Ampicillin resistance gene: bases 1856-2716 (C)
URA3 gene: bases 2734-3841 (C)
2 micron (μ) origin: bases 3845-5316
f1 origin: bases 5384-5839 (C)
(C) = complementary strand



```

1  5' end of GAL1 promoter
   GAL4 binding site
ACGGATTAGA AGCCGCCGAG CGGGTGACAG CCTCCGAAG GAAGACTCTC CTCCGTGCGT

61  CCTCGTCTTC ACCGGTCGCG TTCCTGAAAC GCAGATGTGC CTCGCGCCGC ACTGTCCGA
      GAL4 binding site

121  ACAATAAAGA TTCTACAATA CTAGCTTTTA TGGTTATGAA GAGGAAAAAT TGGCAGTAAC

181  CTGGCCCCAC AAACCTTCAA ATGAACGAAT CAAATTAACA ACCATAGGAT GATAATGCGA

241  TTAGTTTTTT AGCCTTATTT CTGGGGTAAT TAATCAGCGA AGCGATGATT TTTGATCTAT

301  TAACAGATAT ATAAATGCAA AACTGCATA ACCACTTTAA CTAATACTTT CAACATTTTC
      TATA box

361  GGTITGTATT ACTTCTTATT CAAATGTAAT AAAAGTATCA ACAAAAAATT GTTAATATAC
      start of transcription

421  CTCTATACTT TAACGTCAAG GAGAAAAAAC CCCGGATCGG ACTACTAGCA GCTGTAATAC
      3' end of GAL1 promoter

481  GACTCACTAT AGGGAATATT AAGCTTGGTA CCGAGCTCGG ATCCACTAGT AACGGCCGCC
      T7 promoter/priming site      Hind III      Kpn I      Sac I      BamH I

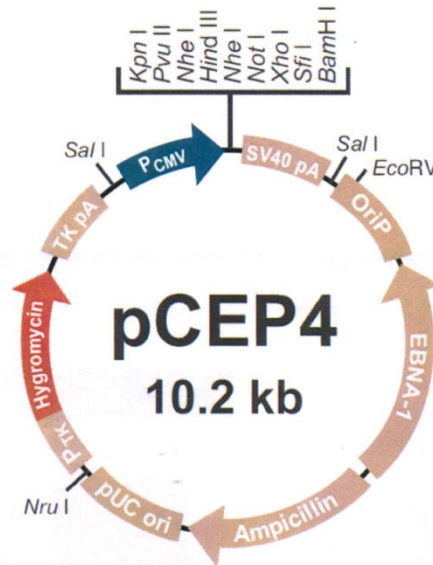
541  AGTGTGCTGG AATTCTGCAG ATATCCATCA CACTGGCGGC CGCTCGAGCA TGCATCTAGA
      BstX I*      Eco RI      BsaB I      BstX I*      Not I      Xho I      Sph I      Xba I

601  GGGCCGCATC ATGTAATTAG TTATGTCACG CTTACATTCA CGCCCTCCCC CCACATCCGC
      5' end of CYC1 transcription terminator
  
```


Mammalian expression vector pCEP4 (Invitrogen):

Comments for pCEP4:
10186 nucleotides

CMV promoter: bases 1-588
Multiple cloning site: bases 619-676
SV40 polyadenylation signal: bases 685-926
OriP: bases 1344-3319
EBNA-1 gene (complementary strand): bases 3620-5545
Ampicillin resistance gene: bases 6171-7031
pUC origin: bases 7040-7815
TK promoter: bases 8183-8345
Hygromycin resistance gene: bases 8409-9419
TK polyadenylation signal: bases 9431-9702



```

1   GTTGACATTG ATTATTGACT AGTTATTAAT AGTAATCAAT TACGGGGTCA TTAGTTCATA
    — enhancer region (5' end)
61  GGGCATATAT GGAGTTCCGC GTTACATAAC TTACGGTAAA TGGCCCGCCT GGCTGACCGC
121 CCAACGACCC CCGCCCATTG ACGTCAATAA TGACGTATGT TCCCATAGTA ACGCCAATAG
181 GGACTTTCCA TTGACGTCAA TGGGTGGAST ATTTACGGTA AACTGCCCCAC TTGGCAGTAC
241 ATCAACTGTA TCATATGCCA ACTCCGCCCC CTATTGACGT CAATGACCGT AAATGGCCCC
301 CCTGGCATTG TGGCCAGTAC ATGACCTTAC GGGACTTTCC TACTTGGCAG TACATCTACG
361 TATTAGTCAT CGCTATTACC ATGGTGATGC GGTTTTGGCA GTACACCAAT GGGCGTGGAT
    AP1
421 AGGCGTTTGA CTCACGGGGA TTTCCAAGTC TCCACCCCAT TGACGTCAAT GGGAGTTTGT
    enhancer region (3' end)
481 TTTGGCACCA AAATCAACGG GACTTTCCAA AATGTCGTAA TAACCCCGCC CCGTTGACGC
    CAAT          TATA          pCEP Forward primer
541 AAATGGGCGG TAGGCGTGTA CGGTGGGAGG TCTATATAAG CAGAGCTCGT TTAGTGAACC
    Transcription start
601 GTCAGATCTC TAGAAGCTGG GTACCAGCTG CTAGCAAGCT TGCTAGCGGC CGCTCGAGGC
    Sfi I   BamH I
661 CGGCAAGGCC GGATCCAGAC ATGATAAGAT ACATTGATGA GTTTGGACAA ACCACAATA
    EBV Reverse primer
  
```

Appendix F: Synthetic Complete (SC) and Drop-out media

SYNTHETIC COMPLETE (SC) AND DROP-OUT MEDIA

In order to test the growth requirements of strains, it is useful to have media in which each of the commonly encountered auxotrophies is supplemented except the one of interest (drop-out media). Dry growth supplements are stored premixed.

SC is a medium in which the drop-out mix contains all possible supplements (i.e., nothing is "dropped out").

• Bacto-yeast nitrogen base without amino acids (0.67%)	6.7 g
• Glucose (2%) <i>→ or 1.7g YNB w/o a.a. + 5.0g (NH₄)₂SO₄</i>	20 g
Bacto-agar (2%)	20 g
• Drop-out mix (0.2%)	2 g
Distilled H ₂ O	1000 ml

Drop-out mix:

Drop-out mix is a combination of the following ingredients minus the appropriate supplement. It should be mixed very thoroughly by turning end-over-end for at least 15 minutes; adding a couple of clean marbles helps.

Adenine	- 0.5 g ✓	✓ Leucine	- 10.0 g
Alanine	- 2.0 g ✓	Lysine	- 2.0 g
Arginine	- 2.0 g ✓	Methionine	- 2.0 g
Asparagine	- 2.0 g ✓	✓ para-Aminobenzoic acid	- 0.2 g
Aspartic acid	- 2.0 g ✓	Phenylalanine	- 2.0 g
Cysteine	- 2.0 g ✓	Proline	- 2.0 g
Glutamine	- 2.0 g ✓	Serine	- 2.0 g
Glutamic acid	- 2.0 g ✓	✓ Threonine	- 2.0 g
Glycine	- 2.0 g ✓	✗ Tryptophan	- 2.0 g
Histidine	- 2.0 g ✓	Tyrosine	- 2.0 g ✓
Inositol	- 2.0 g ✓	Uracil	2.0 g ✓
Isoleucine	- 2.0 g ✓	Valine	- 2.0 g

Drop-out mix (-uracil) for protein expression in yeast contains everything above except uracil.

For yeast two-hybrid experiment, drop-out mix (-leu -trp -his -ade) contains everything above except leucine, tryptophan, histadine, and adenine. The same drop-out mix was used for preparing other media for yeast two-hybrid:

- 1) -leu
- 2) -trp
- 3) -leu, -trp
- 4) -leu, -trp, -his
- 5) -leu, -trp, -ade

and the appropriate supplement were added according to the following table:

Constituent	Stock concentration (g/100 ml)	Volume of stock for 1 liter of medium (ml)	Final concentration in medium (mg/liter)	Volume of stock to spread on plate (ml)
Adenine sulfate	0.2 ^a	10	20	0.2
Uracil	0.2 ^a	10	20	0.2
L-Tryptophan	1	2	20	0.1
L-Histidine HCl	1	2	20	0.1
L-Arginine HCl	1	2	20	0.1
L-Methionine	1	2	20	0.1
L-Tyrosine	0.2	15	30	0.2
L-Leucine	1	10	100	0.1
L-Isoleucine	1	3	30	0.1
L-Lysine HCl	1	3	30	0.1
L-Phenylalanine	1 ^a	5	50	0.1
L-Glutamic acid	1 ^a	10	100	0.2
L-Aspartic acid	1 ^{a,b}	10	100	0.2
L-Valine	3	5	150	0.1
L-Threonine	4 ^{a,b}	5	200	0.1
L-Serine	8	5	400	0.1

^aStore at room temperature.

^bAdd after autoclaving the medium.

REFERENCE LIST

- Aglipay JA, Lee SW, Okada S, Fujiuchi N, Ohtsuka T, Kwak JC, Wang Y, Johnstone RW, Deng C, Qin J *et al.* (2003) A member of the Pyrin family, IFI16, is a novel BRCA1-associated protein involved in the p53-mediated apoptosis pathway. *Oncogene* **22**(55):8931-8938.
- Albrecht M, Choubey D, Lengauer T. (2005) The HIN domain of IFI-200 proteins consists of two OB folds. *Biochem Biophys Res Commun* **327**(3):679-687.
- Alexander M, Heppel LA, Hurwitz J. (1961) The purification and properties of micrococcal nuclease. *J Biol Chem* **236**:3014-3019.
- Aqvist J, Warshel A. (1989) Calculations of free energy profiles for the staphylococcal nuclease catalyzed reaction. *Biochemistry* **28**(11):4680-4689.
- Arcus V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr Opin Struct Biol* **12**(6):794-801.
- Ashraf SI, Ip YT. (1998) Transcriptional control: repression by local chromatin modification. *Curr Biol* **8**(19):R683-R686.
- Bernardi G. (2007) The neoselectionist theory of genome evolution. *Proc Natl Acad Sci U S A* **104**(20):8385-8390.
- Blackwell LJ, Borowiec JA. (1994) Human replication protein A binds single-stranded DNA in two distinct complexes. *Mol Cell Biol* **14**(6):3993-4001.
- Bochkarev A, Bochkareva E. (2004) From RPA to BRCA2: lessons from single-stranded DNA binding by the OB-fold. *Curr Opin Struct Biol* **14**(1):36-42.
- Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L. (1997) Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. *Nature* **385**(6612):176-181.
- Bochkareva E, Korolev S, Lees-Miller SP, Bochkarev A. (2002) Structure of the RPA trimerization core and its role in the multistep DNA-binding mechanism of RPA. *EMBO J* **21**(7):1855-1863.
- Bochkareva E, Belegu V, Korolev S, Bochkarev A. (2001) Structure of the major single-stranded DNA-binding domain of replication protein A suggests a dynamic mechanism for DNA binding. *EMBO J* **20**(3):612-618.

- Bochkareva E, Frappier L, Edwards AM, Bochkarev A. (1998) The RPA32 subunit of human replication protein A contains a single-stranded DNA-binding domain. *J Biol Chem* **273**(7):3932-3936.
- Briggs RC, Briggs JA, Ozer J, Sealy L, Dworkin LL, Kingsmore SF, Seldin MF, Kaur GP, Athwal RS, Dessypris EN *et al.* (1994) The human myeloid cell nuclear differentiation antigen gene is one of at least two related interferon-inducible genes located on chromosome 1q that are expressed specifically in hematopoietic cells. *Blood* **83**(8):2153-2162.
- Brooks TA, Hurley LH. (2009) The role of supercoiling in transcriptional control of MYC and its importance in molecular therapeutics. *Nat Rev Cancer* **9**(12):849-861.
- Brosh RM, Li JL, Kenny MK, Karow JK, Cooper MP, Kureekattil RP, Hickson ID, Bohr VA. (2000) Replication protein A physically interacts with the Bloom's syndrome protein and stimulates its helicase activity. *J Biol Chem* **275**(31):23500-23508.
- Bürkstümmer T, Baumann C, Blüml S, Dixit E, Dürnberger G, Jahn H, Planyavsky M, Bilban M, Colinge J, Bennett KL *et al.* (2009) An orthogonal proteomic-genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. *Nat Immunol* **10**(3):266-272.
- Choubey D, Panchanathan R. (2008) Interferon-inducible Ifi200-family genes in systemic lupus erythematosus. *Immunol Lett* **119**(1-2):32-41.
- Courey AJ, Jia S. (2001) Transcriptional repression: the long and the short of it. *Genes Dev* **15**(21):2786-2796.
- Coverley D, Kenny MK, Lane DP, Wood RD. (1992) A role for the human single-stranded DNA binding protein HSSB/RPA in an early stage of nucleotide excision repair. *Nucleic Acids Res* **20**(15):3873-3880.
- Datta B, Li B, Choubey D, Nallur G, Lengyel P. (1996) p202, an interferon-inducible modulator of transcription, inhibits transcriptional activation by the p53 tumor suppressor protein, and a segment from the p53-binding protein 1 that binds to p202 overcomes this inhibition. *J Biol Chem* **271**(44):27544-27555.
- Dawson MJ, Trapani JA. (1995) IFI 16 gene encodes a nuclear protein whose expression is induced by interferons in human myeloid leukaemia cell lines. *J Cell Biochem* **57**(1):39-51.
- DeYoung KL, Ray ME, Su YA, Anzick SL, Johnstone RW, Trapani JA, Meltzer PS, Trent JM. (1997) Cloning a novel member of the human interferon-inducible gene family associated with control of tumorigenicity in a model of human melanoma. *Oncogene* **15**(4):453-457.

- Di Filippo M, Bernardi G. (2009) The early apoptotic DNA fragmentation targets a small number of specific open chromatin regions. *PloS one* **4**(4):e5010.
- Ding Y, Wang L, Su LK, Frey JA, Shao R, Hunt KK, Yan DH. (2004) Antitumor activity of IFIX, a novel interferon-inducible HIN-200 gene, in breast cancer. *Oncogene* **23**(26):4556-4566.
- Egistelli L, Chichiarelli S, Gaucci E, Eufemi M, Schininà ME, Giorgi A, Lascu I, Turano C, Giartosio A, Cervoni L *et al.* (2009) IFI16 and NM23 bind to a common DNA fragment both in the P53 and the cMYC gene promoters. *J Cell Biochem* **106**(4):666-672.
- Fairbrother WJ, Gordon NC, Humke EW, O'Rourke KM, Starovasnik MA, Yin JP, Dixit VM. (2001) The PYRIN domain: a member of the death domain-fold superfamily. *Protein Sci* **10**(9):1911-1918.
- Fanning E, Klimovich V, Nager AR. (2006) A dynamic model for replication protein A (RPA) function in DNA processing pathways. *Nucleic Acids Res* **34**(15):4126-4137.
- Felsenfeld G, Groudine M. (2003) Controlling the double helix. *Nature* **421**(6921):448-453.
- Fernandes-Alnemri T, Yu JW, Datta P, Wu J, Alnemri ES. (2009) AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. *Nature* **458**(7237):509-513.
- Flati V, Frati L, Gulino A, Martinotti S, Toniato E. (2001) The murine p202 protein, an IFN-inducible modulator of transcription, is activated by the mitogen platelet-derived growth factor. *J Interferon Cytokine Res* **21**(2):99-103.
- Garcia-Lozano R, Wichmann I, Garcia A, Sanchez-Roman J, Gonzalez-Escribano F, Nuñez-Roldan A. (1996) Presence of antibodies to replication protein A in some patients with systemic lupus erythematosus (SLE). *Clin Exp Immunol* **103**(1):74-76.
- Gariglio M, Azzimonti B, Pagano M, Palestro G, De Andrea M, Valente G, Voglino G, Navino L, Landolfo S. (2002) Immunohistochemical expression analysis of the human interferon-inducible gene IFI16, a member of the HIN200 family, not restricted to hematopoietic cells. *J Interferon Cytokine Res* **22**(7):815-821.
- Gariglio M, De Andrea M, Lembo M, Ravotto M, Zappador C, Valente G, Landolfo S. (1998) The murine homolog of the HIN 200 family, Ifi 204, is constitutively expressed in myeloid cells and selectively induced in the monocyte/macrophage lineage. *J Leukoc Biol* **64**(5):608-614.

- Gasior SL, Olivares H, Ear U, Hari DM, Weichselbaum R, Bishop DK. (2001) Assembly of RecA-like recombinases: distinct roles for mediator proteins in mitosis and meiosis. *Proc Natl Acad Sci U S A* **98**(15):8411-8418.
- Georgaki A, Strack B, Podust V, Hübscher U. (1992) DNA unwinding activity of replication protein A. *FEBS Lett* **308**(3):240-244.
- Ginalski K, Elofsson A, Fischer D, Rychlewski L. (2003) 3D-Jury: a simple approach to improve protein structure predictions. *Bioinformatics* **19**(8):1015-1018.
- Goldberger A, Brewer G, Hnilica LS, Briggs RC. (1984) Nonhistone protein antigen profiles of five leukemic cell lines reflect the extent of myeloid differentiation. *Blood* **63**(3):701-710.
- Hashimoto T, Morioka H, Izuta S. (2000) Binding activity of replication protein A to UV-damaged single-stranded DNA. *Nucleic Acids Symp Ser* (44):177-178.
- Heins JN, Suriano JR, Taniuchi H, Anfinsen CB. (1967) Characterization of a nuclease produced by *Staphylococcus aureus*. *J Biol Chem* **242**(5):1016-1020.
- Henricksen LA, Umbricht CB, Wold MS. (1994) Recombinant replication protein A: expression, complex formation, and functional characterization. *J Biol Chem* **269**(15):11121-11132.
- Hiller S, Kohl A, Fiorito F, Herrmann T, Wider G, Tschopp J, Grütter MG, Wüthrich K. (2003) NMR structure of the apoptosis- and inflammation-related NALP1 pyrin domain. *Structure* **11**(10):1199-1205.
- Hornung V, Ablasser A, Charrel-Dennis M, Bauernfeind F, Horvath G, Caffrey DR, Latz E, Fitzgerald KA. (2009) AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. *Nature* **458**(7237):514-518.
- Isono K, Satoh K, Kobayashi H. (2000) Molecular cloning of a cDNA encoding a novel Ca(2+)-dependent nuclease of *Arabidopsis* that is similar to staphylococcal nuclease. *Biochim Biophys Acta* **1491**(1-3):267-272.
- Jacobs DM, Lipton AS, Isern NG, Daughdrill GW, Lowry DF, Gomes X, Wold MS. (1999) Human replication protein A: global fold of the N-terminal RPA-70 domain reveals a basic cleft and flexible C-terminal linker. *J Biomol NMR* **14**(4):321-331.
- Johnstone RW, Kerry JA, Trapani JA. (1998) The human interferon-inducible protein, IFI 16, is a repressor of transcription. *J Biol Chem* **273**(27):17172-17177.

- Johnstone RW, Kershaw MH, Trapani JA. (1998) Isotypic variants of the interferon-inducible transcriptional repressor IFI 16 arise through differential mRNA splicing. *Biochemistry* **37**(34):11924-11931.
- Johnstone RW, Wei W, Greenway A, Trapani JA. (2000) Functional interaction between p53 and the interferon-inducible nucleoprotein IFI 16. *Oncogene* **19**(52):6033-6042.
- Karmakar P, Seki M, Kanamori M, Hashiguchi K, Ohtsuki M, Murata E, Inoue E, Tada S, Lan L, Yasui A *et al.* (2006) BLM is an early responder to DNA double-strand breaks. *Biochem Biophys Res Commun* **348**(1):62-69.
- Kerr ID, Wadsworth RIM, Cubeddu L, Blankenfeldt W, Naismith JH, White MF. (2003) Insights into ssDNA recognition by the OB fold from a structural and thermodynamic study of *Sulfolobus* SSB protein. *EMBO J* **22**(11):2561-2570.
- Kim C, Paulus BF, Wold MS. (1994) Interactions of human replication protein A with oligonucleotides. *Biochemistry* **33**(47):14197-14206.
- Koul D, Obeyesekere NU, Gutterman JU, Mills GB, Choubey D. (1998) p202 self-associates through a sequence conserved among the members of the 200-family proteins. *FEBS Lett* **438**(1-2):21-24.
- Krissinel E, Henrick K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. *Acta Crystallogr D Biol Crystallogr* **60**(Pt 12 Pt 1):2256-2268.
- Landolfo S, Gariglio M, Gribaudo G, Lembo D. (1998) The Ifi 200 genes: an emerging family of IFN-inducible genes. *Biochimie* **80**(8-9):721-728.
- Lavrik OI, Kolpashchikov DM, Weissbart K, Nasheuer HP, Khodyreva SN, Favre A. (1999) RPA subunit arrangement near the 3'-end of the primer is modulated by the length of the template strand and cooperative protein interactions. *Nucleic Acids Res* **27**(21):4235-4240.
- Liepinsh E, Barbals R, Dahl E, Sharipo A, Staub E, Otting G. (2003) The death-domain fold of the ASC PYRIN domain, presenting a basis for PYRIN/PYRIN recognition. *J Mol Biol* **332**(5):1155-1163.
- Liu T, Rojas A, Ye Y, Godzik A. (2003) Homology modeling provides insights into the binding mode of the PAAD/DAPIN/pyrin domain, a fourth member of the CARD/DD/DED domain family. *Protein Sci* **12**(9):1872-1881.
- Ludlow LEA, Johnstone RW, Clarke CJP. (2005) The HIN-200 family: more than interferon-inducible genes? *Exp Cell Res* **308**(1):1-17.
- Luu P, Flores O. (1997) Binding of SP1 to the immediate-early protein-responsive element of the human cytomegalovirus DNA polymerase promoter. *J Virol* **71**(9):6683-6691.

- Meiss G, Franke I, Gimadutdinow O, Urbanke C, Pingoud A. (1998) Biochemical characterization of *Anabaena* sp. strain PCC 7120 non-specific nuclease NucA and its inhibitor NuiA. *Eur J Biochem* **251**(3):924-934.
- Mondini M, Vidali M, Airò P, De Andrea M, Riboldi P, Meroni PL, Gariglio M, Landolfo S. (2007) Role of the interferon-inducible gene IFI16 in the etiopathogenesis of systemic autoimmune disorders. *Ann N Y Acad Sci* **1110**:47-56.
- Mondini M, Vidali M, De Andrea M, Azzimonti B, Airò P, D'Ambrosio R, Riboldi P, Meroni PL, Albano E, Shoenfeld Y *et al.* (2006) A novel autoantigen to differentiate limited cutaneous systemic sclerosis from diffuse cutaneous systemic sclerosis: the interferon-inducible gene IFI16. *Arthritis Rheum* **54**(12):3939-3944.
- Mou TC, Shen MC, Terwilliger TC, Gray DM. (2003) Binding and reversible denaturation of double-stranded DNA by Ff gene 5 protein. *Biopolymers* **70**(4):637-648.
- Murzin AG, Brenner SE, Hubbard T, Chothia C. (1995) SCOP: a structural classification of proteins database for the investigation of sequences and structures. *J Mol Biol* **247**(4):536-540.
- Ohta S, Shiomi Y, Sugimoto K, Obuse C, Tsurimoto T. (2002) A proteomics approach to identify proliferating cell nuclear antigen (PCNA)-binding proteins in human cell lysates. Identification of the human CHL12/RFCs2-5 complex as a novel PCNA-binding protein. *J Biol Chem* **277**(43):40362-40367.
- Pant K, Karpel RL, Williams MC. (2003) Kinetic regulation of single DNA molecule denaturation by T4 gene 32 protein structural domains. *J Mol Biol* **327**(3):571-578.
- Pawłowski K, Pio F, Chu Z, Reed JC, Godzik A. (2001) PAAD - a new protein domain associated with apoptosis, cancer and autoimmune diseases. *Trends Biochem Sci* **26**(2):85-87.
- Pfuetzner RA, Bochkarev A, Frappier L, Edwards AM. (1997) Replication protein A. Characterization and crystallization of the DNA binding domain. *J Biol Chem* **272**(1):430-434.
- Postel EH. (1999) Cleavage of DNA by human NM23-H2/nucleoside diphosphate kinase involves formation of a covalent protein-DNA complex. *J Biol Chem* **274**(32):22821-22829.
- Robbins JB, McKinney MC, Guzman CE, Sriratana B, Fitz-Gibbon S, Ha T, Cann IKO. (2005) The euryarchaeota, nature's medium for engineering of single-stranded DNA-binding proteins. *J Biol Chem* **280**(15):15325-15339.

- Saccone S, Federico C, Bernardi G. (2002) Localization of the gene-richest and the gene-poorest isochores in the interphase nuclei of mammals and birds. *Gene* **300**(1-2):169-178.
- Salas TR, Petruseva I, Lavrik O, Bourdoncle A, Mergny JL, Favre A, Saintomé C. (2006) Human replication protein A unfolds telomeric G-quadruplexes. *Nucleic Acids Res* **34**(17):4857-4865.
- Scully R, Ganesan S, Vlasakova K, Chen J, Socolovsky M, Livingston DM. (1999) Genetic analysis of BRCA1 function in a defined tumor cell line. *Mol Cell* **4**(6):1093-1099.
- Seelig HP, Ehrfeld H, Renz M. (1994) Interferon-gamma-inducible protein p16. A new target of antinuclear antibodies in patients with systemic lupus erythematosus. *Arthritis Rheum* **37**(11):1672-1683.
- Seroussi E, Lavi S. (1993) Replication protein A is the major single-stranded DNA binding protein detected in mammalian cell extracts by gel retardation assays and UV cross-linking of long and short single-stranded DNA molecules. *J Biol Chem* **268**(10):7147-7154.
- Shell SM, Hess S, Kvaratskhelia M, Zou Y. (2005) Mass spectrometric identification of lysines involved in the interaction of human replication protein a with single-stranded DNA. *Biochemistry* **44**(3):971-978.
- Singh KK, Samson L. (1995) Replication protein A binds to regulatory elements in yeast DNA repair and DNA metabolism genes. *Proc Natl Acad Sci U S A* **92**(11):4907-4911.
- Song LL, Ponomareva L, Shen H, Duan X, Alimirah F, Choubey D. (2010) Interferon-inducible IFI16, a negative regulator of cell growth, down-regulates expression of human telomerase reverse transcriptase (hTERT) gene. *PloS one* **5**(1):e8569.
- Stenlund A. (2003) Initiation of DNA replication: lessons from viral initiator proteins. *Nat Rev Mol Cell Biol* **4**(10):777-785.
- Theobald DL, Mitton-Fry RM, Wuttke DS. (2003) Nucleic acid recognition by OB-fold proteins. *Annu Rev Biophys Biomol Struct* **32**:115-133.
- Trapani JA, Browne KA, Dawson MJ, Ramsay RG, Eddy RL, Show TB, White PC, Dupont B. (1992) A novel gene constitutively expressed in human lymphoid cells is inducible with interferon-gamma in myeloid cells. *Immunogenetics* **36**(6):369-376.
- Treuner K, Ramsperger U, Knippers R. (1996) Replication protein A induces the unwinding of long double-stranded DNA regions. *J Mol Biol* **259**(1):104-112.
- Varriale A, Bernardi G. (2010) Distribution of DNA methylation, CpGs, and CpG islands in human isochores. *Genomics* **95**(1):25-28.

- Wang H, Chatterjee G, Meyer JJ, Liu CJ, Manjunath NA, Bray-Ward P, Lengyel P. (1999) Characteristics of three homologous 202 genes (Ifi202a, Ifi202b, and Ifi202c) from the murine interferon-activatable gene 200 cluster. *Genomics* **60**(3):281-294.
- Wang Y, Cortez D, Yazdi P, Neff N, Elledge SJ, Qin J. (2000) BASC, a super complex of BRCA1-associated proteins involved in the recognition and repair of aberrant DNA structures. *Genes Dev* **14**(8):927-939.
- Watson JD, Laskowski RA, Thornton JM. (2005) Predicting protein function from sequence and structural data. *Curr Opin Struct Biol* **15**(3):275-284.
- Wold MS. (1997) Replication protein A: a heterotrimeric, single-stranded DNA-binding protein required for eukaryotic DNA metabolism. *Annu Rev Biochem* **66**:61-92.
- Wold MS, Kelly T. (1988) Purification and characterization of replication protein A, a cellular protein required for in vitro replication of simian virus 40 DNA. *Proc Natl Acad Sci U S A* **85**(8):2523-2527.
- Xin H, Curry J, Johnstone RW, Nickoloff BJ, Choubey D. (2003) Role of IFI 16, a member of the interferon-inducible p200-protein family, in prostate epithelial cellular senescence. *Oncogene* **22**(31):4831-4840.
- Xin H, Pereira-Smith OM, Choubey D. (2004) Role of IFI 16 in cellular senescence of human fibroblasts. *Oncogene* **23**(37):6209-6217.
- Yan DH, Wen Y, Spohn B, Choubey D, Gutterman JU, Hung MC. (1999) Reduced growth rate and transformation phenotype of the prostate cancer cells by an interferon-inducible protein, p202. *Oncogene* **18**(3):807-811.
- Yan H, Dalal K, Hon BK, Youkharibache P, Lau D, Pio F. (2008) RPA nucleic acid-binding properties of IFI16-HIN200. *Biochim Biophys Acta* **1784**(7-8):1087-1097.
- Yin L, Sit KH. (1997) Micrococcal nuclease (endonuclease) digestion causes apoptosis and mitotic catastrophe with interphase chromosome condensation in human Chang liver cells. *Cell Death Differ* **4**(8):796-805.
- Zhang Y, Howell RD, Alfonso DT, Yu J, Kong L, Wittig JC, Liu CJ. (2007) IFI16 inhibits tumorigenicity and cell proliferation of bone and cartilage tumor cells. *Front Biosci* **12**:4855-4863.
- Zuker M. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**(13):3406-3415.

de Laat WL, Appeldoorn E, Sugasawa K, Weterings E, Jaspers NG, Hoeijmakers JH. (1998) DNA-binding polarity of human replication protein A positions nucleases in nucleotide excision repair. *Genes Dev* **12**(16):2598-2609.