# Molecular characterisation of cytosolic DNA recognition

Doctoral thesis at the Medical University of Vienna
for obtaining the academic degree

**Doctor of Philosophy**

Submitted by
**Dipl.-Ing. Evren Karayel**

Supervisor:
Giulio Superti-Furga, PhD
CeMM, Research Center for Molecular Medicine
Lazarettgasse 14/ AKH BT 25.3 1090 Wien/AUSTRIA

Vienna, September 2011

# Danksagung

An erster Stelle möchte ich mich bei **Prof. Giulio Superti-Furga** dafür bedanken, dass er mir die Möglichkeit gegeben hat in seinem Labor zu arbeiten. Es war eine der wichtigsten Erfahrungen meines Lebens. Noch mehr danke ich ihm dafür, dass er in allen professionellen und persönlichen Anliegen mein Mentor war.

Mein ganz besonderer Dank gilt **Tilmann Bürckstümmer** für seine unermüdliche Betreuung und für die Ordnung und Effizienz, die er in mein Leben gebracht hat. Ich danke ihm für die große Hilfe, ohne die diese Arbeit in dieser Form nicht zustande gekommen wäre.

**Adriana Goncalves** danke ich für ihre über vier Jahre anhaltende Treue und Freundschaft. Wir sind gemeinsam durch Höhen und Tiefen gegangen.

Ich danke dem **GSF-Lab** für die Unterstützung, Beratung, und phantastische Zusammenarbeit. Außerdem danke ich dem gesamten **CeMM** für die kreative, intelligente, kompetente, freundliche Atmosphäre.

Weiters bedanke ich mich bei dem **CCHD PhD Programm** für die finanzielle Unterstützung meiner Doktorarbeit. Ich danke Professoren und Freunden aus dem Programm für spannende Diskussionen und interdisziplinären Austausch.

Schließlich danke ich **Familie Karayel, Federico Dalpiaz, Daniela Herta** und **Mustafa Tavşanlı** für ihre ununterbrochene Unterstützung.

# Table of contents

# Abstract

Recognition of foreign DNA by cytosolic innate immune receptors triggers the production of interferon-β (IFN-β) and the subsequent antimicrobial response. However, it is unclear whether different types of DNA ligands are recognized by similar receptors and whether the resulting response is distinct from the response brought about by the so-called Toll-like receptors (TLRs) in the endosomes. To address these questions, we compared the two most commonly used types of DNA ligands (Interferon-stimulatory DNA (ISD) and poly(dAdT)) and assessed the minimal structural requirements for stimulatory capacity in RAW264.7 murine macrophage cells. Gene expression signatures and competition experiments suggest that ISD and poly(dAdT) are qualitatively indistinguishable and differ from the CpG-containing oligonucleotides triggering the TLR9 pathway. Structure-activity relationship analyses revealed that a minimal length of two helical turns is sufficient for ISD-mediated IFN-β induction, while phosphorylation at the 5' end is dispensable. Altogether, our data suggest that in murine macrophages only one major cytosolic DNA recognition pathway is operational.

After characterizing the response in RAW264.7 cells in detail, we aimed at identifying the molecular mechanism and in particular the DNA sensor, responsible for type I interferon induction. For this purpose, we took a systematic approach: First, we captured cytosolic DNA binding proteins from RAW264.7 cells and peripheral blood mononuclear cells (PBMCs) by affinity purification using a synthetic DNA (ISD) and Calf thymus DNA as "baits". The eluates from these purifications were analyzed by mass spectrometry and resulted in 1,606 distinct putative DNA binding proteins. Next, we prioritized the list of captured proteins according to certain defined criteria and selected 46 candidates for a subsequent loss-of-function validation. To this end, we stably transduced RAW264.7 cells with six shRNAs per gene and measured the IFN-β levels after DNA stimulation. Out of 46 DNA sensor candidates tested, at least ten reduced the IFN-β production significantly with two or more shRNAs making them primary candidates for the long-sought-after DNA sensor in macrophages. These candidates will be further validated in future to assess their gain-of-function phenotype, their DNA binding specificity and their mechanisms of action.

# Zusammenfassung

Die Erkennung von fremder DNA durch zytosolische Rezeptoren des angeborenen Immunsystems löst die Produktion von Interferon-β und die anschließende antimikrobielle Antwort aus. Allerdings ist es unklar, ob verschiedene Arten von DNA-Liganden von denselben Rezeptoren erkannt werden und ob sich die daraus resultierende Immunantwort von der durch die endosomalen *Toll-like* Rezeptoren hervorgerufenen Antwort unterscheidet. Um diese Fragen zu beantworten, haben wir zwei der am häufigsten verwendeten DNA Liganden (die Interferon stimulierende DNA (ISD) und poly(dAdT)) miteinander verglichen und die minimalen Strukturanforderungen für die Stimulation von murinen RAW264.7 Makrophagen ermittelt. Genexpressionssignaturen und Kompetitionsexperimente deuten darauf hin, dass ISD und poly(dAdT) qualitativ nicht voneinander zu unterscheiden sind. Dagegen induzierten CpG-haltige Oligonukleotide, die den TLR9-Signalweg stimulieren, deutlich unterscheidbare Genexpressionsmuster. Struktur-Funktionsanalysen zeigten, dass eine minimale Länge von zwei spiralförmigen Windungen ausreichend für die ISD-vermittelte IFN-β Induktion ist, während die Phosphorylierung am 5'-Ende unwesentlich ist. Insgesamt weisen unsere Daten darauf hin, dass in murinen Makrophagen nur ein zytosolischer DNA-Erkennungsweg vorhanden ist.

Im zweiten Teil der Arbeit haben wir die Identifizierung des für die Typ I-Interferon-Induktion verantwortlichen DNA-Sensors angestrebt. Zu diesem Zweck haben wir den folgenden systematischen Ansatz gewählt: Zuerst haben wir zytosolische DNA-bindende Proteine in RAW264.7-Zellen und mononukleären Zellen des peripheren Blutes (PBMCs) durch Affinitätsreinigung unter Verwendung einer synthetischen DNA (ISD) und Kalbsthymus-DNA als "Köder" eingefangen. Die Eluate aus diesen Reinigungen wurden mittels Massenspektrometrie analysiert und führten zur Identifikation von 1606 potenziellen DNA-bindenden Proteinen. Als nächstes haben wir die Liste der identifizierten Proteine nach bestimmten Kriterien gereiht, und 46 Kandidaten für eine anschließende *Loss-of-Function* Validierung ausgewählt. Zu diesem Zweck haben wir RAW264.7-Zellen mit sechs shRNAs pro Gen stabil transduziert und die IFN-β-Produktion nach DNA Stimulation gemessen. Von 46 getesteten DNA-Sensor Kandidaten reduzierten mindestens zehn die IFN-β Produktion mittels zweier oder mehr shRNAs, was sie zu heißen Kandidaten für den lange gesuchten DNA-Sensor in Makrophagen machte. Diese Kandidaten werden in zukünftigen Experimenten näher erforscht, um ihren *Gain-of-Function*-Phänotyp, ihre DNA-Bindungsspezifität und ihre Wirkmechanismen zu charakterisieren.

# 1 Introduction

## 1.1 Immune System

Pathogens are microorganisms that cause diseases in their host and the range of pathogens may vary from viruses over bacteria to parasites. The immune system consists of all the biological structures and processes that provide protection to an organism from invading pathogens. The functions of the immune system are to recognize an infection (immunological recognition), to limit and fight the infection (effector function), to regulate the response (immune regulation) and to protect an individual from infection by the same pathogen (immunological memory) [Janeway 2008].

The mammalian immune system can be divided into two major branches: innate and adaptive immunity. While adaptive immunity is found exclusively in jawed vertebrates, innate immunity is a defense mechanism present in all multicellular organisms. In mammals, the majority of invading pathogens are removed by the innate immune system within minutes or hours. Only if a pathogen passes the innate immune system undamaged and if the innate immune system alarms the adaptive immune system, an adaptive immune response occurs. The onset of adaptive immune response takes days but it eliminates the pathogen more efficiently than the innate immune system. One of the major differences between the innate immunity and adaptive immunity is that the innate immune system recognizes conserved patterns that are associated with pathogens whereas adaptive immune system recognizes pathogens based on specific protein sequences called antigens.

While the majority of cell types of the human body have some components of the innate immune system, adaptive immune system depends mainly on two cell types: B-leukocytes and T-leukocytes. Both cell types carry similar antigen receptors on their surface binding antigens specifically. It is important to mention that in a given cell all antigen receptors bind the same antigen. Upon binding to antigens, the leukocytes differentiate into different effector cells, which eliminate infected cells by themselves or by recruiting other immune cells. T-cells can also differentiate into regulatory T-cells which regulate the immune response. When a pathogen infects a cell, the cell degrades the proteins into short peptides, a process called antigen processing. After that, the peptides are bound to the major histocompatibility complex (MHC) and transferred to the cell surface. T-leukocytes recognize these presented antigens with their surface receptors and initiate the destruction of the pathogen. The

specificity of adaptive immune systems comes from the variation of antibodies and MHCs, each one of them recognizing only one antigen.

A unique feature of adaptive immune system is the generation of immunological memory. During any infection, a fraction of the B-and T-cells differentiate into memory cells, which provide a long-lasting immunity against already encountered pathogens. In case of a repeated infection, they immediately differentiate into active B- or T-cells.

## 1.2 Innate immunity

Innate immunity represents the first line of defense against invading pathogens. The innate immune system recognizes repetitive patterns that are unique to pathogens, the so-called pathogen-associated molecular patterns (PAMPs). PAMPs are recognized by pattern recognition receptors (PRR). PRRs are germline-encoded and are, therefore, inherited. Some of these receptors are expressed in every cell type and not only in dedicated immune cells. Upon recognition of PAMPs, PRRs induce different signaling pathways which eventually elicit cytokine response. The range of cytokines varies from inflammatory cytokines like TNF-α, IL-1 and IL-6 to type I interferons (IFN). While inflammatory cytokines activate immune cells to fight the infection and repair tissue damage, IFNs are produced by infected cells and activate the IFN receptor to trigger an antimicrobial state. This antimicrobial state is created by expression of specific genes which inhibit microbial growth and replication. Overall, cytokines are also means of communication. Cytokines produced by infected and damaged cells can alert the neighboring cells, can induce specific gene expression to protect the cells from infection, and can attract immune cells to the site of infection.

The innate immune system does not acquire immunological memory during the lifespan meaning, unlike the adaptive immune system, the exposure to an infectious agent does not provide protective immunity for subsequent infections by the same agent.

## 1.3 Cells of the innate immune system

Even though almost every cell type of the human body contains some components of the innate immune system, certain cells have specialized innate immune functions.

**Macrophages**

Macrophages are white blood cells that differentiate from monocytes. Upon encounter, they phagocytose the pathogens and induce a plethora of proinflammatory cytokines like TNF-α, IL-1, IL-6 and chemokines like IL-8. While the proinflammatory cytokines elicit a local inflammation, the chemokines attract other innate immune cells to fight against infection.

**Neutrophils**

Together with eosinophils and basophils, neutrophils belong to the family of granulocytes. Neutrophils contain densely staining granule in their cytoplasm and they are relatively short-lived compared the macrophages. Unlike other innate immune cells, they do not reside in the peripheral tissue before the infection. They are rather attracted to the site of infection by chemokines. Recruited at the site of infection, they kill the pathogens either by phagocytosis or by secreting the antimicrobial content of its granules into the extracellular matrix.

**Eosinophils and basophils**

Eosinophils and basophils are innate immune cells that are involved in fighting parasite infections. They contain numerous granules and produce a variety of cytokines. Unlike macrophages and neutrophils, eosinophils and basophils cannot phagocytose pathogens, but they secrete the content of their granule into extracellular matrix. Besides the antimicrobial peptides, eosinophils produce several cationic effector proteins that act against parasitic worms [Levy O. 2000].

**Dendritic Cells (DC)**

In an immature form, DCs reside in the peripheral tissue and express the majority of pattern recognition receptors (PRR). Upon interaction with pathogens, several PRRs in DCs get activated and lead to a number of antimicrobial responses. But the more pronounced role of DC is being an “antigen presenting cell”; once activated, DCs phagocytose the pathogens and present pathogen-derived antigens to T-cells. Therefore, DCs are also seen as a crucial bridge between innate and adaptive immunity.

## 1.4 Receptors of Innate immunity

Receptors of innate immunity are pathogen recognition receptors (PRRs) that recognize molecular features that are conserved among many pathogens, so called pathogen-associated molecular patterns (PAMPs). PRRs are localized on the plasma membrane, in the endosomes or the cytosol and belong to distinct classes, most notably the Toll-like receptors (TLRs), the NOD-like receptors (NLRs) and the RIG-I-like helicases (RLHs) [Palsson-McDermott et al. 2007, Akira 2006]. Engagement of PRRs by specific ligands triggers intracellular signaling cascades that culminate in the production and secretion of type I IFNs, cytokines and chemokines.

### 1.4.1 Toll-like receptors

Toll-like receptors are the best-studied group of innate immune receptors [Foster et al. 2008]. All TLRs consist of an intracellular Toll/IL-1 receptor (TIR) domain and extracellular Leucine-rich repeats (LRR). Mammalian TLRs comprise 13 members whereas TLR11, 12, 13 are pseudogenes in humans [Akira et al. 2006]. TLRs recognize PAMPs deriving from bacteria, virus, fungi and protozoa. Except in certain exceptional occasions, TLRs can be divided in two groups according to their subcellular localization and ligands: While TLR1, 2, 4, 5 and 6 are localized on the plasma membrane and recognize bacterial and fungal PAMPs, TLR3, 7, 8, 9 are localized in endosomes and recognize viral nucleic acids (**Fig. 1.1**).

**TLR1, TLR2, TLR6, TLR5, TLR4**

TLR2 recognizes a wide variety of ligands. These include lipoproteins from various bacterial pathogens, peptidoglycan from Gram-positive bacteria, Zymosan from yeast cell wall, and two kinds of atypical LPS [Takeda et al. 2003, Hirschfeld et al. 2001, Werts et al. 2001, Smith et al. 2003]. The wide spectrum of ligands could be explained by two factors: Firstly, TLR2 forms separate heterodimers with two structurally related TLRs: TLR1 and TLR6. Interestingly, the TLR2/TLR1 heterodimer recognizes a different set of ligands than TLR2/TLR6 heterodimer. Secondly, TLR2 has also been shown to interact with dectin-1, a lectin family receptor, which recognizes β-glucan deriving from fungal cell wall [Akira et al. 2006].

TLR5 recognizes flagellin, the major component of bacterial flagella [Hayashi et al. 2001]. It is expressed on epithelial cells, macrophages, DCs and on the basolateral side of intestinal epithelium, where normally, only the pathogenic but not commensal bacteria are present. It is

tempting to speculate that polarised expression of TLR5 may contribute to the distinction between pathogenic and commensal microbes [Paul 2003, chapter 14].

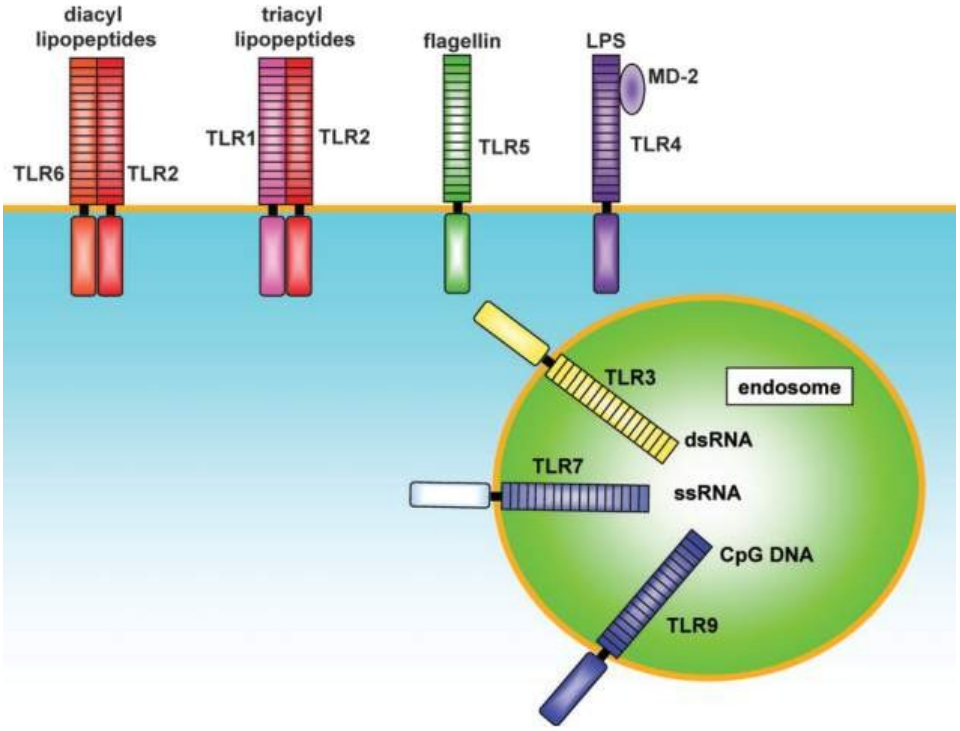



**Fig. 1.1: Toll-like receptors and their ligands.** TLR2 recognizes a lipoproteins and peptidoglycans. While the heterodimer TLR1/TLR2 recognizes triacyl lipopeptides, the heterodimer TLR6/TLR2 recognizes diacyl lipopeptides. TLR4 senses LPS deriving from gram-negative bacteria. TLR5 recognizes the flagelin, the major component of bacterial flagella. On the other hand, the endosomal TLRs sense nucleic acids. TLR3 and TLR7 recognize ss- and dsRNA , respectively. TLR9 senses DNA containing unmethylated CpG motivs (from Takeda et al. 2005).

TLR4 is the receptor for Lipopolysaccharide (LPS) deriving from the outer membrane of Gram-negative bacteria. LPS first builds a complex with LPS-binding protein (LBP), and is transferred on CD14, a GPI-linked protein. CD14 delivers LPS to TLR4. Upon LPS binding, TLR4 forms a heterodimer with MD-2 and gets activated [Park et al. 2009].

**TLR3, TLR7, TLR8, TLR9**

TLR3 is localized on the endosomes and is the receptor for viral dsRNA [Alexopoulou et al. 2002]. The recognition of dsRNA by TLR3 leads to production of type I IFN and proinflammatory cytokines. The majority of the viruses produce double-stranded RNA as replication intermediate, which renders TLR3 a major player in the viral innate immunity.

Highly similar in their structure, TLR7 and TLR8 are also localized on the endosomes but they sense viral ssRNAs [Heil et al. 2004, Diebold et al. 2004, Lund et al. 2004]. TLR7 also recognizes small antiviral compounds, like imidazol and loxoribine. The activation TLR7 and TLR8 creates an antiviral environment by producing antiviral cytokines, especially type I IFN.

TLR9 was one of the first identified sensors triggering innate immune response upon DNA recognition [Hemmi et al. 2000]. It is predominantly expressed in plasmacytoid dendritic cells and it senses DNA containing unmethylated CpG motifs [Gilliet et al. 2008, Kawai et al. 2006]. Unmethylated CpG-DNA is prevalent in viral and bacterial genome and is highly reduced in vertebrates. Inactive TLR9 is located in the endoplasmic reticulum (ER). Upon internalization of CpG-DNA, TLR9 traffics to the endosomes [Tabeta et al. 2006]. Most probably, the trafficking of TLR9 to the endosomes requires the proteolytic cleavage of TLR9 even though the exact mechanism is still unclear [Ewald et al. 2008]. Upon CpG-DNA recognition, TLR9 interacts with MyD88 starting a downstream signaling involving IRAK1, IRAK4, TRAF3, TRAF6 and IRF7 [Barber 2011]. Depending on its localization and downstream interaction partners, TLR9 can induce either type I IFN production or NF-κB activation and IL-12p40 production [Sasai et al. 2010].

It is believed that the endosomal localization of nucleic acid-sensing TLRs protects the host from mounting an immune response to self nucleic acid as self nucleic acids are not found in this subcellular compartment.

**Toll-like receptor signaling pathways**

Upon activation, TLRs elicit a specific response according to the PAMP they recognize. The specificity of the response derives from the recruitment of a single or combination of four adaptor proteins, namely MyD88, TIRAP, TRIF and TRAM and their localization. All of the adaptor proteins contain a TIR domain, which interacts with the TIR domain of TLRs. While all the TLRs, except TLR3, use MyD88 as adaptor protein, TLR4 is the only TLR which recruits all four adaptors.

TLR4 activates two distinct pathways upon activation. When LPS binds to TLR4 on the plasma membrane, TIRAP and MyD88 are recruited to the TLR4 (**Fig. 1.2**). This binding initiates a signaling cascade over IRAKs, TRAF6 and TAK1 recruitment, NF-κB and MAP kinases which eventually leads to proinflammatory cytokine production [Kawai and Akira, 2010]. This is also the pathway used by TLR1/TLR2 and TLR2/TLR6. On the other hand, after the activation of proinflammatory pathway, TLR4 is internalized and recruits the adaptors TRAM and TRIF in the phagosomes. This MyD88 independent activation leads to

activation of the kinases TBK1 and IKKi, subsequent translocation of the transcription factor IRF3 to the nucleus and production of type I IFN [Barton and Kagan 2009]. The complex of TRAM and TRIF also stimulates proinflammatory cytokine production via TRAF6 recruitment [Husebye et al. 2010]. When TLR2 is internalized, it recruits TIRAP and MyD88 but this time with the consequence of type I IFN production [Barbalat et al. 2009]. TLR5 recruits only MyD88 leading to proinflammatory cytokine production over NF-κB activation (**Fig. 1.2**).

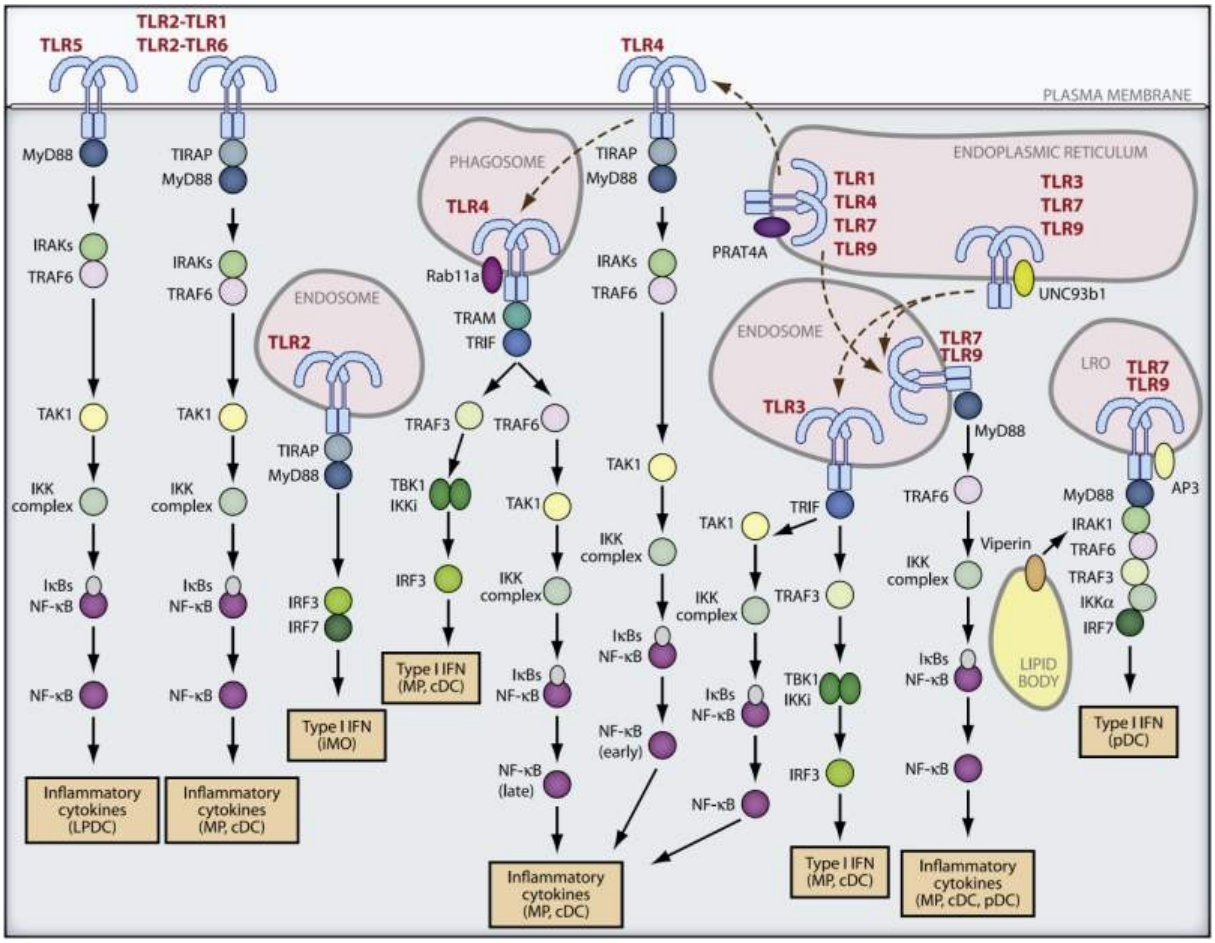



**Fig. 1.2: TLR signaling pathways.** TLRs induce cytokines in various cell types in distinct compartments using overlapping or distinct pathways. MP: macrophages, cDC: conventional denderitic cells (DC), pDC: plasmacytoid DC, LPDC: lamina propria DC, iMO: inflammatory monocytes., LRO: lysosome-related organelle (from Kawai *et al.* 2011)

Like TLR4, TLR3 signals through TRIF leading to the activation of IRF3 and NF-κB and production of proinflammatory cytokines and type I IFN. Finally, both TLR7 and TLR9 recruit MyD88 leading to NF-κB dependent proinflammatory cytokine production in macrophages and conventional DCs.

TLR9 has been shown to elicit two different cellular responses according to its localization [Sasai *et al.* 2010]. Upon activation by CpG-DNA, TLR9 traffics to early endosomes where MyD88 recruitment leads to NF-κB activation and IL-12 production. After that, TLR9 traffics to the lysosome-related organelles where it signals over TRAF3 and IRF3 leading to type I IFN induction.

### 1.4.2 RIG-I like helicases

RIG-I-like helicases are IFN-inducible DExD/H box RNA helicases, which act as cytosolic sensors for viral RNA. This recognition results in the induction of type I IFN and antiviral gene expression creating an antiviral state. To date, three RIG-I like helicases have been identified: RIG-I, MDA5, LGP2 [Rothenfusser et al. 2005, Yoneyama et al. 2004, Loo and Gale 2011].

RIG-I and MDA5 share many similarities in structure. They both contain an RNA binding helicase domain, flanked by two caspase activation and recruitment domains (CARD) on the N-terminus and a regulatory domain in the C-terminus which in case of RIG-I acts as autorepressor (**Fig. 1.3**) [Saito et al. 2007, Yoneyama et al. 2005]. On the other hand, LGP2 lacks the N-terminal CARD domains and, therefore, it is thought not to interact with the common downstream adaptor IPS1. Furthermore, even though LGP2 can bind to RNA, it is still not clear if it is directly involved in viral RNA recognition [Loo and Gale 2011].


**Fig. 1.3: Structural representation of RIG-I like helicases.** RIG-I like helicases consist of an ATPase containing DEAD box helicase domain (DEAD helicase) and a C terminal domain (CTD), which acts in RIG-I as repressor domain. Except LGP2, the RIG-I like helicases contain caspase activation and recruitment domain (CARD) [modified from Loo and Gale 2011].

RIG-I preferentially recognizes shorter RNA pieces with a 5'triphosphate (5'ppp) end [Hornung et al. 2006, Pichlmair et al. 2006]. 5'ppp is believed to be the critical determinant for the distinction between self and viral RNA. In mammals, 5'ppp is removed from mRNAs and tRNA, and is present in rRNA, but is most likely obscured by ribosomal proteins [Alberts

et al. 2007, Ramakrishnan et al. 2002, Pichlmair 2007]. Recently, it has been demonstrated that apart from the 5'ppp, polyuridine motifs with interspersed C nucleotides and panhandle structure are also required to induce a type I IFN response [Saito et al. 2008, Rehwinkel et al. 2010]. While RIG-I has been demonstrated to be necessary for a variety of virus families (Paramyxoviridae, Orthomyxoviridae, Rhabdoviridae, etc), MDA5 has been critical for a limited number of viruses (Picornaviridae). Although not well characterized, long poly(I:C) structures and RNA of high molecular weight has also been suggested as MDA5 ligands [Pichlmair et al. 2009, Kato et al. 2008]

Unlike RIG-I and MDA5, LGP2 has initially been thought as a negative regulator of viral recognition as it lacked the CARD domains and overexpression leads to an impaired type I IFN production [Rothenfusser et al. 2005]. Indeed, mice lacking LGP2 showed enhanced production of type I IFN upon VSV infection. However, EMCV infection in LGP2 deficient mice led a reduced type I IFN production [Venkataraman et al. 2007]. Taken together, LGP2 may play an ambiguous role in viral recognition [Pichlmair et al. 2007].

Upon activation, RIG-I and MDA5 associate with the common adaptor protein IPS-1/CARDIF/MAVS/VISA (**Fig. 1.4**) [Kawai et al. 2005; Meylan et al. 2005; Seth et al. 2005; Xu et al. 2005] IPS-1 is associated with the mitochondrial membrane through its C-terminal domain and contains a N-terminal CARD domain that allows a homotypic interaction with the CARD domains of RIG-I and MDA5. Through the interaction with MDA5 and RIG-I, IPS-1 recruits TRAF6 and NAP1. As a complex they activate the kinases TBK1 and IKKi which phosphorylate IRF3 and IRF7 leading to its dimerization and translocation [Hacker et al. 2006; Kawai et al. 2005; Kumar et al. 2006; Meylan et al. 2005; Oganesyan et al. 2006; Sasai et al. 2006; Seth et al. 2005; Sun et al. 2006; Xu et al. 2005]. Together with the activation of NF-κB pathway, the IRF3 and IRF7 initiate in the nucleus transcription of type I IFN genes (**Fig. 1.4**) [Honda et al. 2005, Honda et al. 2006].

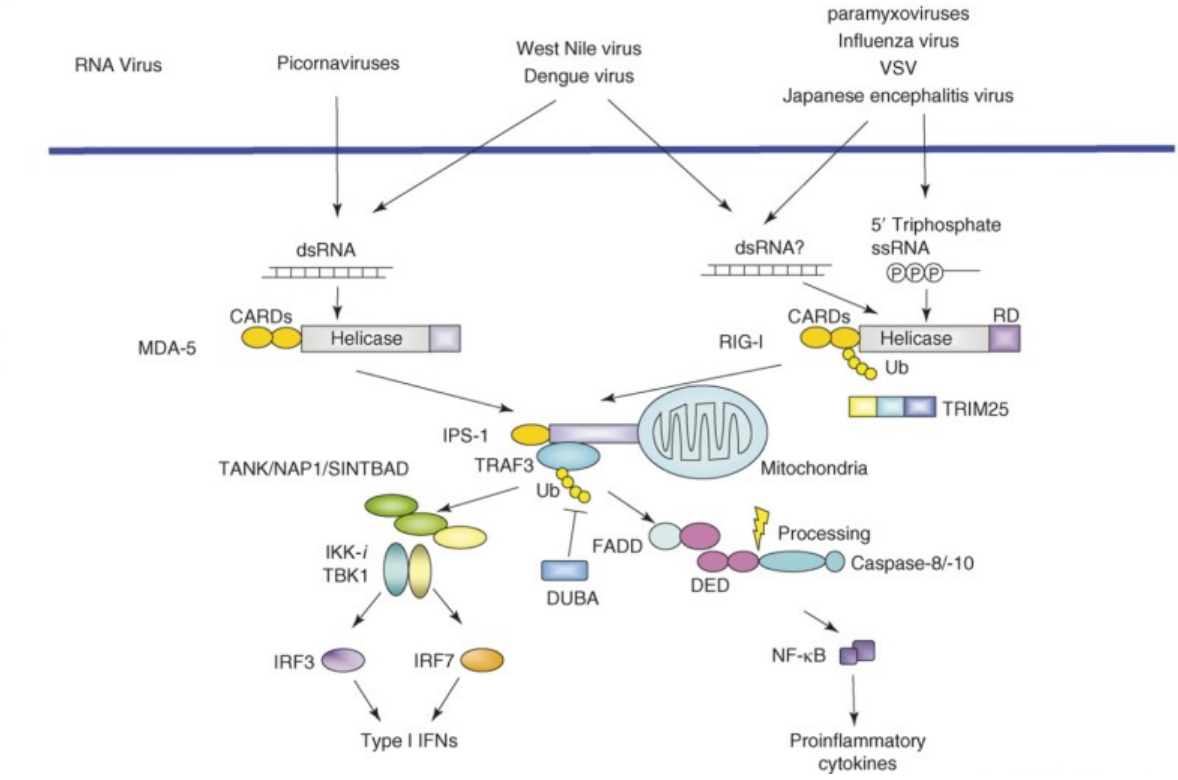



**Fig. 1.4: Recognition of RNA viruses by RIG-I like helicases(RLH).** RIG-I recognizes 5'triphosphate RNA and dsRNA from various viruses. MDA5 recognizes Picornaviruses. Upon activation, RIG-I and MDA5 interact with IPS-I associated to the mitochondria. IPS-1 form together RLHs and downstream signaling molecules an IPS-1 signalosome which induces type I IFN and proinflammatory cytokine production [Loo and Gale 2011] [from Takeuchi and Akira 2008].

### 1.4.3 Cytosolic DNA recognition

Despite “thousands of man-years worth of DNA transfections”, it has only recently been noticed that introduction of dsDNA into the cytosol triggers a potent innate immune response, leading to the cleavage of pro-IL-1β into active IL-1β, and production of IFN-β and other cytokines [Muruve et al. 2008, Ishii et al. 2006, Stetson and Medzhitov 2006, Pichlmair and Reis e Sousa 2007]. Even though significant progress has been made in the last years due to intensive investigations in the field and many components of the cytosolic DNA recognition have been unveiled, some receptors initiating the signaling pathways are still elusive.

Based on the current knowledge, Hornung and Latz classify the cytosolic DNA recognition system in three pathways: first, DNA is directly sensed by cytosolic DNA receptor(s) leading to type I IFN and proinflammatory cytokine production. Second, DNA is transcribed to RNA by RNA polymerase III and sensed by RIG-I with similar transcriptional outcomes. Finally, DNA is sensed by AIM2 leading to inflammasome formation and subsequent cleavage of pro-IL-1β and pro-IL-18 [Hornung and Latz, 2010] (**Fig. 1.5**).

**IFN inducing cytosolic DNA recognition**

In 2006, Ishii and colleagues reported that when B-DNA, DNA in right-handed spiral confirmation, is transfected into the cytosol, type I IFN and other cytokines were induced in a TLR-independent but IRF3-dependent manner. Especially, poly(dAdT), a synthetic DNA consisting of repetitive dA-dT sequences, was by far the most potent activator of transcription, even activating human embryonic kidney cells (HEK293), a cell line normally non-responsive to DNA transfection [Ishii et al. 2006]. In the same year, Stetson and Medzhitov reported that a synthetic, well-defined, 45bp-long piece of DNA (Interferon stimulatory DNA, ISD) leads to IFN production in a sequence independent manner when transfected into the cytosol of primary macrophages or DCs but not of HEK293 cells [Stetson and Medzhitov 2006]. These two observations led to the assumption that two ligands activate different receptors in different cell types that both trigger type-I interferon induction [Hornung and Latz, 2010].

The first ever protein to be suggested as IFN inducing cytosolic DNA receptor was DAI/ZBP1 [Takaoka et al. 2007]. DAI is an IFN inducible protein which is capable to bind B-DNA. Experiments have shown that cytosolic DNA triggers type I IFN response and NF-κB activation in a DAI/ZBP1-dependent manner. However, DAI-deficient mice had a normal IFN response when challenged with B-DNA or infected with DNA viruses [Ishii et al. 2008]. Furthermore, the impact of DAI-specific siRNAs and shRNAs was limited to L929 fibroblasts,

but was not observed in mouse embryonic fibroblasts [Wang et al. 2008]. Taken together, these reports suggested that while DAI is required for DNA recognition in certain cell types, additional, possibly redundant, DNA receptors and recognition pathways exist (**Fig. 1.5**).

In 2010, Unterholzner and her colleagues identified IFI16, a pyrin domain protein, as a DNA sensor for the IRF pathway [Unterholzner et al. 2010]. The study showed that IFI16 binds to a DNA motif derived from vaccinia virus and interacts directly with STING (see below). Loss-of-function experiments demonstrated that DNA-induced type I IFN production was dependent on IFI16, TBK1 and STING. DNA stimulation experiments with geneteically deficient mice will shed light on the question if IFI16 is a cytosolic DNA receptor. However, there are several murine orthologs and the closest homologue of IFI16, IFI204 shares an amino acid identity of only 37%. Although the loss of IFI204 in mouse cell lines has similar effects as loss of IFI16 in human cells, it remains to be seen whether the IFI204 knockout mice will have a phenotype.

The view that the DNA is sensed only by a specific DNA receptor was challenged when, in 2009, two groups simultaneously discovered a novel DNA recognition pathway which involved the cytosolic RNA sensor RIG-I [Ablasser et al. 2009, Chiu et al. 2009]. Both studies demonstrated that poly(dAdT), when transfected into the cytosol of human cells, was first transcribed by RNA polymerase III into RNA. The newly transcribed RNA was carrying 5' triphosphates, thereby rendering the RNA an optimal ligand for RIG-I. As RNA polymerases are ubiquitous and function in the nucleus, it raises the possibility that the initiation of DNA recognition may occur in the nucleus, even though it is also known that RNA polymerases are also functional in the cytoplasm (**Fig. 1.5**) [Jaehning et al 1977]. However, the physiological relevance of this pathway remains a matter of debate as this response seems to be confined to DNAs with AT content.

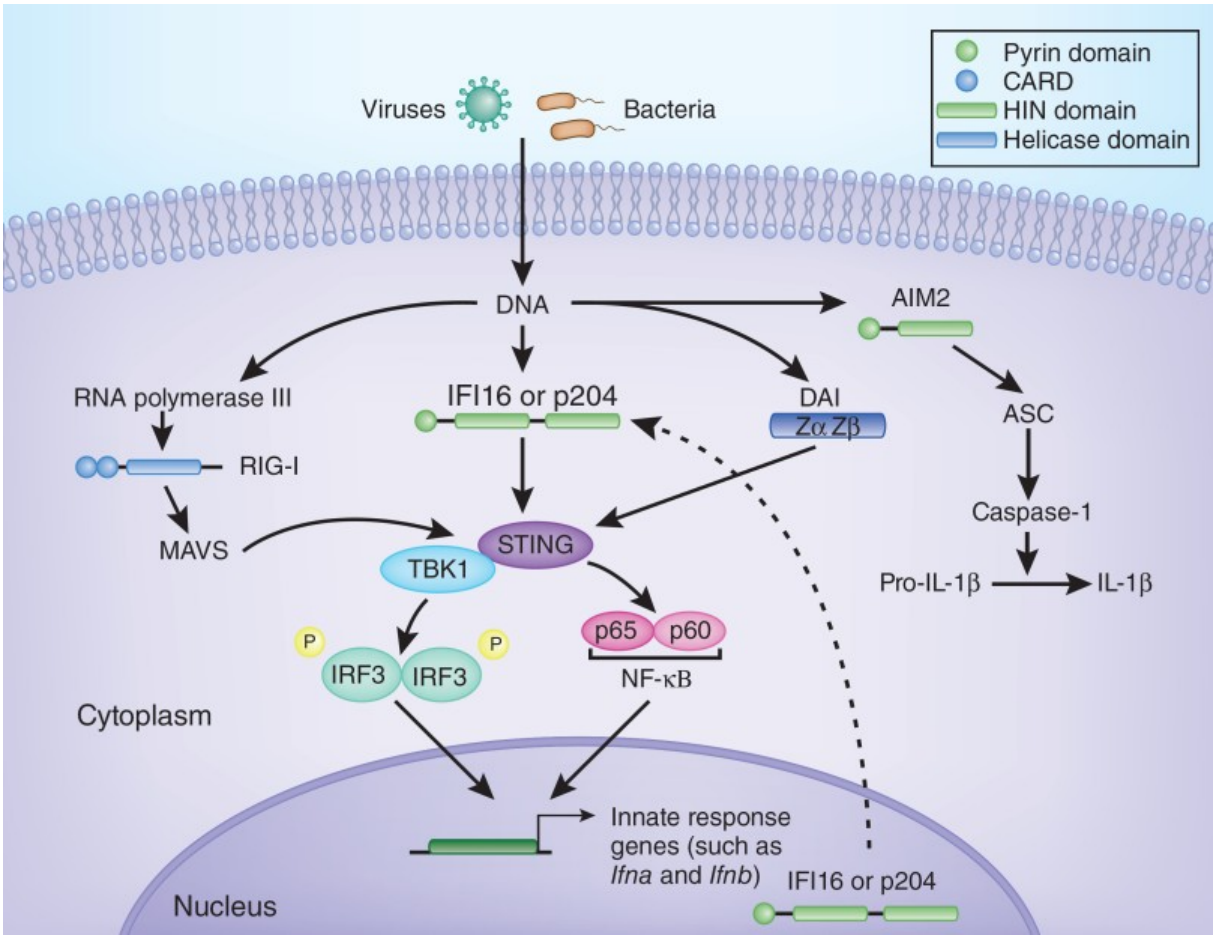



**Fig. 1.5: Cytosolic DNA induced pathways leading to Type I IFN production.** Upon infection with bacteria or viruses, DNA enters the cytosol and is recognized by different receptors. In human cells AT-rich sequences can be transcribed by RNA polymerase III to RNA and recognized by RIG-I. Other proposed recognitions occur via DAI and IFI16. All the so far identified recognition pathways converge on interacting with STING and activating downstream signaling molecules which leads to type I IFN production. Interestingly, IFI16 is predominantly located in the nucleus. Additionally, cytosolic DNA can be sensed by AIM2 leading to formation of inflammasome with ASC and caspase-1 to mediate the cleavage of proIL-1β into active IL-1β [from Goubau et al. 2010]

It is not unexpected that the evolution resulted in a number of partly redundant DNA sensors given the importance of DNA recognition in the host defense against viruses and bacteria. Recent developments in the field of cytosolic DNA answered many open questions but the sensor in certain cell types or for certain types of DNA is still not determined. While the major signaling components downstream of STING have been identified, the sensor(s) for ISD recognition in macrophages or non-AT-rich DNA recognition in PBMCs and mouse DC is (are) still elusive (**Table 1.1**). Pinpointing the remaining pieces of this “puzzle” will certainly contribute to identify novel drug targets in DNA-induced inflammations hence will improve current therapies.

**Table 1.1: Mechanisms of IFN-inducing cytosolic DNA sensing** (modified from Hornung and Latz, 2010)

| Type of DNA | Cell type | The sensor |
|---|---|---|
| AT-rich DNA | HEK293 cells, human DCs | RNA polymerase III, RIG-I |
| Non AT-rich DNA | PBMCs, mouse DC, MEFs, | IFI16, Not identified |
| ISD | mDCs, macrophages, MEFs | Not identified |
| Poly(dAdT) | HEK293, human DC, HeLa, | RNA polymerase III, RIG-I |
| Poly(dAdT) | Mouse macrophages, MEFs | Not identified |

**IFN inducing cytosolic DNA pathway**

Focusing on the downstream signaling pathway shows that the type I IFN production upon any type of cytosolic DNA recognition depends on TBK1 activation and IRF3 and IRF7 translocation [Stetson and Medzhitov 2006, Ishii et al. 2006]. STING has been identified as an additional player of this pathway acting upstream of TBK1 but downstream of the nucleic acid sensor as it does not have a nucleic acid binding capacity. [Ishikawa et al. 2008, Zhong et al. 2008, Sun et al. 2009]. STING is a membrane bound protein localized in the endoplasmatic reticulum (ER). Interestingly, STING has been shown to act downstream of DNA as well as RNA recognition. STING deficient mice failed to mount an IFN response to RNA and DNA viruses [Ishikawa et al. 2008, Ishikawa et al. 2009]. Upon interacting with IPS-1 or IFI16 (and probably also unknown DNA sensor), STING triggers the TBK1 dependent phosphorylation of IRF3 and activation of NF-κB. These transcription factors move then to the nucleus to initiate the type I IFN production (**Fig. 1.5**) [Goubau et al. 2010].

Next to cytosolic IFN-inducing DNA receptors, some proteins have been identified to recognize the cytosolic DNA but to play an auxillary role in the IFN induction. First such a protein was high mobility group box (HMGB) proteins [Yanai et al. 2009]. The innate immune response to RNA and DNA in cells lacking HMGB proteins is severely impaired. In this case, HGMBs are thought to act upstream of TLRs and cytosolic receptors and operate as sentinels. Yanai and colleagues propose a mechanism where HMGBs bind to every type nucleic acid and transfer it to the respective sensor even though where the binding of HMGBs to nucleic acids occurs and how this complex activates their respective receptor signaling cascade remains unclear. Similar to HMGBs, LRRFIP1 has been identified to bind RNA and DNA recruiting β-catenin to enhance IFN-β response [Yang et al. 2010]. Yang and colleagues demonstrate that upon association with LRRFIP1, β-catenin is phosphorylated

and translocates to the nucleus leading to recruitment of histone acyltransferase p300 to the *ifnb1* promoter enhancing the subsequent IFN-β production. LRRFIP1 enhances the IFN-β response to VSV, a single-stranded RNA virus, and to *Listeria monocytogenes*, a cytosolic bacterium. Experiments in LRRFIP1 and β-catenin deficient cells will shed light on the questions if LRRFIP1 recognizes every nucleic acid or if it has ligand specificity and show if the β-catenin-enhanced IFN-β response is limited to LRRFIP1 pathway [Rathinam et al. 2010].

**Inflammasome activating cytosolic DNA recognition**

After much of the research on DNA recognition was focused on the type I IFN pathway, four independent groups identified AIM2, a cytosolic DNA sensor which led to IL-1β release [Buerckstuemmer et al. 2009, Fernandes-Alnemri et al. 2009, Hornung et al. 2009, Roberts et al. 2009]. AIM2-induced IL-1β release occurs in an IRF3- and TBK1-independent manner. AIM2, like IFI16, belongs to the family of PYHIN proteins. PYHIN proteins contain a HIN200 domain which binds DNA and a pyrin domain for protein-protein interaction. After the DNA binding by HIN200 domain, the pyrin domain recruits ASC, a common adaptor for those NOD-like receptors that have a pyrin domain (called NLRPs). The formation of inflammasome leads to activation of caspase 1 which in turn cleaves pro-IL-1β and pro-IL-18 leading to secretion of these cytokine into the extracellular matrix (**Fig. 1.5**).

## 1.5 Diseases related to aberrant DNA recognition

In the majority of times and cases, the immune system is a protective barrier against invading pathogens. However, sometimes it is the immune system itself that poses a threat to an organism [Paul et al. chapter 29]. A malfunction in immune system results in severe abnormalities. While a “too diligent” immune system may lead to autoimmune diseases, a “too lazy” immune system may lead to infections. This chapter will focus on the autoimmune diseases due to aberrant DNA recognition.

As a general definition, it can be stated that autoimmune diseases are diseases in which the immune system fails to distinguish self from foreign and mounts an immune response against self. The resulting tissue damage is mediated by T cells and antibodies [Paul et al. 2003, chapter 41]. Even though this definition suggests autoimmune diseases as a consequence of dysfunctional adaptive immunity, it is clear that it is the interplay of the innate and adaptive immunity that leads to systemic autoimmune diseases. Generally, autoimmune disorders occur if the immune system fails to distinguish between pathogenic and non-pathogenic agents or if it is activated in the absence of an antigen.

One of the most important features of immune system is its ability to discriminate between self and non-self antigens. For a correctly functioning immune system, the innate immune system and adaptive immunity implemented different strategies. The innate immune system recognizes the PAMPs, which are unique to pathogens and are not present in the host [Janeway *et al.* 1989, Janeway *et al.* 1992]. The recognition of PAMPs elicits an immediate immune response. In contrast to that, some parts of the innate immune system recognize molecules that are only present on healthy and uninfected cells. For example, natural killer cells (NK cells) survey cells for MHC-1, which is normally constitutively expressed and only in case of an infection, the expression is lost. Loss of MHC-1 leads to activation of NK cells, which eliminates the target cells [Kärre *et al.* 1997]. On the other hand, the adaptive immune system developed “immunological tolerance”. Immunological tolerance consists of various mechanisms including elimination, inactivation and suppression of immune cells. For example, adaptive immune system actively eliminates B- and T-cells that recognize “self” patterns by negative selection. Additionally, it evolved regulatory cells which dampens the immune response in case of a response to “self antigens”.

Systemic Lupus Erythematosus (SLE) is an autoimmune disease characterized by the presence of antinuclear antibodies, microvascular inflammation, tissue damage, and antibodies against dsDNA. While the exact cause of SLE is unknown, the combination of

multiple factors like the genetic make-up, the hormonal system and the environment are thought to play a key role. The prevalence of SLE worldwide ranges from 40 to 50 per 100,000. Recently, improved diagnosis and therapy have increased the 10-year-survival rate to over 90% [Trager J *et al.* 2001, Kasitanon N *et al.* 2006]. It is difficult to make correlation between the onset of the disease and presence of autoantibodies but all the SLE patients produce autoantibodies and about 50 to 70% of all SLE patients produce antibodies against dsDNA [Paul et al. 2003, chapter 41]. One of the other important symptoms of SLE is the upregulation of type I IFNs. The elevated levels of type I IFN and the presence of antibodies against dsDNAs led to the assumption that this may be due to RNA or DNA containing immune complexes [Ronnblom *et al.* 2006]. Under healthy conditions, the presence of DNA is limited to the nucleus and mitochondria. DNA encountered in other compartments or extracellular space is digested by cellular DNases. For example, DNase I is found in the extracellular space is thought to digest DNA deriving from apoptotic cells. DNase I deficiency or mutations are associated with SLE in mice and human [Napirei et al. 2000, Yasutomo et al. 2001, Barber 2011]. The presence of DNase II digests the DNA of engulfed apoptotic and necrotic cells and its presence is limited to lysosomes. Mice deficient in DNase II succumb due to type I IFN overproduction [Yoshida et al. 2005]. The current treatment is based on suppressing the immune system by corticosteroids and hydroxychloroquine which blocks the TLR activation hence the nucleic acid related cytokine production. Recently, a member of DNase III family, TREX1 has been reported to digest DNA, reverse transcribed from nuclear retroelements, thus, preventing their accumulation in the cytosol [Stetson et al. 2008]. It has been known that mutations in the human Trex1 gene cause SLE and Aicardi-Goutières syndrome, a SLE like disease [Crow et al. 2006, Lee-Kirsch et al. 2007]. Mice deficient in Trex1 develops an autoimmune condition which can be rescued by genetic ablation of IRF3 suggesting the IFN pathway responsible for the autoimmune condition [Stetson et al. 2008]. Taken together, these reports suggest the aberrant recognition of self DNA or insufficient sequestering of cytosolic DNA could generate autoimmune diseases.

## 1.6 Functional proteomics

Proteomics is the systematic analysis of proteins in a tissue or a cell, mostly using mass spectrometry (MS) as technological basis [Domon and Aebersold 2006]. Two principal approaches in proteomics are expression-based and functional proteomics. The first approach aims to determine the expression of every protein in a cell or tissue. The latter approach employs some affinity purification steps prior to protein MS to decrease the complexity of the proteome under investigation. Furthermore, the term “functional” refers to the aim of identifying proteins that are functionally connected by determining physical protein-protein or protein-agent interactions. The nature of the agent can range from biological (e.g. nucleic acids, fatty acids…) to chemical compounds (e.g. drugs).

The affinity purification exploits unique features either of a protein of interest (like DNA binding) or of a tag attached to the protein of interest to purify the interacting partners. In affinity purification, the proteins of interest are purified from a heterogeneous mixture of proteins like cell lysate. For example, we aimed to capture cytosolic DNA binding proteins. For that purpose, we took advantage of the DNA binding property of these proteins. We immobilized DNA on a matrix and cytoplasmic extracts of different cell types went through this matrix. In ideal case, only proteins which have the desired feature bind the matrix. Subsequently, the bound proteins are removed from the matrix by a step called elution.

The captured proteins are separated by gel electrophoresis prior the mass spectrometric analysis in order to decrease further the complexity of the proteome. The separated proteins are cut out of the gel and digested for MS-analysis (*in gel* digest). Gel electrophoresis increases the possibility of identification of low abundant proteins by creating protein groups of the same molecular weight. Conversely, substantial sample loss may occur during the extraction of peptides from the gel. Alternatively, the sample may be digested into peptides directly without any separation (*in solution* digest).

The mass spectrometer measures the mass-over-charge ratio (m/z) of ions in the gas phase. Therefore, the peptides need to be transferred into the gas phase and then ionized. Two major methods for ionization are matrix assisted laser desorption/ionization (MALDI) and electrospray ionization (ESI). After ionization, the m/z ratio of peptides is determined by measuring the time of peptides to travel over a fixed distance (Time of flight, (TOF)) which is a function of m/z [Mallick and Kuster 2010]. Another way to determine the m/z is to use oscillating electrostatic fields forcing ions into spiraling trajectories (Quadrupole). This technique takes advantage of the fact that at any given field frequency and amplitude only

species of a single m/z can travel and be recorded [Mallick and Kuster 2010]. After the first round of MS analysis, selected peptides are fragmented and undergo a second MS analysis for identification of the peptide sequence (tandem mass spectrometry (MS/MS)) The complexity of the protein mixtures generated by proteomic experiments are generally so high that an additional chromatographic separation step prior the tandem mass spectrometry is indispensible (Liquid chromatography (LC)-MS/MS). Finally, in order to identify the proteins the measured peptide masses are searched against a database which consists computed masses of *in silico* digested proteins (**Fig. 1.6**).

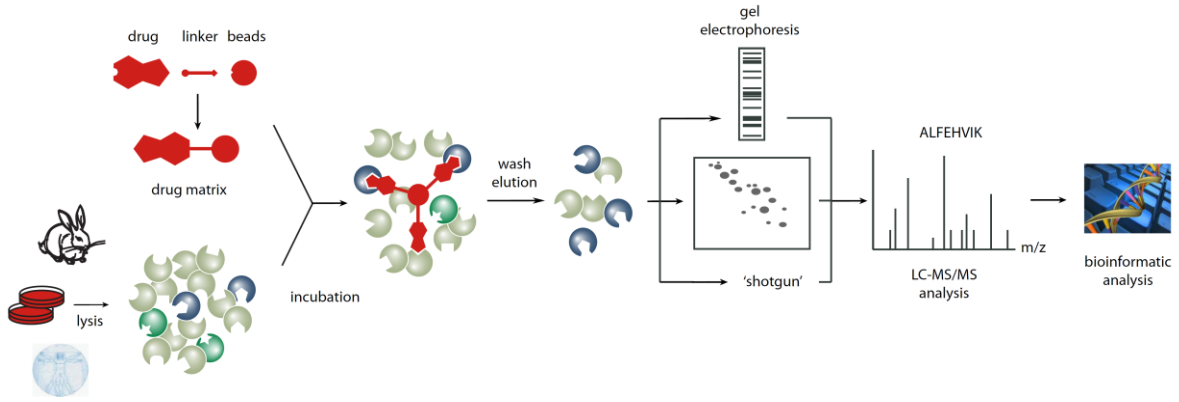



**Fig. 1.6.: A showcase for proteomic flow.** A heterogeneous mixture of proteins like cell lysate or human samples are incubated with a drug which is immobilized on a resin in order to find drug targets. After several washes the bound proteins are eluted and either separated by electrophoresis or digested directly in solution (shotgun approach). The peptides are separated by LC which is coupled to an MS/MS. The measured peptides masses are searched against databases which consist of computed masses of in silico digested proteins [with the courtesy of Uwe Rix].

# 2 Aim of the study

Cytoplasmic DNA triggers the production and secretion of proinflammatory cytokines and interferons. However, it is unclear whether different types of DNA ligands are recognized by similar receptors and whether the resulting response is distinct from the endosomal TLR response. As our understanding of cytosolic DNA recognition is still limited and critical DNA sensors have remained elusive, we first aim at characterizing cytosolic DNA recognition in more detail. To this end, we will compare different types of DNA ligands and assess the minimal structural requirements for stimulatory capacity as well as the consequences of cytosolic DNA recognition at the level of cellular signal transduction in RAW264.7 cells.
The second aim of this study is to identify the elusive IFN-inducing DNA receptor. We envisage to take a systematic approach: First, we capture cytosolic DNA binding proteins from different cell types by affinity purification using different DNA ligands as “baits”. The eluates from these purifications will be analyzed by mass spectrometry to provide a comprehensive list of DNA-binding proteins. Next, we will prioritize the list of captured proteins by the means of a scoring scheme and select candidates for a validation screen from highly ranked proteins. Finally, in order to test the effect of candidate proteins on the IFN-β response, we will carry out a loss-of-function validation by creating shRNA cell lines for selected candidates and measuring the IFN-β levels after DNA stimulation.

# 3 Materials and Methods

## 3.1 Materials

### 3.1.1 Reagents

Interferon stimulatory DNA (ISD) was synthesized by IBA (Germany) (ISD sense 5`-tacagatctactagtgatctatgactgatctgtacatgatctaca-3', ISD antisense 5'-tgtagatcatgtacagatcagtcatagatcactagtagatctgta-3'). The ISD as described originally contains a 5' hydroxyl group [Stetson and Medzhitov 2006]. The phosphorylated version was generated by incubation with Clp1 in the presence of ATP [Weitzer and Martinez 2007]. The blocked versions were labeled at the 5' end with an amino group attached to a 6-carbon spacer or a biotin group (IBA) [Agrawal et al. 1986].

For all the experiments following reagents were used, if not otherwise indicated:

**Table 3.1: List of Reagents**

| Reagent | Manufacturer |
|---|---|
| Acrylamide/Bis-acrylamide 30% | Sigma-Aldrich |
| Adenosine 5' triphosphate (ATP) | GE Healthcare |
| Ammoniumperoxodisulfate | Merck |
| Biospin column | Bio-Rad |
| Blotting-Grade Blocker, nonfat dry milk | Bio-Rad |
| Bromphenol Blue | Sigma-Aldrich |
| CellTiter-Glo® Luminescent Cell Viability Assay | Promega |
| Chloroquine | Sigma-Aldrich |
| CoenzymeA | Sigma |
| CpG (ODN 1826) | Invivogen |
| CT-DNA sodium salt | Sigma-Aldrich |
| d-Luciferin | Promega |
| DMEM (Dulbecco's modified Eagle's Medium) | PAA |
| DNA-cellulose double-stranded, from CT-DNA | Sigma-Aldrich |
| dNTP | Fermentas |
| DTT | Sigma-Aldrich |
| ECL Western Blotting Detection Reagents | GE Healthcare |
| EDTA | Sigma-Aldrich |
| EGTA | Serva |
| Fetal Calf Serum (FCS) | Gibco |
| genome-wide GeneChip Mouse Gene 1.0 ST array | Affymetrix |
| Glycerol | Merck |
| Glycin | Merck |

| | |
|---|---|
| HEPES | Merck |
| Hydrochloric acid | Merck |
| IFN-β specific ELISA | PBL Biomedical |
| KCl | Sigma-Aldrich |
| Lipofectamine 2000 | Invitrogen |
| LPS | Sigma-Aldrich |
| Lymphoprep | Nycomed |
| Methanol | Sigma-Aldrich |
| $MgSO_4$ | Merck |
| Mouse Interferon-β | PBL Biomedical |
| NaCl | Merck |
| NaF | Sigma-Aldrich |
| Nitrocellulose transfer membrane | Whatman |
| NP-40 substitute | Calbiochem |
| $Oligo(dT)_{18}$ primers | Fermentas |
| OptiMEM® Reduced Media | Invitrogen |
| Passive lysis buffer | Promega |
| PBS | PAA |
| Penicillin/Streptomycin | PAA |
| Plasmid Plus Midi Kit | Qiagen |
| Poly (I) | Sigma-Aldrich |
| Poly(C) cross-linked 4% beaded agarose | Sigma-Aldrich |
| poly(I:C) sodium salt | Sigma-Aldrich |
| Polybrene | Sigma-Aldrich |
| Puromycin dihydrochloride | Sigma-Aldrich |
| Reaction buffer (5X) | Fermentas |
| RevertAid M-MuLV Reverse Transcriptase | Fermentas |
| RNasin Plus RNase Inhibitor | Promega |
| Sodium citrate | Sigma-Aldrich |
| Sodium orthovanadate | Sigma-Aldrich |
| Sodium pyrophosphate | Sigma-Aldrich |
| Sodiumdodecylsulfate (SDS) | Serva |
| Spermidine | Sigma-Aldrich |
| Spermine | Sigma-Aldrich |
| Streptavidin Plus UltraLink® Resin | Pierce® |
| Sucrose | Merck |
| SV Total RNA Isolation System | Promega |
| Sybr Green | Bioline |
| Tetramethylethylendiamin (TEMED) | Merck |
| TNF-α specific ELISA | BD Biosciences |
| Tricine | Sigma |
| Tris[hydroxymethyl]aminomethane (Tris) | Sigma-Aldrich |
| Triton X-100 | Sigma-Aldrich |
| Trypsin | PAA |
| Water (pro analysi) | Merck |
| Z-FA-FMK | Sigma-Aldrich |
| β-Mercaptoethanol | Sigma-Aldrich |

### 3.1.2 Cells

RAW264.7, LL171 cells were maintained in DMEM, supplemented with 10% fetal calf serum, in a 5% $CO_2$ humidified atmosphere. LL171 cells were a kind gift from Mireia Pelegrin (Montpellier, France). RAW264.7 and LL171 cells were detached from plates after a single wash with PBS using Trysin and RAW dissociation buffer containing 270 mM KCl and 0,3 M sodium citrate, respectively.

### 3.1.3 Antibodies

The DAI-specific antiserum was custom synthesized (Eurogentec) using recombinant purified DAI (full-length) as an antigen. Rabbit polyclonal antibody against actin protein (#AAN01) was purchased from Cytoskeleton. Rabbit polyclonal TBK1/NAK antibody (#3013) was purchased from Cell signaling technology. Monoclonal antibody against Tubulin protein (#T9026) was purchased from Sigma. Mouse monoclonal antibody against RCC1 (E-6) (#sc55559) was purchased from Santa Cruz Biotechnology, Inc. Peroxidase-conjugated AffiniPure goat anti-rabbit antibody and peroxidase-conjugated AffiniPure goat anti-mouse antibody were purchased from Jackson Immuno Research.

# 3.2 Methods

## 3.2.1 DNA stimulation

RAW264.7 cells were seeded on a 6-well plate and 24-well plate at a density of $10^6$ and $2x10^5$ cells per well, respectively. On the next day, the cells were transfected with differently modified ISDs, poly(dAdT) at a concentration of 1µg/ml using Lipofectamine 2000 according the manufacturer's instructions, if not otherwise indicated. 1 µg/ml LPS and 10 µg/ml poly(I:C) and 20 µM CpG were added directly to the medium. Cells and supernatants were collected after 4 hours. For immunoblot analysis, cells were lysed using Frackelton buffer (10 mM Tris/HCl pH7.5, 50 mM NaCl, 30 mM sodium pyrophosphate, 1% Triton X-100, 1 mM DTT, 100 µM sodium orthovanadate, 50 µM NaF and protease inhibitors).

## 3.2.2 mRNA extraction and reverse transcription

For quantitative real-time PCR, total RNA was extracted from RAW264.7 cells using Promega SV total RNA isolation system according to manufacturer's instructions. 1 µg of total RNA in 10 µl $ddH_2O$ was incubated at 65°C for 5 minutes in the presence of $oligo(dT)_{18}$ primers allowing it to bind to polyA tails of mRNA's. Subsequently, the reverse transcriptase master mix was added and the samples were incubated at 42°C for 1h before inhibiting the reaction at 65°C for 5 minutes (**Table 3.2**).

**Table 3.2: Reverse transcriptase Reaction mix**

| RT Reaction mix | Concentration |
|---|---|
| Reaction buffer | 1x |
| dNTPs | 1 mM |
| RNase inhibitors | 0,5 U/µl |
| $dH_2O$ | Ad 20 µl |
| RevertAid Reverse Transcriptase | 20 U/µl |

### 3.2.3 Quantitative Real-Time PCR

In order to test the transcription levels of respective genes via RT-PCR, cDNA was mixed in a dilution of 1:20 with 10 µM primer pair and Sybr Green according the manufacturer's instructions (**Table 3.3**). The final volume was adjusted to 10 µl.

**Table 3.3: Primers for RT-PCR**

| | forward primer sequence (5' to 3') | reverse primer sequence (5' to 3') |
|---|---|---|
| *Ifn-β* | TCAGAATGAGTGGTGGTTGC | GACCTTTCAAATGCAGTAGATTCA |
| *tnf-α* | CAAAATTCGAGTGACAAGCCTG | GAGATCCATGCCGTTGGC |
| *cycb* | CAGCAAGTTCCATCGTGTCATCAAGG | GGAAGCGCTCACCATAGATGCTC |
| *cxcl2* | ACATCCAGAGCTTGAGTGTGA | GCCCTTGAGAGTGGCTATG |
| *saa3* | CTGGGCTGCTAAAGTCATCA | TGAGTCCTCTGCTCCATGTC |
| *csf3* | TCCTGCTTAAGTCCCTGGAG | GACACAGCTTGTAGGTGGCA |
| *tgtp* | CCCTAAGAGGAAAGCCATCA | CATGGCTCTGTATGGTAGAAGC |
| *serpina3g* | AGAGACCCTGAGGAAGTGGA | GTCAGCCTGTGTGGAGAAGA |
| *cxcl11* | CAGGAAGGTCACAGCCATAG | GCTTTCTCGATCTCTGCCAT |
| *themis2(#1)* | AGCCAGGGTATCAATAAGAAACAG | GTTCTACAACTTGAGGCTTTACAC |
| *themis2(#2)* | GCCAGGGTATCAATAAGAAACAG | GTTCTACAACTTGAGGCTTTACAC |
| *rbms2(#1)* | GAGTAGCAGCACCTTCTGAC | TTGGCTCTGTCGTTTCTTTGG |
| *rbms2(#2)* | ATGGCCTTGACCTATGACCC | CAGAACTGAGCCCTGATAGGA |
| *sbno1(#1)* | ACAGTTCACCAAGAGATAGTCC | GATTTCTTCACCTTTCCGCT |
| *sbno1(#2)* | AAATGTGGCTGAGATGACTGG | ATGAACCTTTGCTTCTCTGTG |
| *fiz1(#1)* | CTAAGGGATTCCGAGACTCCA | GCTCACAGACCAAACAGCAG |
| *fiz1(#2)* | CTTCAACTTGGCTAACCACC | ACAGACCAAACAGCAGTAGG |
| *ecsit(#1)* | TCACATCCTAAGAGCTGACCTG | GCCTTCTGTCACTTCATCCA |
| *ecsit(#2)* | TAAGGTCACTGTCTACCAGATGTC | TCTCCTCTACTTTCTCTTCCTCAG |
| *preb(#1)* | GACAGTTCCACCTTCTTGCC | TACCCGAATCACTGACACTG |
| *preb(#2)* | CAACCATGATAACACCCTGCT | TGTAACCAGCTTGCCATCAG |
| *rsad2(#1)* | CTAACCAGAAGATGAAAGACTCC | ACGCCAACATCCAGAATAGAC |
| *rsad2(#2)* | TTGAAACATTCTTGGAGCGT | AGGAGTCTTTCATCTTCTGGT |
| *zbp1(#1)* | CAAAGAAGTGAACCCACTCC | CCAGAATGAGCTATGTCTTGG |
| *zbp1(#2)* | GTAACGGCAACAAGATGACC | GAGCTTGTACCTGTGTCTTCC |
| *srbd1(#1)* | TCATTTCTCCTACTAGTCAGATCC | GCAGTTCCATTTCCGATCAC |
| *srbd1(#2)* | CTTTCTTCTCTGAGTTGTCATCTG | AGGCTGTCTACGATTTCTGG |
| *rbms1(#1)* | AGATGGAGAGGCTGGAATGAC | TAGGGTGTAAGAGAAGTTTGAGTG |
| *rbms1(#2)* | AGTTTCTGCTCCTACAGAACCT | CTGTAGTCGGGTCATAAGTGAG |
| *tbk1(#1)* | AATACATCCACGCTATGATCTG | CTTAACCAGTTCAACCAGCC |
| *tbk1(#2)* | TAGTCTTTCTCAGGGTCTTCAGG | AAGCACATCACTGGTCTCTG |
| *Cyclophilin B* | CAGCAAGTTCCATCGTGTCAAGG | CTCGTAGATACCACTCGCGAAGG |

Subsequently, RT-PCR reactions were carried out using the Rotorgene 6500 (Corbett) using the following light cycler program:

**Table 3.4 Light cycler program**

| Steps | Temperatur [°C] | Duration |
|---|---|---|
| Hold | 95 | 10 min. |
| Cycling (x40) | 94<br>60<br>72 | 60 sec.<br>15 sec.<br>30 sec. |
| Hold | 25 | 30 sec. |
| Melt | from 65 to 95 | 30 min. |

Each sample was normalized against Cyclophilin B, a housekeeping gene.

## 3.2.4 SDS-PAGE

The proteins samples were diluted with Lämmli buffer (4x) and boiled at 100°C for 3 minutes. The samples were electrophoresed in a discontinuous SDS-polyacrylamide gel and the proteins are separated according to their sizes. Later, the stacking gel was removed and the separation gel was used for Western blot.

**Table 3.5: The composition of the stacking gel**

| | 5% gel |
|---|---|
| Stacking Gel Buffer (4x) | 2,5 ml |
| Acrylamid 30% | 1,6 ml |
| $H_2O$ | 5,8 ml |
| TEMED | 15 µl |
| Ammonium peroxo disulfate 10% | 100 µl |

**Table 3.6: The composition of the running gel**

| | 8% gel | 10% gel |
|---|---|---|
| Running Gel Buffer (4x) | 2.50 ml | 2.50 ml |
| Acrylamid 30% | 2.67 ml | 3.33 ml |
| $H_2O$ | 4.83 ml | 4.17 ml |
| TEMED | 10 µl | |
| Ammonium peroxo disulfate 10% | 100 µL | |

| | |
|---|---|
| Running Gel Buffer (4x) | 1.5 M Tris-HCL pH 8.8<br>10% SDS |
| Stacking Gel Buffer (4x) | 0.5 M Tris-HCl pH 6.8<br>10% SDS |
| Lämmli Sample Buffer (4x) | 20 ml 0.5M Tris-HCl pH 6.8<br>20 ml 20% glycerol<br>4 g SDS<br>10 mL water (pro analysi)<br>40 mg bromphenol blue<br>100 µl β-mercaptoethanol (14.29 M) |
| SDS Running Buffer (5x) | 250 mM Tris,<br>1.9 M Glycin<br>35 mM SDS<br>Add water to 20 L |

### 3.2.5 Western blot

The separated proteins from the SDS-PAGE are transferred by blotting to a nitrocellulose membrane. The transfer was performed in semi-dry conditions (1 mA/cm$^2$ for 1.5h).

| | |
|---|---|
| Western blot buffer (1x) | 6.07 g Tris<br>28.84 g Glycine<br>10% Methanol<br>Add water to 2 L |

Once the proteins are transferred to the nitrocellulose membrane, unspecific binding sites were blocked by incubating the membrane with blocking solution for an hour. Then, membranes were incubated with the primary antibody diluted in the blocking solution for two hours. After that, the membranes were washed the three times with PBS-T and the antibody-protein complex on the membrane was detected by incubating with a secondary antibody conjugated with horseradish peroxidise for an hour.

| | |
|---|---|
| PBS-T | 0.05% Tween 20 in PBS |
| Blocking Solution | 5% Nonfat dry milk in PBS-T |

After rinsing the membrane three times with PBS-T, it was developed using ECL Western blotting Reagent according to the manufacturer's instructions.

### 3.2.6 ELISA

Cell culture supernatants were collected as described above. Supernatants were cleared by centrifugation. The concentration of IFN-β and TNF-α in these supernatants was determined using an IFN-β specific ELISA or a TNF-α specific ELISA according to the manufacturer's instructions.

### 3.2.7 Phosphorylation of DNA

In order to determine the susceptibility of the ISD variants for 5' phosphorylation, recombinant GST-hClp1 [Weitzer and Martinez 2007] was added at a final concentration of 2 µM to reaction mixtures (100 mM KCl, 5 mM $MgCl_2$, 10 mM DTT, 0.1 µM [γ-$^{32}$P] adenosine 5' triphosphate and RNasin) containing 100 nM ISD-OH, ISD-P, ISD-Bio and ISD-NH2. The reactions were incubated at 30ºC for 30 or 60 minutes, the reaction products were separated on a 15% denaturing acrylamide gel and analyzed by Phosphorimaging**.**

### 3.2.8 Microarray

The microarray analysis was carried out by Martin Bilban from the Department of Laboratory Medicine, Medical University of Vienna.

Microarray analysis from poly(dAdT) and CpG stimulated RAW264.7 cells: RAW264.7 cells were stimulated with poly(dAdT) or CpG for 4h. cRNA Synthesis and Gene Expression Profiling Total RNA was isolated per manufacturer's instructions (RNeasy mini kit (Qiagen, Valencia, CA). Preparation of cRNA, hybridization to murine 430 2.0 GeneChips (Affymetrix, Santa Clara, CA), and scanning of the arrays were carried out according to manufacturer's protocols (https://www.affymetrix.com). Images were analyzed with GeneChip software (Affymetrix, version 5.0).

Microarray analysis from IFN-β treated RAW264.7, L929, NIH3T3 cells: The respective cell lines were treated with IFN-β for 4h. The total RNA was isolated with SV Total RNA isolation system according the manufacturer's instructions. Using 200 ng total RNA GeneChip analysis was carried out. Terminally labeled cDNA was prepared and hybridized to genome-wide GeneChip Mouse Gene 1.0 ST arrays. The arrays were scanned according to

manufacturer's instructions. For signal extraction and normalization, the RMA algorithm was used [Irizarry et al. 2003].

### 3.2.9 Isolation of PBMCs

The buffy coat was obtained from Austrian Red Cross. The PBMCs were isolated using a density gradient centrifugation approach taking advantage of different densities of different cell types. The buffy coat is diluted 1:2 with PBS and every 35 ml of the diluted sample is layered over 15 ml of Lymphoprep in a centrifuge tube without mixing. The mixture is centrifuged at 400 g for 25 minutes at 21°C without brake . After centrifugation the PBMCs form a distinct band at the interface between the plasma phase and the high density phase. The PBMCs were removed carefully using a Pasteur pipette. The PBMCs were collected and transferred to a new centrifugation tube and washed with PBS. The pellet is resuspended in 2ml cold erythrocytes-lysis buffer (0.15 M $NH_4Cl$, 10 mM $NaHCO_3$ and 1 mM EDTA, pH7.3) to remove the remaining erythrocyte contamination and washed twice with PBS.

### 3.2.10 Cytosolic extract

Respective cell lines were treated with 100 U IFN-β over night and were harvested, centrifuged and resuspended in 3 volumes of Buffer N. After briefly vortexing, the suspension is left on ice for 3 minutes and centrifuged at 4°C and 500 g for 5 minutes. The supernatant contains the cytosolic extract. For nuclear extract, the pellet is washed once with buffer N and resuspended in 3 volumes of Buffer C420. The suspension is vortexed vigorously at 4°C for 10-15 minutes and centrifuged for 5 minutes at maximum speed. The supernatant contains the nucleic extract.

| | |
|---|---|
| Buffer N | 10 mM HEPES (pH 7.9) |
| | 10 mM KCl |
| | 0.1 mM EDTA |
| | 0.1 mM EGTA |
| | 1 mM DTT |
| | 300 mM Sucrose |
| | 0.75 mM Spermidine |
| | 0.15 mM, Spermine |
| | 0.1% (w/v) NP-40 substitute |
| | 50 µM NaF |

100 µM sodium orthovanadate
Protease inhibitors

Buffer C420
20 mM HEPES (pH7.9)
420 mM NaCl
1 mM EDTA
1 mM EGTA
1 mM DTT
25% glycerol
50 µM NaF
100 µM sodium orthovanadate
Protease inhibitors

### 3.2.11 Nucleic acid affinity purification

Biotinylated sense ISD was incubated at 65°C with antisense ISD in TE-buffer in the presence of 5 M NaCl for 20 minutes and allowed to cool to room temperature to obtain final concentration 20 µM double stranded oligonucleotides. The 120 µl of 50% slurry Streptavidin beads were washed 3 times with PBS. 2,5 nmol biotinylated ISD was added and incubated at 4°C for 2h on a shaker to allow binding of the biotinylated nucleotide to the resin. In case of poly(I:C)-coupled beads, an aliquot of lyophilized poly(C)-agarose was dissolved in 400 µl PBS creating a 50% slurry suspension. The suspension is incubated on ice for 15 minute allowing the resin to swell followed by the wash of the beads twice in a buffer containing 50 mM Tris/HCl pH7.5 and 100 mM NaCl. After the wash the poly(C)-agarose was resuspended in 2 ml of wash buffer. On the other hand, poly(I) is diluted in PBS to a final concentration of 2 mg/ml. 2 ml of poly(I) were added to 1 ml of poly(C)-agarose and were incubated on the rotary wheel at 4°C for 2h. Following this step, all the coupled beads were washed 3 times with buffer N containing 150 mM NaCl. The prepared cytosolic extracts from RAW264.7 and PBMC's were diluted to final protein concentration of 5 mg/ml. 5 mg extract was incubated with ISD-PO, ISD-PS or CT-DNA coupled beads at 4°C for 2 hours in the presence of 10 µg/ml poly(I:C) as competitor or incubated with poly(I:C) coupled beads in the presence of 10 µg/ml calf thymus DNA as competitor. The incubated resins were transferred to 5 ml biospin columns and washed with 10 ml Buffer N allowing the buffer to enter the column by gravity flow. Subsequently, the columns were dried by centrifugation and bound proteins were eluted

by boiling the beads in 50 µl SDS sample buffer (50 M Tris, 380 mM Glycin, 7 mM SDS). In case of the purifications from DNA transfected cells, the RAW264.7 cells were stimulated with 1 µg/ml biotinylated ISD for 4h prior to cytosolic extract preparation and incubated with empty streptavidin beads in the presence of 10 µg/ml poly(I:C) during the purification.

Subsequently, the eluates were analysed by the mass spectrometry department at CeMM.

### 3.2.12 Mass spectrometry analysis

Mass spectrometry analysis of the eluates was carried out in the mass spectrometry department at CeMM by Keiryn Bennett and Melanie Planyavsky. Samples were analyzed by LC-MS/MS and the results were searched in the human Swiss prot (ver.2010.09_20100812) containing 35149 sequences and mouse Swiss prot (ver.2010.09_20100812) containing 24048 sequences with the search engine MASCOT and Phenyx as described in Buerckstuemmer et al. 2009 with following modifications: high confidence peptide identifications are used to recalibrate all precursor and fragment ion masses before a second search with narrower mass tolerances (30 p.p.m. and 0,1 Da).

### 3.2.13 Cloning of shRNA vectors

In collaboration with the laboratory of Louis Staudt from *National Institutes of Health, National Cancer Institute, Bethesda, USA*, six shRNA vectors per DNA sensor candidate were designed and cloned (For the sequences, *see Appendix*) [Reynolds et al. 2004]. Out of 46 candidate proteins, Ifi204, which is the mouse homologue of IFI16, was the only protein where the algorithm failed to provide six shRNAs. Only in this case, we decided to work with two shRNAs provided by the algorithm instead of six.

62-mer RNAi nucleotides were synthesized at Invitrogen (UK). The sense and antisense strands are mixed together with an annealing buffer, heated up to 95°C for 2min and cooled down to room temperature allowing the nucleotides to anneal obtaining a final concentration of 200ng/µl. The annealed nucleotides were diluted in $H_2O$ to a concentration of 0.05 ng/µl. The diluted annealed nucleotides were ligated into pRSMX_PG. After that, the ligation reaction was transformed into TOP10 *E.coli* by heatshocking. After incubation in SOC Medium at 37°C for 1h, the *E.coli* was plated onto 12-lane trays. The next day, 2 to 3 colonies were picked and inoculated in LB-Medium with ampicillin. After 24h, 2 µl of

transformed bacterial cultures were mixed with the PCR mix (**Table 3.7**) and a PCR-run was performed in order to screen the presence of the insert (**Table 3.8**).

**Table 3.7. PCR-mix**

| PCR Mix | Concentration/Volumes |
|---|---|
| Bac culture | 2 µl |
| pMSCV-5 | 4 pmol |
| pMSCV-3 | 4 pmol |
| dNTP mix | 125 µM |
| Taq buffer | 1x |
| Taq polymerase | 1.25 U |
| $H_20$ | ad 27 µl |

**Table 3.8: PCR protocol**

| Steps | Temperature [°C] | Duration |
|---|---|---|
| Hold | 95 | 10 min |
| Cycling (29x) | 94 | 20 sec |
| Hold | 72 | 5 min |

The PCR products were loaded on an agarose gel and the right size of the insert was tested under the UV-light, the samples with the right size were sequenced by Microsynth (Switzerland).

### 3.2.14 Loss-of-Function Validation

1ml transfection mix per well containing VSV-G and the respective shRNA-vectors was prepared using Lipofectamine 2000 according the manufacturer's instructions in OptiMEM® Reduced Media and pipetted on empty 6-wells. Producer cell line 293gp cells were seeded on the same 6-wells at a density of $1x10^6$ using the OptiMEM® Reduced Media on top of the transfection mix. After 4-6 hours, 1ml DMEM medium (containing 10% FCS and 1% Pen/Strep) was added to each well. The next day, the medium of 293gp cells was changed and RAW264.7 cells were seeded in 12-well plate at a density of $1x10^5$ per well. After 24h, the medium of RAW264.7 cells was changed and 1 ml DMEM (+ 10% FCS, 1% P/S, 8 µg/ml polybrene) was added on the wells. Subsequently, the supernatant of 293gp cells were transferred onto RAW264.7 cells and incubated for 24h at 37°C. The next day, the cells were selected with DMEM containing 10% FCS, 1% P/S and 7.5 µg/ml Puromycin until 70% of the

cells reach GFP-positivity. At this point, the cells were stimulated using 1 µg/ml ISD as described in the section *DNA transfection.* After 4h of stimulation, the cell viability was measured using CellTiter-Glo® Luminescent Cell Viability Assay according the manufacturer's instructions.

### 3.2.15 Bioassay

LL171 cells were seeded in 96-well-plates at a density of $4x10^4$. The next day, 5 µl of the supernatant from DNA stimulated RAW267.4 cells were added on LL171 cells and incubated for 5h at 37°C and 5% $CO_2$. Subsequently, the cells were lysed with 50 µl passive lysis buffer for 15 minutes at RT and 20 µl of the lysate is mixed with 50 µl of Luciferase Assay Substrate buffer and the luminescent signal was measured.

| | |
|---|---|
| Luciferase Assay Buffer | 270 uM CoenzymeA |
| | 470 uM d-Luciferin |
| | 530 uM ATP |
| | 20 mM Tricine |
| | 0.1 mM EDTA |
| | 3.74 mM $MgSO_4$ |
| | 33.3 mM DTT |

# 4 Results

## 4.1 Different DNA ligands elicit a common cytosolic DNA recognition pathway in murine macrophages

### 4.1.1 Raw264.7 macrophages as model organism

In order to study the structural requirements for DNA recognition in the cytosol, we chose to analyze DNA-mediated interferon-β (IFN-β) production in RAW264.7 macrophages. These cells are mouse leukaemic monocyte macrophages [Raschke et al. 1978]. Due to their ability to produce high levels of IFN-β in response to LPS and moderate levels of IFN-β in response to CpG-DNA and poly(I:C) treatment (**Fig. 4.1**), they are widely used in innate immunity research. Transfection of RAW264.7 cells with ISD or poly(dAdT) triggered the production of high levels of IFN-β, comparable to those achieved by LPS treatment.

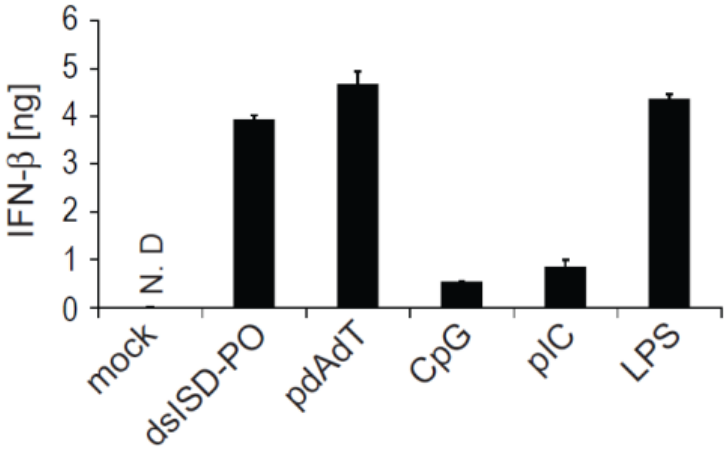



**Fig. 4.1: Transfection of innate immunity ligands leads to IFN-β production in RAW264.7 cells.** RAW264.7 cells were mock transfected or transfected with dsISD-PO (1µg/ml) or poly(dAdT) (1µg/ml). In parallel, RAW264.7 cells were treated by addition of 1µg/ml LPS, 10µg/ml poly(I:C) or 20µM CpG to the medium. After 4h, supernatants were collected and IFN-β concentrations determined by ELISA.

Furthermore, IFN-β production elicited by ISD or poly(dAdT) was concentration-dependent (**Fig. 4.2**), with ISD being slightly more potent than poly(dAdT). As robust IFN-levels were detectable at a concentration of 1µg/ml ISD (**Fig. 4.2**), we decided to use this concentration for all the following experiments.

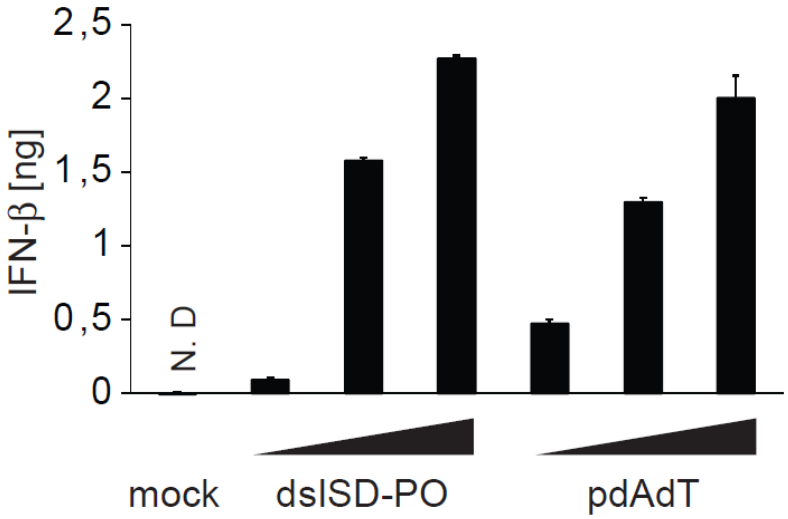



**Fig. 4.2: Transfection of ISD leads to concentration-dependent IFN-β production** RAW264.7 cells were mock treated or transfected with dsISD-PO (10ng/ml, 100ng/ml, or 1000ng/ml) or poly(dAdT) (250ng/ml, 1µg/ml, or 2µg/ml). After 4h, cell culture supernatants were collected and IFN-β concentrations determined by ELISA. Data are shown as mean ± SD of two replicates in one representative experiment out of three independent experiments. N.D: not detectable.

### 4.1.2 ISD and poly(dAdT)-mediated responses are TLR9 independent

RAW264.7 cells are macrophage-like cells that can respond to DNA stimulation by the means of TLR9. In general, TLR9 recognizes CpG motifs present on bacterial DNA. As shown by *Hacker* and colleagues, CpG-DNA can only induce a response if endocytosed and bound to TLR9 in the mature endosomes. Although ISD contains no contiguous CpG motifs, we wanted to rule out the possibility that TLR9 contributes to ISD recognition by applying ISD in the absence of transfection reagents [Stetson and Medzhitov 2006, Takaoka et al. 2007]. This treatment did not trigger IFN-β production in RAW264.7 cells (**Fig. 4.3 A**). For TLR9 dependent pathway activation, acidification of the endosome and subsequent proteolytic cleavage by cathepsins are two prerequisites which can be blocked by chloroquine or Z-FA-FMK, respectively [Park et al. 2008]. CpG and LPS trigger the production of interferon and, subsequently, of interferon-inducible genes. We therefore decided to use DAI (DLM-1/ZBP1), one of the most strongly interferon-induced genes as an indirect read-out for the interferon response [Takaoka et al. 2007]. In agreement with plasma membrane localization of TLR4, LPS-mediated DAI production was unaffected by chloroquine or Z-FA-FMK (**Fig. 4.3 B**). CpG-mediated DAI production, however, was completely abrogated in the presence of either of the two inhibitors (**Fig. 4.3 B**), implying that CpG recognition occurs in the endosome and

hence through TLR9. ISD and poly(dAdT) signaling, in contrast, were unaffected by chloroquine or Z-FA-FMK (**Fig. 4.3 B**).

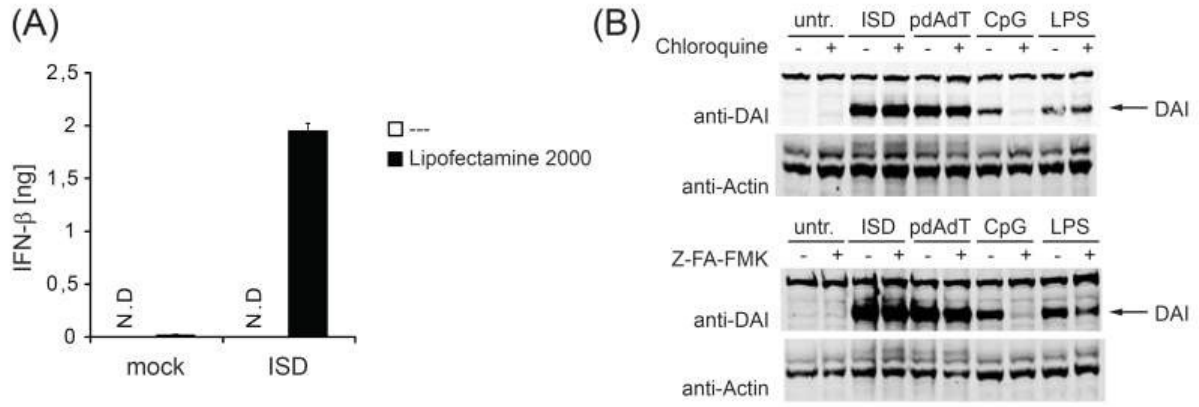



**Fig. 4.3: ISD and poly(dAdT)-mediated responses are TLR9 independent.** (A) RAW264.7 cells were either mock treated or treated with ISD in the presence/absence of Lipofectamine 2000. After 4h, the cell culture supernatants were collected and IFN-β concentrations determined by ELISA. N.D: not detectable. (B) RAW 264.7 cells were either left untreated or were treated with 10μM Chloroquine or 50μM Z-FA-FMK for 1h. Subsequently, the cells were either left unstimulated or stimulated with ISD, pd(dAdT), CpG or LPS. Cells were harvested 24h after stimulation and DAI and Actin levels were analyzed by immunoblotting.

Similar results were obtained when directly assaying for IFN-β or TNF-α by ELISA (**Fig. 4.4**), indicating that the recognition of ISD and poly(dAdT) occurs in the cytosol in a TLR9-independent manner.

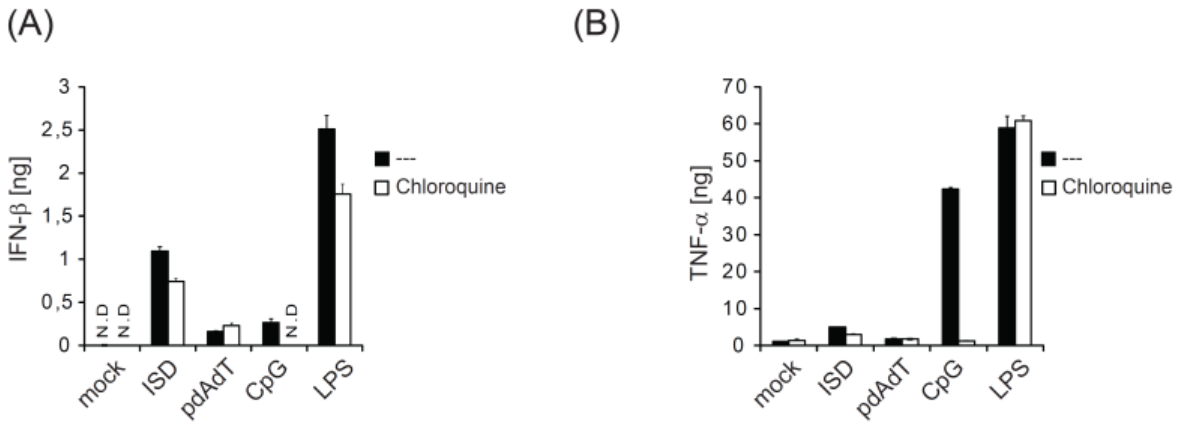



**Fig. 4.4: ISD and poly(dAdT)-mediated responses are endocytosis independent.** RAW 264.7 cells were either left untreated or were treated with 10μM Chloroquine. Subsequently, the cells were either left unstimulated or stimulated with ISD, poly(dAdT), CpG or LPS. Cell culture supernatants were collected 4h after stimulation and (A) IFN-β and (B) TNF-α concentrations determined by ELISA. Data are shown as mean ± SD of two replicates in one representative experiment out of two independent experiments.

### 4.1.3 Modification of 5'end does not affect DNA recognition in the cytosol

ISD, unlike other DNAs that are commonly used to study DNA-induced IFN-β production, is well-defined in terms of length (45bp), sequence (no CpG motifs), ends (5' and 3' hydroxyl group, no 5' or 3' overhangs) and etiology (generated by chemical synthesis and therefore unmodified and free of biological contaminants). Since it was recently shown that RIG-I recognizes 5' triphosphate-containing RNAs [Pichlmair et al. 2006, Hornung et al. 2006], we raised the hypothesis that the 5' end is a critical determinant of cytosolic DNA recognition. To test this hypothesis, we generated four different ISD variants: (i) ISD containing a free 5' hydroxyl group (ISD-OH), (ii) ISD *in vitro* phosphorylated at the 5' end (ISD-P; [Weitzer and Martinez 2007]), (iii) ISD blocked at the 5' end by addition of a biotin group (ISD-Bio) and (iv) ISD blocked at the 5' end by addition of an amino group (ISD-NH2). We opted for two blocked variants to exclude any effect of one particular blocking group. All ISD variants migrated as a distinct band of ~50bp in an agarose gel (**Fig. 4.5**), indicating proper annealing. As expected, only ISD-OH could be phosphorylated by the nucleic acid kinase Clp1 *in vitro* [Weitzer and Martinez 2007], since ISD-P has already been phosphorylated and ISD-NH2 and ISD-Bio were blocked at the 5' end (**Fig. 4.5**).

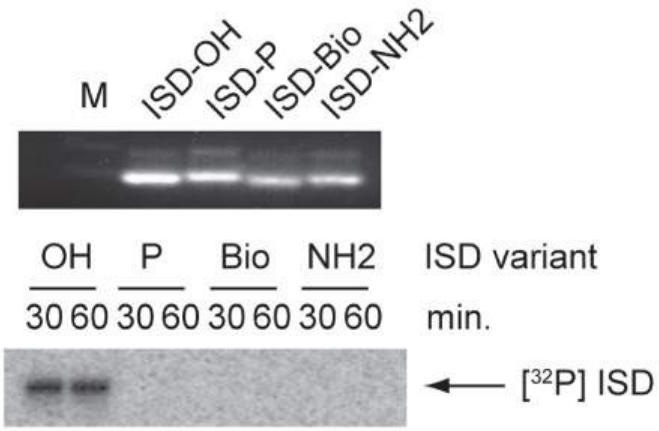



**Fig. 4.5: ISD-OH is the only oligonucleotide to be phosphorylated by Clp1 *in vitro*.** ISDs modified as indicated were electrophoresed on a 3% agarose gel (upper panel) or phosphorylated *in vitro* by addition of the nucleic acid kinase Clp1 in the presence of [$\gamma$-$^{32}$P]ATP for the indicated times. Phosphorylated ISDs were separated on a 15% denaturing acrylamide gel and visualized by autoradiography (lower panel).

Next, we addressed whether the modification of the 5' end affects DNA recognition in the cytosol. To this end, we transfected RAW264.7 cells with two concentrations of the four ISD variants and measured IFN-β induction. Again, we observed a strong concentration-dependent induction of IFN-β, both at the level of the mRNA (**Fig. 4.6 A**) and the secreted

protein (**Fig. 4.6 B**). Contrary to our hypothesis, modification of the 5' end did not affect ISD-mediated IFN-β production (**Fig. 4.6 B**). Similar results were obtained for ISD induction of TNF-α (**Fig. 4.6 C, D**), suggesting that all signaling pathways induced by short synthetic DNA are subject to the same structure–function relationship.

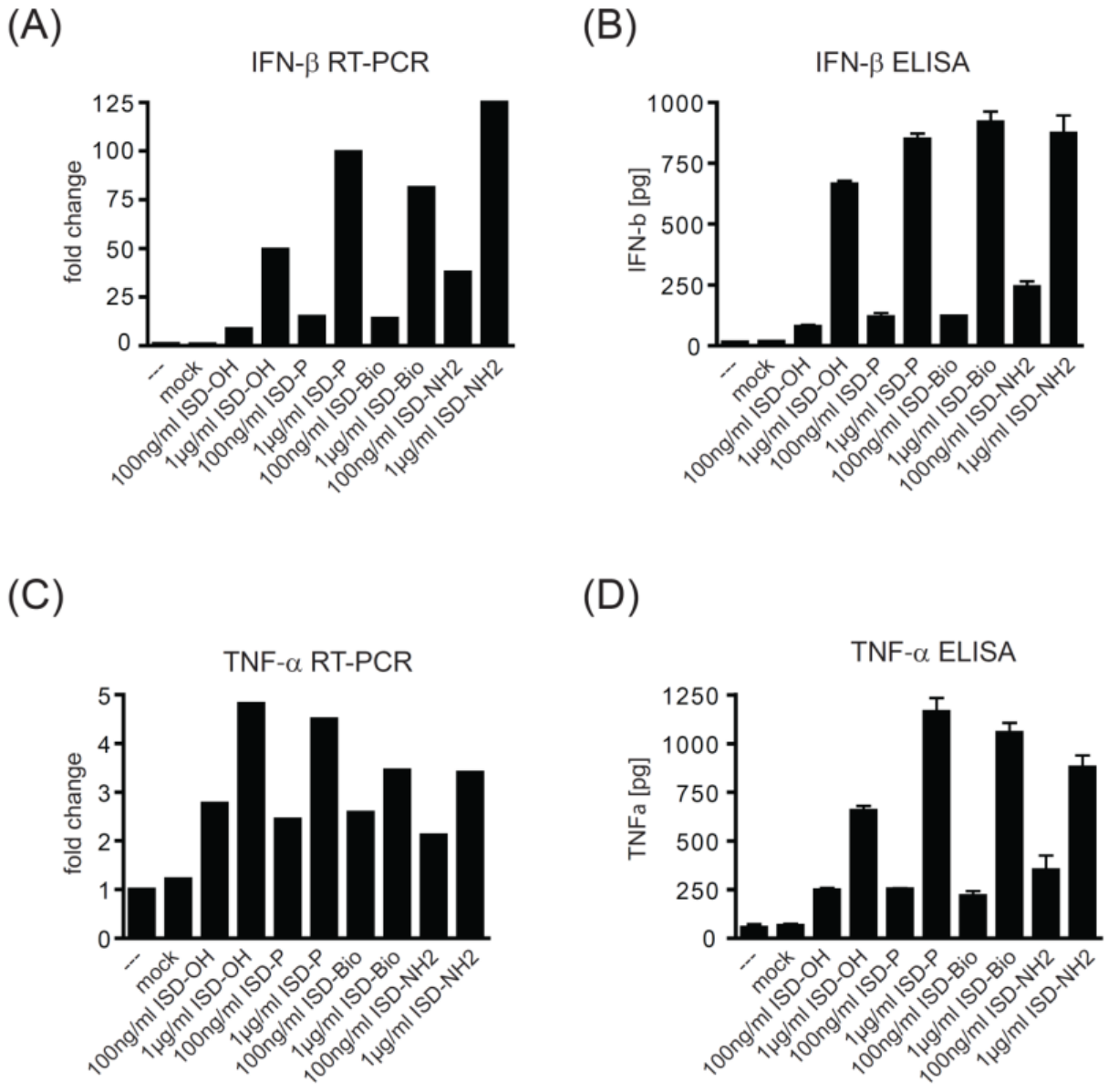



**Fig. 4.6: Modification of 5'end does not affect DNA recognition in the cytosol.** (A-D) RAW264.7 cells were transfected with ISDs modified as indicated at concentrations of 100ng/ml or 1µg/ml. After 4h, the RNA was extracted and (A) IFN-β and (C) TNF-α mRNA levels were analyzed by quantitative RT-PCR. Cell culture supernatants were also collected and (B) IFN-β and (D) TNF-α concentrations determined by ELISA. Data are shown as mean ± SD of two replicates in one representative experiment out of three independent experiments.

### 4.1.4 The DNA-mediated IFN-β response is length dependent

Having shown that DNA recognition does not depend on the integrity of its 5' end, we assessed the minimal length required for efficient IFN-β production. To this end, we created double-stranded ISD oligonucleotides of different sizes (from 45 to 15bp, as indicated in **Fig. 4.7**). Proper annealing was monitored by agarose gel electrophoresis (**Fig. 4.7 A**). DNA oligonucleotides larger than 24bp were potent activators of IFN-β production (**Fig. 4.7 B**). In contrast, we observed a significant reduction in the stimulatory potential when ISD was shrunk to less than 24bp, with ISD of 15bp being essentially inactive (**Fig. 4.7 B**). This suggests that DNA requires at least two helical turns to be recognized by cytosolic sensors.

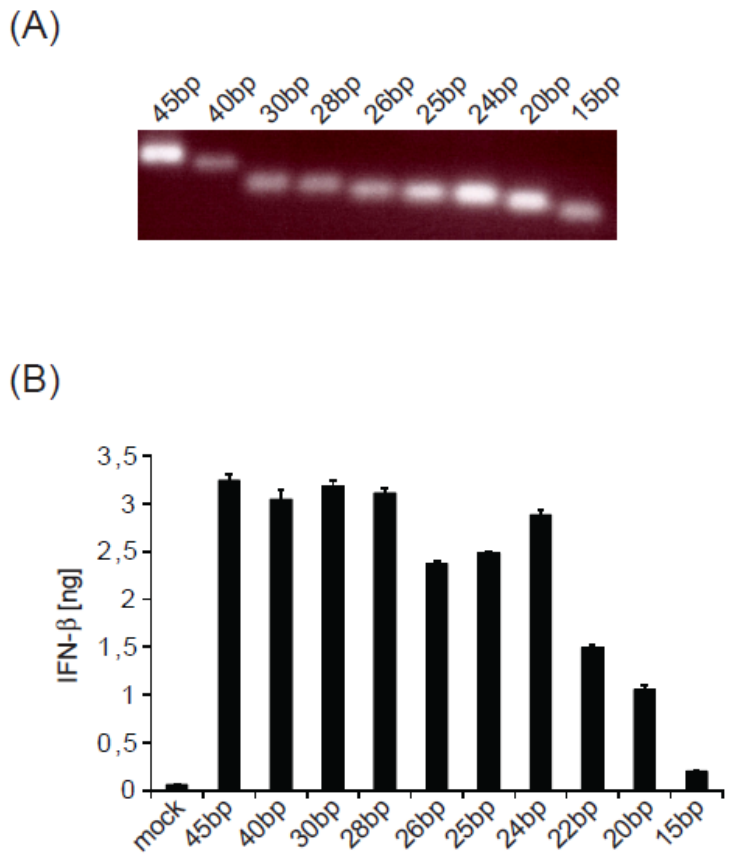



**Fig. 4.7: The DNA-mediated IFN-β response is length dependent.** (A) ISDs of different lengths were electrophoresed on a 3% agarose gel. (B) RAW264.7 cells were transfected with 1µg/ml ISD of indicated lengths and the cell culture supernatants were collected after 4h and IFN-β concentrations determined by ELISA. Data are shown as mean ± SD of two replicates in one representative experiment out of three independent experiments.

### 4.1.5 ssISD-PS antagonizes dsISD-PO- and poly(dAdT)-mediated responses

ISD and poly(dAdT), though both meant to trigger the cytosolic DNA pathway, are likely to be of different nature and were originally identified in two distinct experimental setups [Ishii et al. 2006, Stetson and Medzhitov 2006]. While ISD is a well-defined double-stranded DNA with blunts ends, poly(dAdT) is heterogeneous in size with random order of dA and dT bases and hence is likely to form a mesh-like structure (data not shown). Furthermore, ISD signaling was studied in murine macrophages [Stetson and Medzhitov 2006], whereas poly(dAdT) signaling was assayed in murine embryonic fibroblasts [Ishii et al. 2006]. We therefore raised the question whether the two species of DNA would be recognized by the same receptor. Notably, both single- and double-stranded ISDs (ssISD-PO, dsISD-PO) activated IFN-β production, while ISDs in which the phosphodiester backbone had been replaced with phosphorothioate (ssISD-PS, dsISD-PS) were completely inactive (**Fig. 4.8 A**), in agreement with previously published results [Stetson and Medzhitov 2006]. Therefore, we assessed whether the inactive ssISD-PS could be used as a competitor. ssISD-PS competed with dsISD-PO for the unknown DNA sensor, leading to impaired IFN-β production with higher concentrations of ssISD-PS (**Fig. 4.8 B**). Strikingly, ssISD-PS also blocked poly(dAdT)-mediated IFN-β production (**Fig. 4.8 C**). This suggests that ISD and poly(dAdT) share a common receptor in RAW264.7 macrophages.

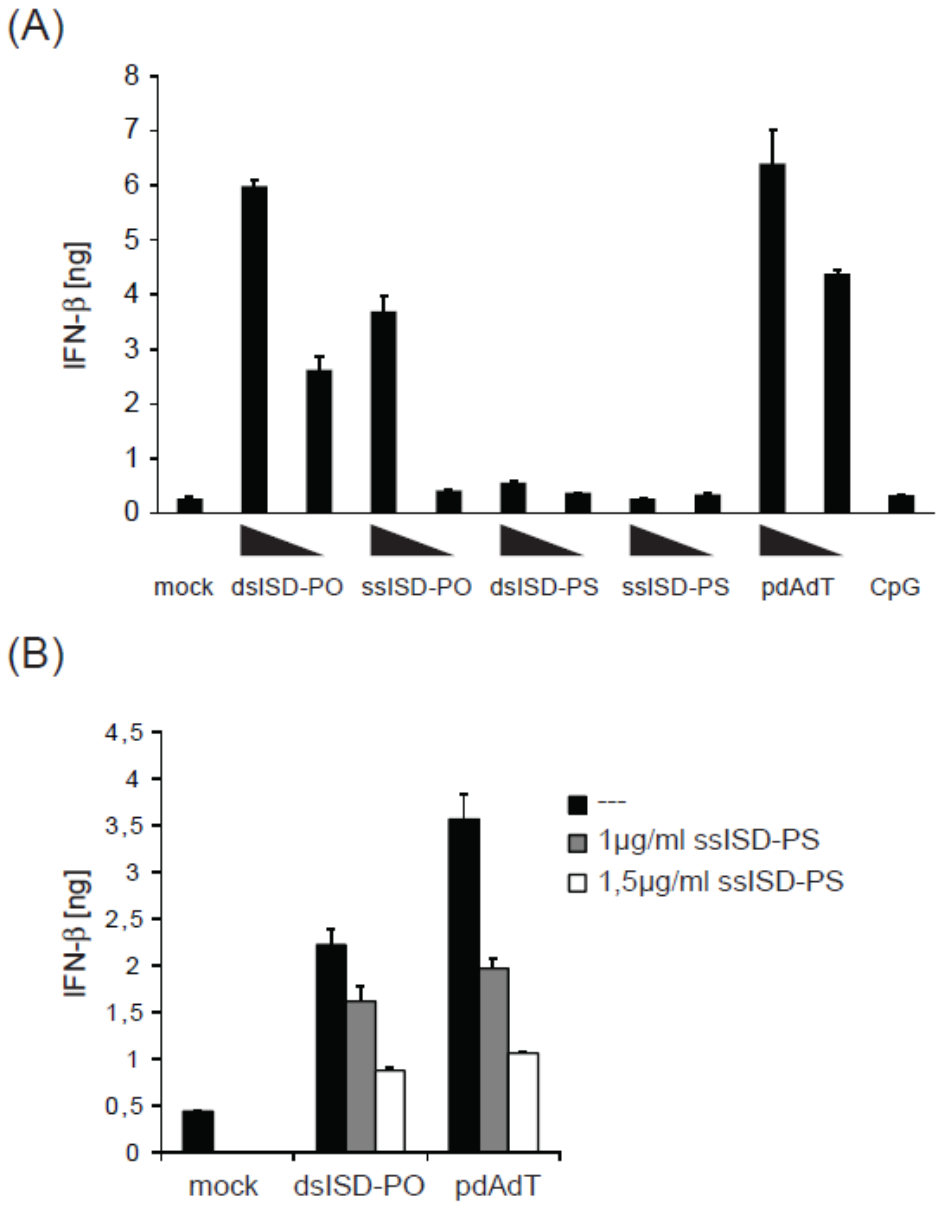



**Fig. 4.8: ssISD-PS antagonizes dsISD-PO- and poly(dAdT)-mediated responses.** (A) RAW264.7 cells were transfected with single- or double-stranded ISDs containing either phosphodiester backbones (ssISD-PO or ds-ISD-PO) or phosphorothioate backbones (ss-ISD-PS or ds-ISD-PS) at concentrations of 1µg/ml or 100ng/ml. As controls, RAW264.7 cells were either mock transfected or transfected with 1µg/ml or 100ng/ml poly(dAdT) or 20µM CpG was added directly to the medium. After 4h, the cell culture supernatants were collected and IFN-β concentrations determined by ELISA. (B) RAW264.7 cells were transfected with 1µg/ml dsISD-PO or poly(dAdT) together with increasing concentrations of ssISD-PS as indicated. After 4h, the cell culture supernatants were collected and IFN-β concentrations determined by ELISA. Data are shown as mean ± SD of two replicates in one representative experiment out of three independent experiments.

### 4.1.6 Distinct gene expression profiles induced by ISD and poly(dAdT) compared to CpG-DNA

Distinct receptors trigger distinct signaling cascades that culminate in the regulation of distinct sets of genes. To understand which pathways are elicited by poly(dAdT), we stimulated RAW264.7 macrophages with poly(dAdT) or CpG-DNA for 4h and analyzed global changes in gene expression by microarray analysis. Clearly, poly(dAdT) and CpG-DNA regulated distinct sets of genes (**Fig. 4.9**).

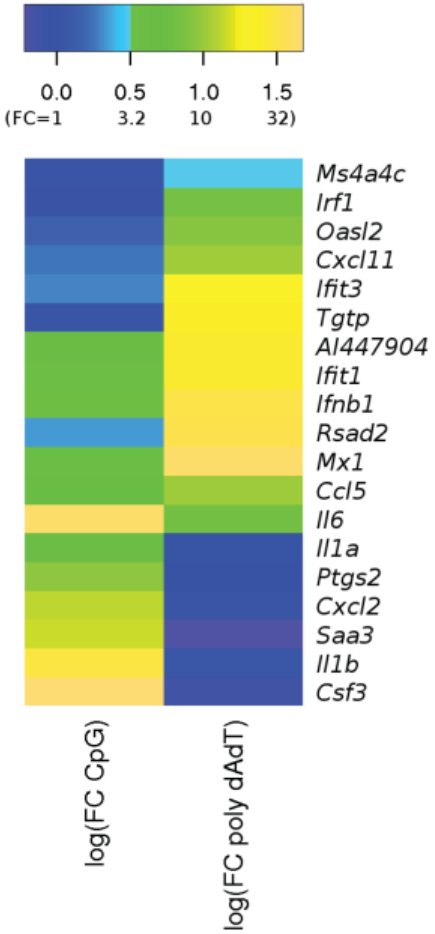



**Fig. 4.9: poly(dAdT) induces a specific gene expression profile that is distinct from the one elicited by CpG-DNA.** RAW264.7 cells were stimulated with CpG-DNA or transfected with 1µg/ml poly(dAdT). RNA was isolated 4h post stimulation and global changes in gene expression were detected by microarray analysis. Clusters of CpG- or poly(dAdT)-regulated genes are represented by a heatmap in logarithmic scale.

Based on the microarray analysis, we selected seven genes that were strongly induced by poly(dAdT) and unaffected by CpG-DNA (*Serpina3a, Ms4a4c, Irf1, Oasl2, Cxcl11, Ifit3, Tgtp*) and six genes that were strongly induced by CpG-DNA and unaffected by poly(dAdT) (*Csf3,*

*IL-1a, IL-1b, Saa3, Cxcl2, Ptgs2*). The microarray data were confirmed by quantitative real-time PCR, thereby defining gene signatures for poly(dAdT) and for CpG-DNA (**Fig. 4.10**). If poly(dAdT) and ISD share the same receptor, they should induce the same set of signature genes. Indeed, ISD transfection lead to elevated expression of poly(dAdT)-induced genes, while CpG-DNA signature genes remained unchanged. This highlights that ISD and poly(dAdT) do not trigger TLR9 activation. Furthermore, it provides strong evidence that ISD and poly(dAdT) trigger the same DNA sensor, leading to a similar gene expression profile.

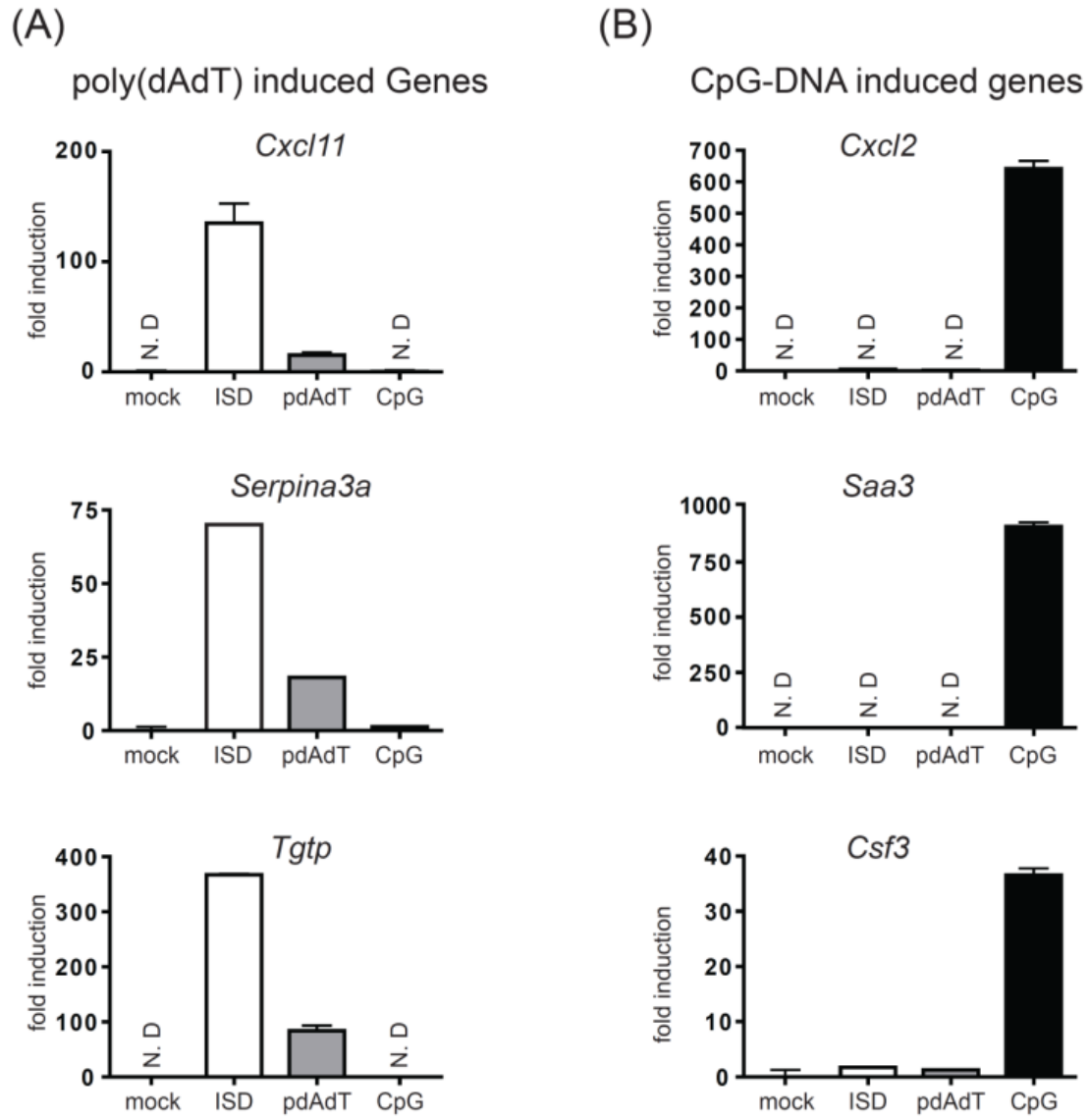



**Fig. 4.10: ISD and poly(dAdT) trigger the same DNA sensor, leading to a similar gene expression profile.** RAW264.7 cells were mock transfected, transfected with 1µg/ml ISD or poly(dAdT) or treated with 20µM CpG-DNA. After 4h, the RNA was extracted and changes in gene expression presented as (A) poly(dAdT)-induced and (B) CpG-induced genes were analyzed by quantitative RT-PCR analysis. Data are shown as mean ± SD of two replicates in one representative experiment out of three independent experiments.

## 4.2 A systematic approach to identify the cytosolic DNA sensor for type I interferon production

We and others showed that cytosolic DNA induces innate immune response independently of TLR9 and the receptor which triggers the production and secretion of proinflammatory cytokines and interferons is still unknown (**Fig. 4.3**) [Stetson and Medzhitov 2006, Ishii et al. 2006]. Furthermore, our gene signature (**Fig. 4.10**) and competition experiments (**Fig. 4.8**) suggested that cytosolic recognition of two different synthetic DNAs (ISD and poly(dAdT)) are qualitatively indistinguishable and differ from the CpG-containing oligonucleotides triggering the TLR9 pathway. These observations implied that, in murine macrophages, only one major IFN-inducing cytosolic DNA recognition pathway is operational.

After characterizing the response in RAW264.7 cells in detail, we aimed at identifying the DNA sensor responsible for type I interferon induction in those cells. At the same time, these cells are of murine origin and their immortalization may have changed the cellular physiology. We therefore decided on peripheral blood mononuclear cells (PBMCs) isolated from human blood as a second source of cells. These cells, although heterogeneous in nature, contain a wide array of immune cells such as NK-cells, B-cells, T-cells, macrophages, and dendritic cells and respond to ISD stimulation with significant IFN-β production (data not shown).

We hypothesized that the elusive DNA sensor should exhibit at least three attributes: Most importantly, it should be able to physically bind to DNA in order to sense it and to activate the downstream pathway. Secondly, the DNA sensor should be in the cytosol as it should sense the cytosolic foreign DNA. Finally, it may be positively regulated by the cytokine it produces, namely IFN-β, as a part of a feedback to regulate the antiviral response. The latter assumption has also been supported by the notion that other nucleic acid receptors such as RIG-I and MDA5 and other signaling components such as IKK-i or IRF7 are transcriptionally regulated in response to IFN-β [Kang et al. 2004, Yoneyama et al, 2004, 2005].

To identify proteins which fulfill these criteria, we took a systematic approach: First, we captured DNA binding proteins from cytosolic extracts of PBMCs and RAW264.7 cells by affinity purifying them using immobilized DNA on a matrix. The eluates from these purifications were analyzed by mass spectrometry. In order to create an “antiviral state” in the cytosol and, following our assumption, to increase the protein levels of the elusive DNA receptor, we treated the respective cell lines with IFN-β prior the cytosolic extraction and purification. In order to identify which of the proteins captured in the affinity purifications are

transcriptionally regulated by IFN-β, we performed microarray analysis from IFN-β treated cells. To complement our microarray data, we extracted interferon-inducible genes from www.interferome.org; a database of IFN regulated genes generated from 28 publicly available microarrays [Samarajiwa et al. 2009]. Finally, we established a scoring scheme to prioritize the candidates to be selected for a validation screen (**Fig. 4.11**).

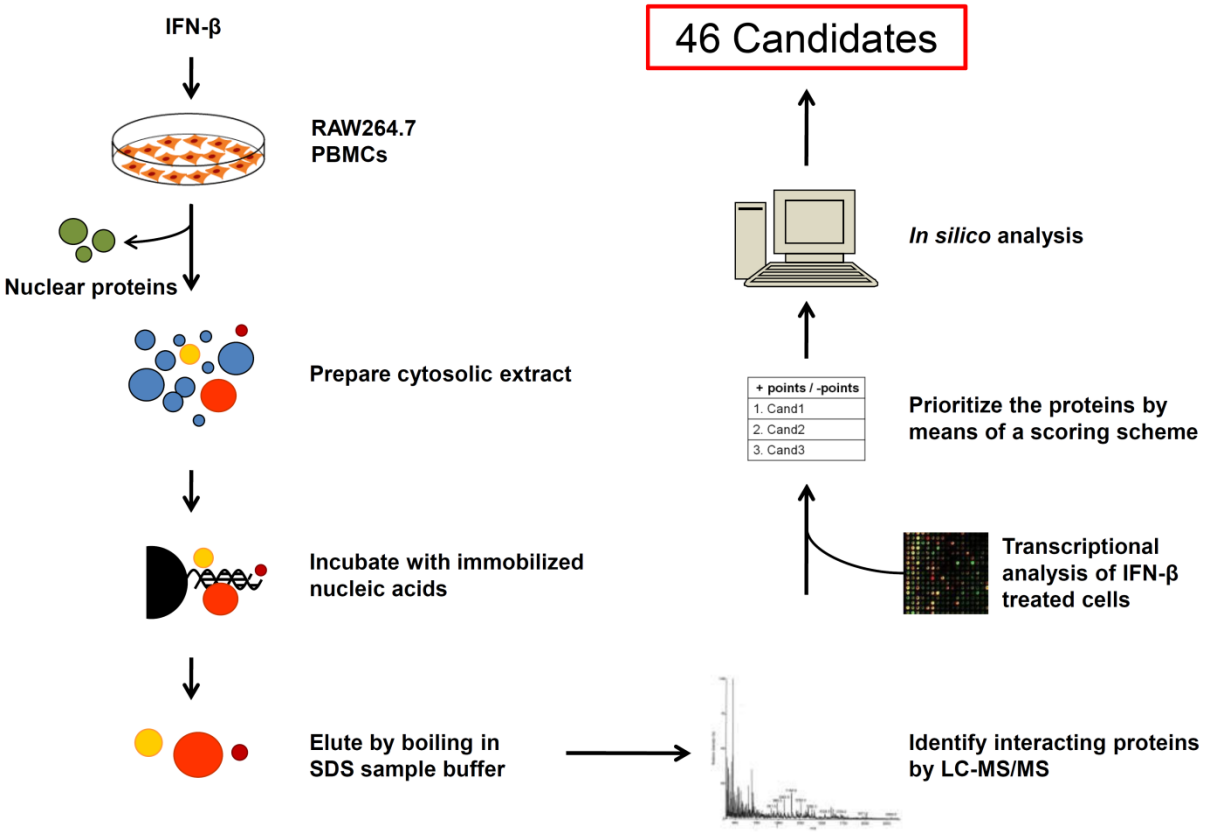



**Fig. 4.11: The systematic approach to identify cytosolic DNA sensor.** Schematic presentation. In order to identify candidate proteins for cytosolic DNA sensor, we combined mass spectrometry analysis of cytosolic extracts incubated with immobilized nucleic acids with microarray data of RNA extracted from IFN-β treated cells.

## 4.3 Identifying DNA binding proteins

### 4.3.1 Cytosolic extraction

Experiments conducted by us and others (Stetson and Medzhitov 2006) showed that ISD is recognized in the cytosol and not in the endosome (**Fig. 4.3**, **Fig. 4.4**) and this recognition does not depend on Toll-like receptors. Based on these two observations, we hypothesized that a cytosolic DNA sensor exists that triggers the production of type I interferons and it is at any time, not necessarily exclusively but at least partly, present in the cytosol to initiate the response cascade.

To identify proteins that are cytosolic and bind to DNA, we prepared cytosolic extracts from IFN-β treated RAW264.7 cells and human PBMCs. We assessed the efficiency of the subcellular fractionation using marker proteins that exclusively occur in the cytosol (the cytoskeletal protein tubulin) or the nucleus (the chromosome condensation regulator protein RCC1). Immunoblotting of the total extract showed a strong band for tubulin and a weak band for RCC1 (**Fig. 4.12**). In contrast, the cytosolic extracts showed a strong signal for tubulin and no detectable signal for RCC1. Vice versa, nuclear extracts were strongly enriched for RCC1 and did not contain detectable amounts of tubulin. Overall, this suggested that the cytosolic extracts used for the nucleic acid affinity purification were largely devoid of nuclear markers and were hence suitable for identifying cytosolic DNA binders.

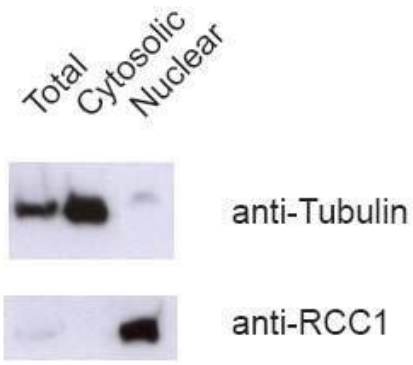



**Fig. 4.12: The cytosolic extracts are free of nuclear markers.** RAW 264.7 cells were harvested and the cytosolic and nuclear fractions were separated. Subsequently, the extracts were immunoblotted for cytosolic (Tubulin) and nuclear (RCC1) markers.

### 4.3.2 Nucleic acid affinity purification

In order to identify the cytosolic proteins which bind DNA, we carried out affinity purifications with nucleic acids as affinity ligands. We chose four different nucleic acids which were immobilized on streptavidin matrices. Firstly, we used interferon-stimulatory DNA (ISD) with a phosphodiester (PO) backbone, **ISD-PO**, an extensively characterized piece of DNA that can be synthesized *in vitro* (**Chapter 4.1.1**). Secondly, we used Calf Thymus-DNA (**CT-DNA**) a "natural" DNA of heterogeneous length and sequence. Thirdly, we used ISD with a phosphorothioate (PS) backbone (**ISD-PS**), rendering ISD resistant to nuclease cleavage, hence more stable. ISD-PS does not induce IFN-β, but has been shown to bind to nucleic acid binding proteins with a higher affinity **(Fig. 4.8**, **Fig. 4.13**). Furthermore, in order to be able to exclude RNA-binding proteins, we used **poly(I:C)**, a synthetic dsRNA composed of a strand of inosinic acid, poly (I), annealed to a strand of cytidylic acid, poly(C). By definition, a DNA sensor should have a higher affinity to DNA than to RNA. To exclude the proteins with a higher affinity for RNA from our affinity purifications, we added free poly(I:C) in excess to the purifications in which immobilized DNA was used as affinity ligand. Conversely, the poly(I:C) affinity purification was carried out in the presence of free calf-thymus DNA. Finally, we used **empty** streptavidin beads, in order to exclude proteins which bind to the affinity matrix rather than to nucleic acids.

Additionally, in order to capture proteins that bind to DNA during an *in vitro* DNA stimulation, we transfected RAW264.7 cells directly with biotinylated dsISD-PO and captured ISD-PO-binding proteins using streptavidin affinity matrix. In order to monitor the efficiency and specificity of the purification, we chose the endonuclease Trex1 and the alleged DNA sensor DAI/Zbp1 as positive controls (**Fig. 4.13**). Trex1 and DAI are cytosolic proteins previously shown to specifically bind to DNA [Stetson et al. 2008, Takaoka et al. 2007]. Indeed, immunoblotting experiments showed strong Trex1 bands in the eluates deriving from incubation with DNA coupled matrices, suggesting Trex1 was specifically retained by the immobilized DNA variants. In contrast, Trex1 does not bind to RNA, as no band can be seen in the eluate obtained from immobilized poly(I:C). As an expected consequence, the strong Trex1 band from the cytoplasmic extract disappears only in the supernatants incubated with immobilized DNA whereas it remains at comparable levels to the original extract in the purifications with RNA coupled resin. In the case of DAI/Zbp1, weak bands were seen only in the eluates from the immobilized ISD-PS and poly(I:C). This may indicate that DAI has a weaker affinity for DNA than Trex1 that is only detectable when ISD-PS is used as a ligand.

Collectively, the results suggest that our nucleic acid affinity purifications are capable of specifically capturing DNA binding proteins.

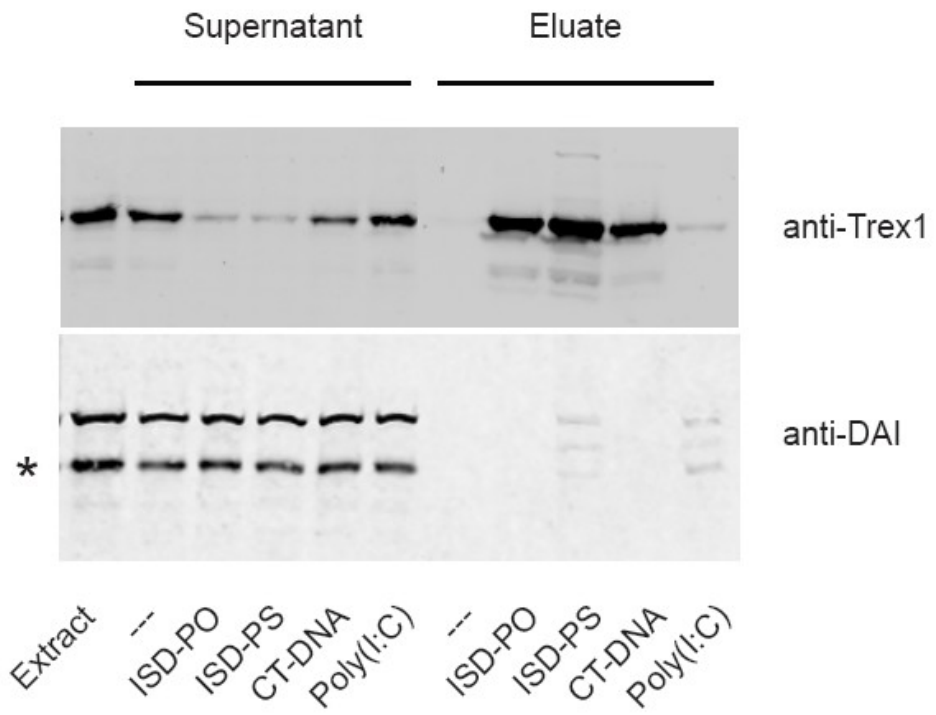



**Fig. 4.13 DNA affinity purification specifically captured DNA-binding proteins.** Cytosolic extracts of RAW264.7 cells were prepared as described. DNA affinity purifications using ISD-PO, ISD-PS and calf thymus DNA (CT-DNA) in the presence of free poly(I:C) as competitor and RNA affinity purification using poly(I:C) in the presence CT-DNA as competitor were carried out. The respective eluates were loaded on a SDS-page and probed for Trex1 and DAI by immunoblotting. (*) indicates the specific band for DAI.

In order to identify nucleic acid binders in an unbiased and comprehensive fashion, these same eluates were digested with trypsin and analyzed via liquid chromatography–tandem mass spectrometry. The mass spectrometry analysis resulted in a total identification of 1,606 proteins, of which 1,340 proteins were captured from RAW264.7 cells and 542 proteins from PBMCs. The overlap between these two datasets was around 280 proteins, i.e. more than 50% of the proteins present in the human dataset were also found in the murine dataset (**Fig. 4.14**).

## 4.4 Microarray and type I IFN-regulated genes

In order to capture transcripts that were regulated by IFN-β, we analyzed the proteins found in our proteomics dataset by two means: a publicly available microarray dataset and microarray analysis carried out in the laboratory.

First, we identified all the IFN-regulated proteins from the proteomics dataset by searching the list against the interferome-database (www.interferome.org) [Samarajiwa et al. 2009]. This database consists of over 2,000 genes that are regulated by different types of IFNs and displays only proteins that are up or down regulated by more than 1.5 fold relative to control samples when treated with type I IFN. When incorporating those data, we found that out of 1,345 proteins in RAW264.7-dataset and 542 proteins in the PBMC-dataset identified by mass spectrometry, 228 Proteins and 99 proteins, respectively, were type I IFN-regulated according to interferome database.

To further complement this analysis, we carried out microarray analysis. We treated three different cell lines, namely RAW264.7, L929 and NIH3T3 with IFN-β for 4h and analyzed changes in gene expression by microarray analysis. In total, over 500 transcripts were found to be upregulated more than 2-fold in the microarray dataset. Of 1340 proteins identified as DNA binders from RAW264.7 cells, 41 were found to be upregulated by IFN at the same time. Similarly, 25 out of 542 human DNA binders were found to be IFN-induced. Interestingly, all the transcripts identified by our microarray datasets were also covered by the interferome dataset (**Fig. 4.14**)

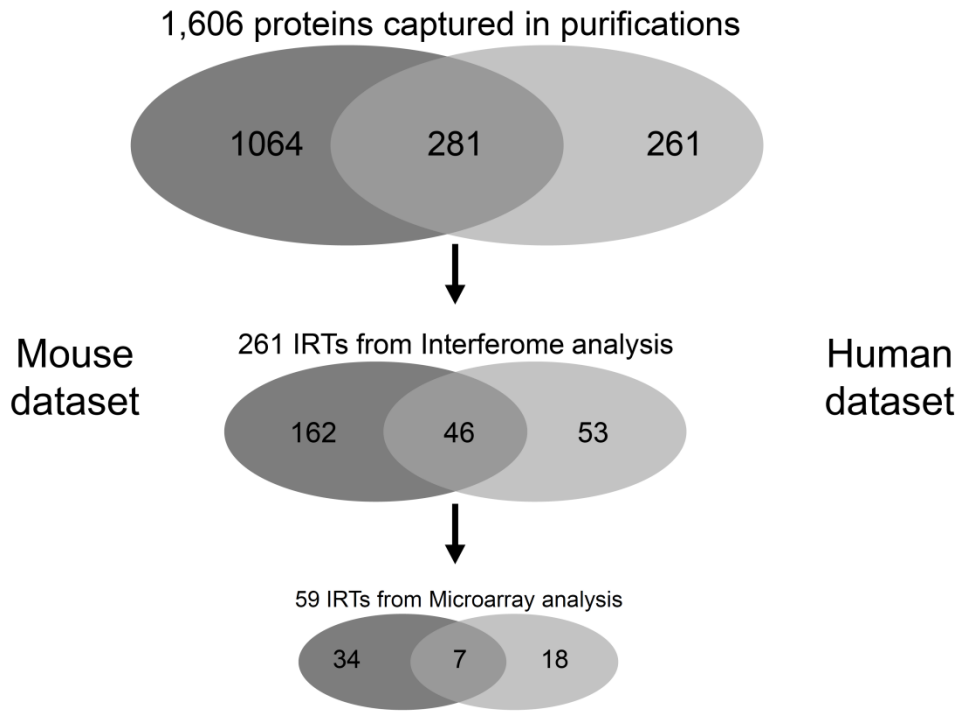



**Fig. 4.14: Summary of nucleic acid affinity purifications and the transcriptome analyses.** The nucleic acid affinity pulldowns captured a total of 1,606 proteins. Out of these, 261 proteins were found to be type I IFN regulated in the Interferome database. Of these, 59 were found to be upregulated by IFN-β in our own microarray experiments.

## 4.5 The scoring scheme

We wanted to take a systematic approach by being completely unbiased in our hunt for cytosolic DNA sensor but it is technically challenging, if not impossible, to functionally validate the entire set of proteins in a systematic manner. We decided to make use of bioinformatics and establish a scoring scheme to prioritize candidates and, eventually choose around 50 high-score candidates for a loss-of-function validation (**Fig. 4.15**).

In this scoring scheme, every protein captured in the nucleic acid affinity purifications initially starts with 100 points and it would gain points for favorable properties and lose points for unfavorable properties. An “ideal” DNA sensor should contain a nucleic acid binding domain, should have a higher affinity for DNA rather than RNA, and should be regulated by the cytokine it induces to enhance the antiviral response. Technically, it should specifically bind to nucleic acids and innate immunity related proteins. We defined innate immunity related proteins as proteins which are seen in pulldowns carried out at CeMM with baits involved in innate immunity (e.g. TBK1, AIM2, RIG-I, DAI, TLRs).

Proteins that were type I IFN-regulated according to the interferome database received 30 points whereas proteins that were IFN-regulated according to the microarray dataset or by both (interferome and microarray) obtained 50 points. The reason for this distinction is on the one hand that our microarray was also performed on RAW264.7 cells and hence it is better comparable. On the other hand, the microarray dataset identifies much less proteins as IFN-β regulated compared to interferome dataset. This may be due to the fact that interferome database considers any gene up or downregulated by 1.5 fold as IFN regulated gene compared to 2-fold upregulation in microarray dataset. Furthermore, for the generation of the microarray dataset the cells were only stimulated with IFN-β while the interferome dataset is generated using several interferons in several different cell types.

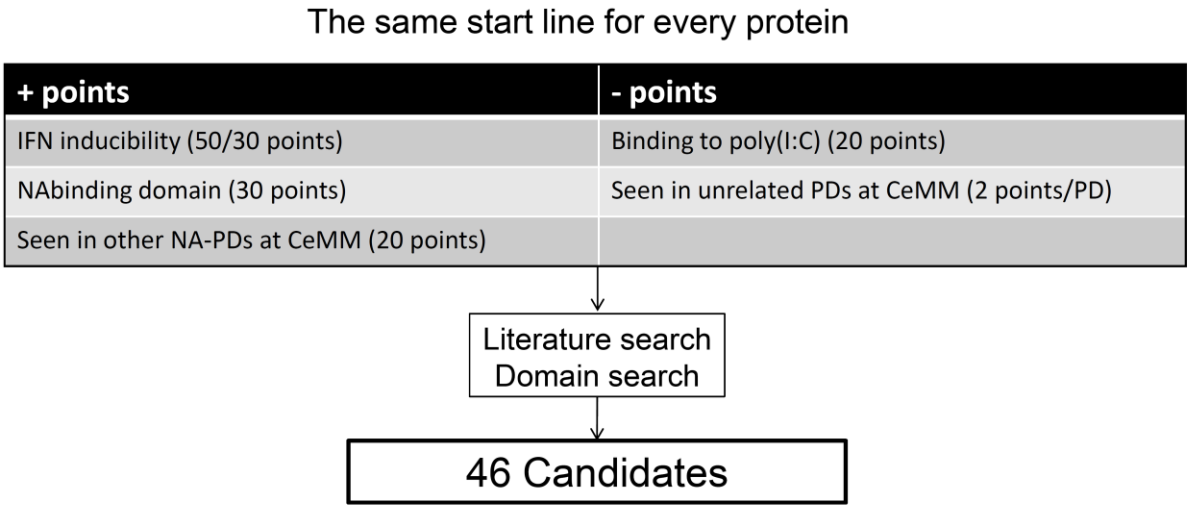



**Fig. 4.15: The scoring scheme.** Every protein identified by the mass spectrometry after the nucleic acid affinity purifications, started with hundred points and gained or lost points according to the above mentioned criteria. From the subsequently generated lists, the first 250 proteins were investigated for further information in the literature and explored for their domain composition. According to the obtained information, we decided on 46 candidate proteins. PD: Pulldown.

Furthermore, the potential candidate should contain a DNA binding domain. Upon *in silico* analysis using "Gene Ontology" the proteins containing a known DNA binding domain obtained 30 points [Ashburner et al. 2000]. In order to increase the specificity of the dataset, we took advantage of the fact that CeMM has created a database of proteins identified in more than 900 proteomics experiments, most of which were unrelated to innate immunity. We concluded that a protein seen in many unrelated purifications is likely to be a "sticky" protein, i.e. bind unspecifically, therefore its presence in the nucleic acid purifications may not be due to its DNA binding properties. Such proteins were penalized with two points per

affinity purification. Proteins that were found in other nucleic acid pulldowns which were carried out previously at CeMM, obtained 20 points. Furthermore, proteins that bound also to RNA in our experimental setup were penalized with 20 points.

In order to improve the efficiency of our scoring scheme, we used two proteins, which were previously shown to be involved in the cytosolic DNA recognition or in viral innate immunity, namely TREX1 and DAI. We adjusted the point weight given to the different parameters in a way that these proteins would rank high in the list, assuming that the DNA sensors would do the same.

From the obtained weighted list of proteins in the human (PBMCs) and mouse (RAW264.7) dataset, we investigated the domain composition and manually evaluated the existing literature about the first 250 proteins. At this stage, we searched for promising candidates with interesting domain composition (e.g. ATP binding domain, protein-protein interaction domain for downstream signaling) or previous connection to innate immunity from the literature. Furthermore, we focused on proteins which remain poorly characterized to date and discarded very well established proteins such as junB, for which a not-yet discovered role as DNA sensor appeared unlikely.

Interestingly, the dataset contained two proteins, SLFN2 and SLFN5, which belonged to the schlafen (SLFN) family. As several SLFN family members had a previous connection to immunity and some members contain a DNA/RNA helicase domain [Geserick et al. 2004, Berger et al. 2010], we decided to add additional family members, namely SLFN8 and SLFN9 to the list although they were not present in the pulldowns.

Eventually, upon investigation of the first 250 proteins in the mouse and human datasets, we decided on 46 candidates to carry out a loss-of-function validation (**Table 4.1**).

**Table 4.1: List of DNA sensor candidates**
The list of 46 DNA sensor candidate proteins was generated by applying the scoring scheme, *in silico* literature and domain analysis.

| Gene name | IFN-inducibilty | NA-binding domain containment | presence in CeMM NA-pulldowns | presence in poly(I:C) pulldown | presence in CeMM non-Innlm pulldowns | Total Score |
|---|---|---|---|---|---|---|
| Ighmbp2 | 30 | 30 | 20 | - | - | 200 |
| Isg20 | 50 | 30 | 20 | - | - | 200 |
| Sp100 | 50 | 30 | 20 | - | -4 | 200 |
| Ifi204 | 50 | 30 | 20 | - | -4 | 196 |
| ASCC3 | 50 | 30 | - | - | - | 180 |
| DAI/Zbp1 | 50 | 30 | - | - | - | 180 |
| SLFN5 | 50 | 30 | - | - | - | 180 |
| SLFN8 | 50 | 30 | - | - | - | 180 |
| SLFN9 | 50 | 30 | - | - | - | 180 |
| AI481105 | 50 | 30 | 20 | -20 | -4 | 176 |
| Bst2 | 50 | - | 20 | - | - | 170 |
| Epsti1 | 50 | - | 20 | - | - | 170 |
| Ifi202b | 50 | - | 20 | - | - | 170 |
| Mpeg1 | 50 | - | 20 | - | - | 170 |
| AI607873 | 50 | - | 20 | - | -8 | 162 |
| Rbm7 | 30 | 30 | 20 | -20 | - | 160 |
| Aebp1 | - | 30 | 20 | - | - | 150 |
| Ank3 | 30 | - | 20 | - | - | 150 |
| Ankhd1 | - | 30 | 20 | - | - | 150 |
| Btaf1 | - | 30 | 20 | - | - | 150 |
| Centg3 | - | 30 | 20 | - | - | 150 |
| Cpsf7 | - | 30 | 20 | - | - | 150 |
| Dna2l | - | 30 | 20 | - | - | 150 |
| Ecsit | - | 30 | 20 | - | - | 150 |
| Fiz1 | - | 30 | 20 | - | - | 150 |
| Irgm | 50 | - | - | - | - | 150 |
| NO66 | - | 30 | 20 | - | - | 150 |
| Nufip1 | - | 30 | 20 | - | - | 150 |
| Obfc1 | - | 30 | 20 | - | - | 150 |
| Ppfibp1 | - | 30 | 20 | - | - | 150 |
| Preb | - | 30 | 20 | - | - | 150 |
| Prrx1 | - | 30 | 20 | - | - | 150 |
| Pyhin1 | 50 | - | - | - | - | 150 |
| Rbms1 | - | 30 | 20 | - | - | 150 |
| Rbms2 | - | 30 | 20 | - | - | 150 |
| RSAD2 | 50 | - | - | - | - | 150 |
| Sbf1 | 30 | - | 20 | - | - | 150 |
| Sbno1 | - | 30 | 20 | - | - | 150 |
| SLFN2 | 50 | - | - | - | - | 150 |
| SNRNP200 | - | 30 | 20 | - | - | 150 |
| Srbd1 | - | 30 | 20 | - | - | 150 |
| Themis2 | 50 | - | - | - | - | 150 |
| Trim56 | 50 | - | - | - | - | 150 |
| Zbtb7a | - | 30 | 20 | - | - | 150 |
| Zcchc11 | - | 30 | 20 | - | - | 150 |
| Zfp143 | - | 30 | 20 | - | - | 150 |

## 4.6 Loss-of-function validation

In order to test the impact of the DNA sensor candidates on IFN-β production, we opted for a loss-of-function validation. Furthermore, we decided to use RAW264.7 cells for this approach. As seen in chapter 4.1, RAW264.7 cells are a perfectly suitable innate immune cell line, which produces high amount of IFN-β upon DNA transfection. We were able to show that these cells recognize the DNA in the cytosol and the receptor is still elusive.

### 4.6.1 The experimental setup and its optimization

To determine the effects of the candidate proteins on the IFN-β levels, we decided to carry out a loss-of-function validation in RAW264.7 cells. For this purpose, we designed and cloned six shRNA vectors per DNA sensor candidate using an algorithm developed at the NIH in collaboration with the laboratory of Louis Staudt *from National Institutes of Health, National Cancer Institute, Bethesda, USA* [Reynolds et al. 2004]. Additional to those 272 shRNAs, we also designed one non-targeting shRNA as a negative control and six shRNAs against TBK1 as positive control. Each vector of shRNA was carrying a GFP gene and a puromycin-resistance-gene for selection of cell lines which stably express the shRNA of choice.

In total, 279 shRNAs were cloned. We optimized the screen by using the shRNAs against TBK1 as a test case. Expression of the different TBK1-specific shRNAs in RAW264.7 cells led to downregulation of TBK1 in the case of shRNA3 and shRNA5 and to an almost completely loss in the case of shRNA2 and ahRNA6 as seen by the disappearance of the TBK1-specific band in immunoblotting assay (**Fig. 4.16**). These results suggest that we can achieve decent knockdown levels in RAW264.7 cells and indicate that this approach is feasible for our screen.

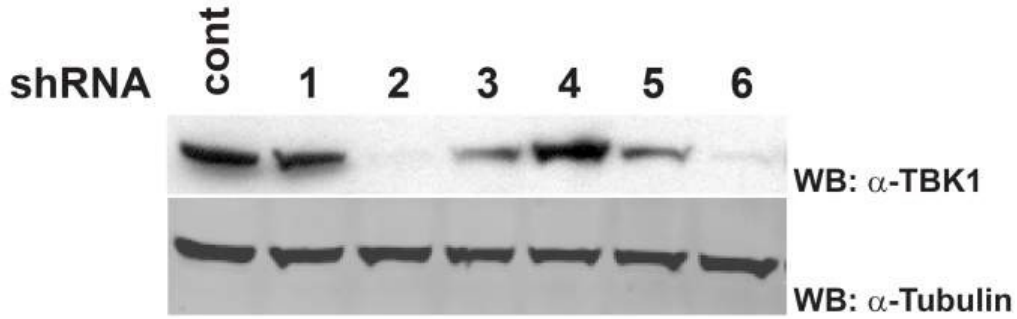



**Fig. 4.16: shRNAs 2 and 6 lead to efficient knockdown of TBK1 in RAW264.7 cells.** 6 different shRNA-vectors against TBK1 were transfected into the producer cell line 293gp cells. The virus - containing supernatants were transferred on RAW264.7 cells and were either incubated for 24h before the first medium change. After keeping the cell under puromycin selection for 72h, the cells were lysed and immunoblotted against TBK1 to test the incubation efficiencies. Tubulin serves as loading control.

For the loss-of-function validation, we created cell lines expressing each of the 279 shRNAs in parallel. To this end, we transfected 293gp cells with respective shRNA and helper constructs. After 48h, the supernatants containing the viruses were transferred on RAW264.7 and the cells were selected until they reach at least 70% GFP positivity. At this point, the created stable cell lines were stimulated with DNA for 4h and subsequently, the supernatants were transferred on LL171 cells, a mouse reporter cell line which originates from L929 cells containing a stable IFN-stimulated response element-luciferase reporter plasmid [ISRE-Luc], hence expressing luciferase linear to IFN-β levels. The measured luciferase levels were normalized to respective RAW264.7 cell number from which the supernatants were originating (**Fig. 4.17**).

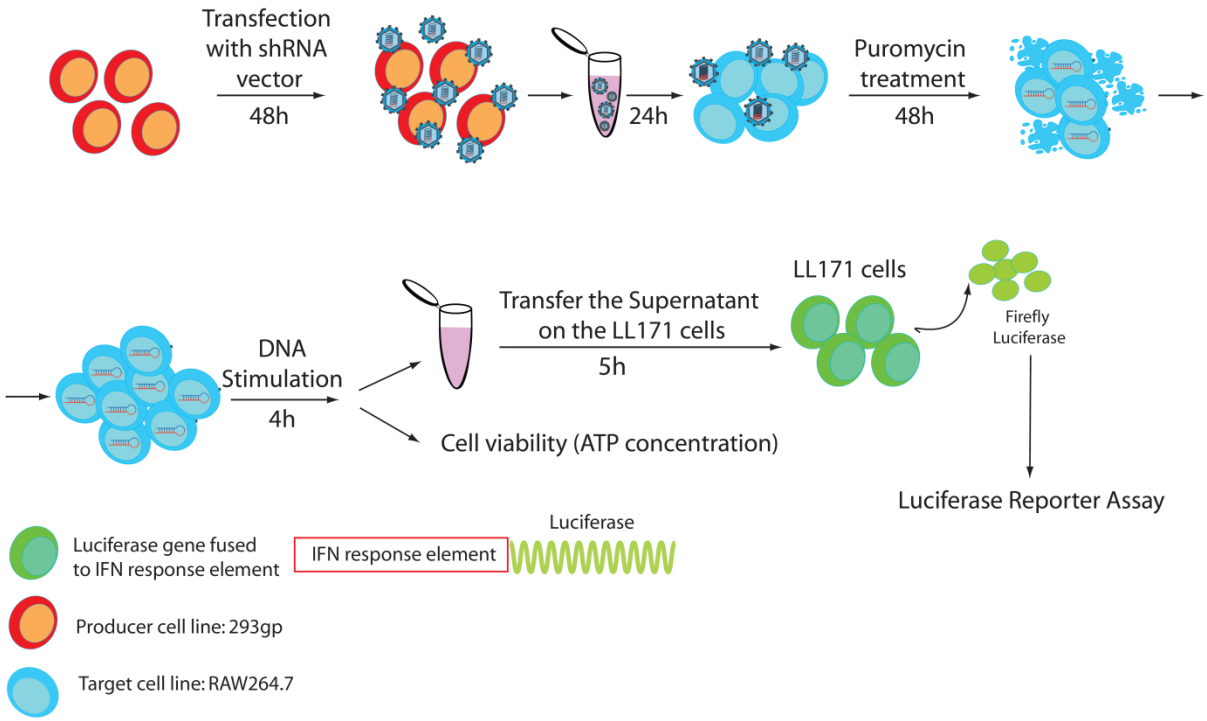


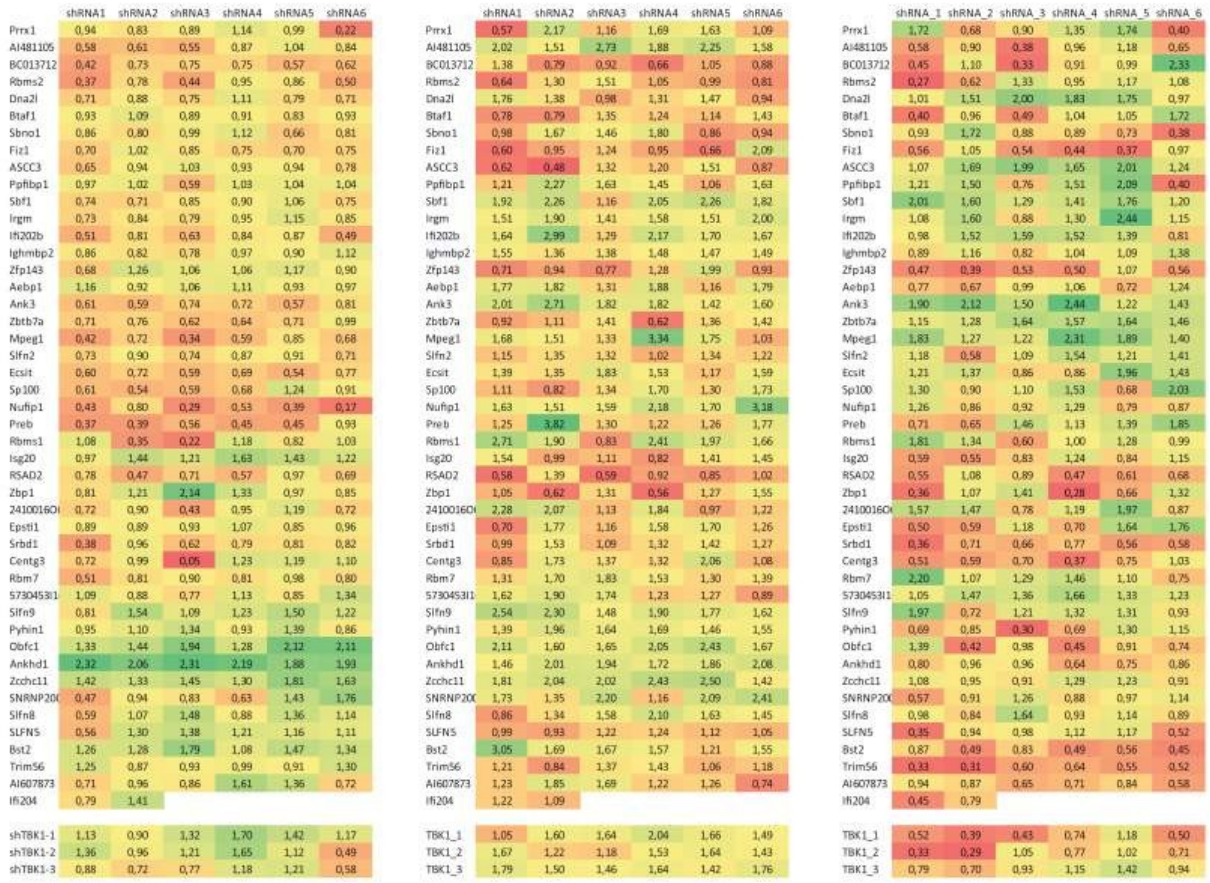

**Fig. 4.17: The schematic representation of the loss-of-function validation.** The producer cell line 293gp was transfected with respective shRNA. After 48 hours the supernatants containing the virus were transferred on RAW264.7 cells. After 24h of incubation, the cells were selected with puromycin until over 70% of the cells were GFP-positive. Thereafter, the RAW264.7 cells were stimulated with ISD for 4h. The supernatants were collected and transferred on reporter cell line LL171 which contains a luciferase gene fused to an IFN-response element and incubated for 5h. The luciferase activity, which is linear to the IFN-β levels, was measured using luciferase reporter assay. The luciferase activity is normalized to cell number using cell viability assay.

## 4.6.2 The validation

In order to monitor the stability of the loss-of-function validation we established several control points in the assay. First of all, we created three independent sets of shTBK1 cell lines. Secondly, we wanted to obtain at least 10-fold change in the IFN-β production after DNA stimulation compared to unstimulated cells. Finally, we included shRNAs 2 and 6 against TBK1 as positive controls. Cells were monitored for morphological changes during the screen and were discarded if gross morphological changes were observed. Finally, cells were examined by FACS for GFP expression: Only those screens, in which more than 70% of the cells express GFP, were stimulated with ISD (data not shown). In order to obtain

sufficient data for a statistical analysis, we carried out three independent screens (**Table 4.2**). shows the results of each shRNA in three independent screens as a heatmap.

**Table 4.2: The heatmaps of three independent screens.**
The supernatants from DNA stimulated RAW264.7 cells were transferred in triplicates on LL171 cells. After 5h, the luciferase signal was measured. Each received signal was first normalized to the cell number and then to a well which contains the same supernatant and is present on each plate. The average of triplicates was calculated and normalized to the signal from the negative control in order to calculate the fold of change compared to negative control. The results are displayed on a heatmap. The color-code reads from red to green for low to high signals, respectively.



| | shRNA1 | shRNA2 | shRNA3 | shRNA4 | shRNA5 | shRNA6 |
|---|---|---|---|---|---|---|
| Prrx1 | 0,94 | 0,83 | 0,89 | 1,14 | 0,99 | 0,22 |
| AI481105 | 0,58 | 0,61 | 0,55 | 0,87 | 1,04 | 0,84 |
| BC013712 | 0,42 | 0,73 | 0,75 | 0,75 | 0,57 | 0,62 |
| Rbms2 | 0,37 | 0,78 | 0,44 | 0,95 | 0,86 | 0,50 |
| Dna2l | 0,71 | 0,88 | 0,75 | 1,11 | 0,79 | 0,71 |
| Btaf1 | 0,93 | 1,09 | 0,89 | 0,91 | 0,83 | 0,93 |
| Sbno1 | 0,86 | 0,80 | 0,99 | 1,12 | 0,66 | 0,81 |
| Fiz1 | 0,70 | 1,02 | 0,85 | 0,75 | 0,70 | 0,75 |
| ASCC3 | 0,65 | 0,94 | 1,03 | 0,93 | 0,94 | 0,78 |
| Ppfibp1 | 0,97 | 1,02 | 0,59 | 1,03 | 1,04 | 1,04 |
| Sbf1 | 0,74 | 0,71 | 0,85 | 0,90 | 1,06 | 0,75 |
| Irgm | 0,73 | 0,84 | 0,79 | 0,95 | 1,15 | 0,85 |
| Ifi202b | 0,51 | 0,81 | 0,63 | 0,84 | 0,87 | 0,49 |
| Ighmbp2 | 0,86 | 0,82 | 0,78 | 0,97 | 0,90 | 1,12 |
| Zfp143 | 0,68 | 1,26 | 1,06 | 1,06 | 1,17 | 0,90 |
| Aebp1 | 1,16 | 0,92 | 1,06 | 1,11 | 0,93 | 0,97 |
| Ank3 | 0,61 | 0,59 | 0,74 | 0,72 | 0,57 | 0,81 |
| Zbtb7a | 0,71 | 0,76 | 0,62 | 0,64 | 0,71 | 0,99 |
| Mpeg1 | 0,42 | 0,72 | 0,34 | 0,59 | 0,85 | 0,68 |
| Slfn2 | 0,73 | 0,90 | 0,74 | 0,87 | 0,91 | 0,71 |
| Ecsit | 0,60 | 0,72 | 0,59 | 0,69 | 0,54 | 0,77 |
| Sp100 | 0,61 | 0,54 | 0,59 | 0,68 | 1,24 | 0,91 |
| Nufip1 | 0,43 | 0,80 | 0,29 | 0,53 | 0,39 | 0,17 |
| Preb | 0,37 | 0,39 | 0,56 | 0,45 | 0,45 | 0,93 |
| Rbms1 | 1,08 | 0,35 | 0,22 | 1,18 | 0,82 | 1,03 |
| Isg20 | 0,97 | 1,44 | 1,21 | 1,63 | 1,43 | 1,22 |
| RSAD2 | 0,78 | 0,47 | 0,71 | 0,57 | 0,97 | 0,69 |
| Zbp1 | 0,81 | 1,21 | 2,14 | 1,33 | 0,97 | 0,85 |
| 2410016O | 0,72 | 0,90 | 0,43 | 0,95 | 1,19 | 0,72 |
| Epsti1 | 0,89 | 0,89 | 0,93 | 1,07 | 0,85 | 0,96 |
| Srbd1 | 0,38 | 0,96 | 0,62 | 0,79 | 0,81 | 0,82 |
| Centg3 | 0,72 | 0,99 | 0,05 | 1,23 | 1,19 | 1,10 |
| Rbm7 | 0,51 | 0,81 | 0,90 | 0,81 | 0,98 | 0,80 |
| 5730453I1 | 1,09 | 0,88 | 0,77 | 1,13 | 0,85 | 1,34 |
| Slfn9 | 0,81 | 1,54 | 1,09 | 1,23 | 1,50 | 1,22 |
| Pyhin1 | 0,95 | 1,10 | 1,34 | 0,93 | 1,39 | 0,86 |
| Obfc1 | 1,33 | 1,44 | 1,94 | 1,28 | 2,12 | 2,11 |
| Ankhd1 | 2,32 | 2,06 | 2,31 | 2,19 | 1,88 | 1,93 |
| Zcchc11 | 1,42 | 1,33 | 1,45 | 1,30 | 1,81 | 1,63 |
| SNRNP20 | 0,47 | 0,94 | 0,83 | 0,63 | 1,43 | 1,76 |
| Slfn8 | 0,59 | 1,07 | 1,48 | 0,88 | 1,36 | 1,14 |
| SLFN5 | 0,56 | 1,30 | 1,38 | 1,21 | 1,16 | 1,11 |
| Bst2 | 1,26 | 1,28 | 1,79 | 1,08 | 1,47 | 1,34 |
| Trim56 | 1,25 | 0,87 | 0,93 | 0,99 | 0,91 | 1,30 |
| AI607873 | 0,71 | 0,96 | 0,86 | 1,61 | 1,36 | 0,72 |
| Ifi204 | 0,79 | 1,41 | | | | |
| | | | | | | |
| shTBK1-1 | 1,13 | 0,90 | 1,32 | 1,70 | 1,42 | 1,17 |
| shTBK1-2 | 1,36 | 0,96 | 1,21 | 1,65 | 1,12 | 0,49 |
| shTBK1-3 | 0,88 | 0,72 | 0,77 | 1,18 | 1,21 | 0,58 |

| | shRNA1 | shRNA2 | shRNA3 | shRNA4 | shRNA5 | shRNA6 |
|---|---|---|---|---|---|---|
| Prrx1 | 0,57 | 2,17 | 1,16 | 1,69 | 1,63 | 1,09 |
| AI481105 | 2,02 | 1,51 | 2,73 | 1,88 | 2,25 | 1,58 |
| BC013712 | 1,38 | 0,79 | 0,92 | 0,66 | 1,05 | 0,88 |
| Rbms2 | 0,64 | 1,30 | 1,51 | 1,05 | 0,99 | 0,81 |
| Dna2l | 1,76 | 1,38 | 0,98 | 1,31 | 1,47 | 0,94 |
| Btaf1 | 0,78 | 0,79 | 1,35 | 1,24 | 1,14 | 1,43 |
| Sbno1 | 0,98 | 1,67 | 1,46 | 1,80 | 0,86 | 0,94 |
| Fiz1 | 0,60 | 0,95 | 1,24 | 0,95 | 0,66 | 2,09 |
| ASCC3 | 0,62 | 0,48 | 1,32 | 1,20 | 1,51 | 0,87 |
| Ppfibp1 | 1,21 | 2,27 | 1,63 | 1,45 | 1,06 | 1,63 |
| Sbf1 | 1,92 | 2,26 | 1,16 | 2,05 | 2,26 | 1,82 |
| Irgm | 1,51 | 1,90 | 1,41 | 1,58 | 1,51 | 2,00 |
| Ifi202b | 1,64 | 2,99 | 1,29 | 2,17 | 1,70 | 1,67 |
| Ighmbp2 | 1,55 | 1,36 | 1,38 | 1,48 | 1,47 | 1,49 |
| Zfp143 | 0,71 | 0,94 | 0,77 | 1,28 | 1,99 | 0,93 |
| Aebp1 | 1,77 | 1,82 | 1,31 | 1,88 | 1,16 | 1,79 |
| Ank3 | 2,01 | 2,71 | 1,82 | 1,82 | 1,42 | 1,60 |
| Zbtb7a | 0,92 | 1,11 | 1,41 | 0,62 | 1,36 | 1,42 |
| Mpeg1 | 1,68 | 1,51 | 1,33 | 3,34 | 1,75 | 1,03 |
| Slfn2 | 1,15 | 1,35 | 1,32 | 1,02 | 1,34 | 1,22 |
| Ecsit | 1,39 | 1,35 | 1,83 | 1,53 | 1,17 | 1,59 |
| Sp100 | 1,11 | 0,82 | 1,34 | 1,70 | 1,30 | 1,73 |
| Nufip1 | 1,63 | 1,51 | 1,59 | 2,18 | 1,70 | 3,18 |
| Preb | 1,25 | 3,82 | 1,30 | 1,22 | 1,26 | 1,77 |
| Rbms1 | 2,71 | 1,90 | 0,83 | 2,41 | 1,97 | 1,66 |
| Isg20 | 1,54 | 0,99 | 1,11 | 0,82 | 1,41 | 1,45 |
| RSAD2 | 0,58 | 1,39 | 0,59 | 0,92 | 0,85 | 1,02 |
| Zbp1 | 1,05 | 0,62 | 1,31 | 0,56 | 1,27 | 1,55 |
| 2410016O | 2,28 | 2,07 | 1,13 | 1,84 | 0,97 | 1,22 |
| Epsti1 | 0,70 | 1,77 | 1,16 | 1,58 | 1,70 | 1,26 |
| Srbd1 | 0,99 | 1,53 | 1,09 | 1,32 | 1,42 | 1,27 |
| Centg3 | 0,85 | 1,73 | 1,37 | 1,32 | 2,06 | 1,08 |
| Rbm7 | 1,31 | 1,70 | 1,83 | 1,53 | 1,30 | 1,39 |
| 5730453I1 | 1,62 | 1,90 | 1,74 | 1,23 | 1,27 | 0,89 |
| Slfn9 | 2,54 | 2,30 | 1,48 | 1,90 | 1,77 | 1,62 |
| Pyhin1 | 1,39 | 1,96 | 1,64 | 1,69 | 1,46 | 1,55 |
| Obfc1 | 2,11 | 1,60 | 1,65 | 2,05 | 2,43 | 1,67 |
| Ankhd1 | 1,46 | 2,01 | 1,94 | 1,72 | 1,86 | 2,08 |
| Zcchc11 | 1,81 | 2,04 | 2,02 | 2,43 | 2,50 | 1,42 |
| SNRNP20 | 1,73 | 1,35 | 2,20 | 1,16 | 2,09 | 2,41 |
| Slfn8 | 0,86 | 1,34 | 1,58 | 2,10 | 1,63 | 1,45 |
| SLFN5 | 0,99 | 0,93 | 1,22 | 1,24 | 1,12 | 1,05 |
| Bst2 | 3,05 | 1,69 | 1,67 | 1,57 | 1,21 | 1,55 |
| Trim56 | 1,21 | 0,84 | 1,37 | 1,43 | 1,06 | 1,18 |
| AI607873 | 1,23 | 1,85 | 1,69 | 1,22 | 1,26 | 0,74 |
| Ifi204 | 1,22 | 1,09 | | | | |
| | | | | | | |
| TBK1_1 | 1,05 | 1,60 | 1,64 | 2,04 | 1,66 | 1,49 |
| TBK1_2 | 1,67 | 1,22 | 1,18 | 1,53 | 1,64 | 1,43 |
| TBK1_3 | 1,79 | 1,50 | 1,46 | 1,64 | 1,42 | 1,76 |

| | shRNA_1 | shRNA_2 | shRNA_3 | shRNA_4 | shRNA_5 | shRNA_6 |
|---|---|---|---|---|---|---|
| Prrx1 | 1,72 | 0,68 | 0,90 | 1,35 | 1,74 | 0,40 |
| AI481105 | 0,58 | 0,90 | 0,38 | 0,96 | 1,18 | 0,65 |
| BC013712 | 0,45 | 1,10 | 0,33 | 0,91 | 0,99 | 2,33 |
| Rbms2 | 0,27 | 0,62 | 1,33 | 0,95 | 1,17 | 1,08 |
| Dna2l | 1,01 | 1,51 | 2,00 | 1,83 | 1,75 | 0,97 |
| Btaf1 | 0,40 | 0,96 | 0,49 | 1,04 | 1,05 | 1,72 |
| Sbno1 | 0,93 | 1,72 | 0,88 | 0,89 | 0,73 | 0,38 |
| Fiz1 | 0,56 | 1,05 | 0,54 | 0,44 | 0,37 | 0,97 |
| ASCC3 | 1,07 | 1,69 | 1,99 | 1,65 | 2,01 | 1,24 |
| Ppfibp1 | 1,21 | 1,50 | 0,76 | 1,51 | 2,09 | 0,40 |
| Sbf1 | 2,01 | 1,60 | 1,29 | 1,41 | 1,76 | 1,20 |
| Irgm | 1,08 | 1,60 | 0,88 | 1,30 | 2,44 | 1,15 |
| Ifi202b | 0,98 | 1,52 | 1,59 | 1,52 | 1,39 | 0,81 |
| Ighmbp2 | 0,89 | 1,16 | 0,82 | 1,04 | 1,09 | 1,38 |
| Zfp143 | 0,47 | 0,39 | 0,53 | 0,50 | 1,07 | 0,56 |
| Aebp1 | 0,77 | 0,67 | 0,99 | 1,06 | 0,72 | 1,24 |
| Ank3 | 1,90 | 2,12 | 1,50 | 2,44 | 1,22 | 1,43 |
| Zbtb7a | 1,15 | 1,28 | 1,64 | 1,57 | 1,64 | 1,46 |
| Mpeg1 | 1,83 | 1,27 | 1,22 | 2,31 | 1,89 | 1,40 |
| Slfn2 | 1,18 | 0,58 | 1,09 | 1,54 | 1,21 | 1,41 |
| Ecsit | 1,21 | 1,37 | 0,86 | 0,86 | 1,96 | 1,43 |
| Sp100 | 1,30 | 0,90 | 1,10 | 1,53 | 0,68 | 2,03 |
| Nufip1 | 1,26 | 0,86 | 0,92 | 1,29 | 0,79 | 0,87 |
| Preb | 0,71 | 0,65 | 1,46 | 1,13 | 1,39 | 1,85 |
| Rbms1 | 1,81 | 1,34 | 0,60 | 1,00 | 1,28 | 0,99 |
| Isg20 | 0,59 | 0,55 | 0,83 | 1,24 | 0,84 | 1,15 |
| RSAD2 | 0,55 | 1,08 | 0,89 | 0,47 | 0,61 | 0,68 |
| Zbp1 | 0,36 | 1,07 | 1,41 | 0,28 | 0,66 | 1,32 |
| 2410016O | 1,57 | 1,47 | 0,78 | 1,19 | 1,97 | 0,87 |
| Epsti1 | 0,50 | 0,59 | 1,18 | 0,70 | 1,64 | 1,76 |
| Srbd1 | 0,36 | 0,71 | 0,66 | 0,77 | 0,56 | 0,58 |
| Centg3 | 0,51 | 0,59 | 0,70 | 0,37 | 0,75 | 1,03 |
| Rbm7 | 2,20 | 1,07 | 1,29 | 1,46 | 1,10 | 0,75 |
| 5730453I1 | 1,05 | 1,47 | 1,36 | 1,66 | 1,33 | 1,23 |
| Slfn9 | 1,97 | 0,72 | 1,21 | 1,32 | 1,31 | 0,93 |
| Pyhin1 | 0,69 | 0,85 | 0,30 | 0,69 | 1,30 | 1,15 |
| Obfc1 | 1,39 | 0,42 | 0,98 | 0,45 | 0,91 | 0,74 |
| Ankhd1 | 0,80 | 0,96 | 0,96 | 0,64 | 0,75 | 0,86 |
| Zcchc11 | 1,08 | 0,95 | 0,91 | 1,29 | 1,23 | 0,91 |
| SNRNP20 | 0,57 | 0,91 | 1,26 | 0,88 | 0,97 | 1,14 |
| Slfn8 | 0,98 | 0,84 | 1,64 | 0,93 | 1,14 | 0,89 |
| SLFN5 | 0,35 | 0,94 | 0,98 | 1,12 | 1,17 | 0,52 |
| Bst2 | 0,87 | 0,49 | 0,83 | 0,49 | 0,56 | 0,45 |
| Trim56 | 0,33 | 0,31 | 0,60 | 0,64 | 0,55 | 0,52 |
| AI607873 | 0,94 | 0,87 | 0,65 | 0,71 | 0,84 | 0,58 |
| Ifi204 | 0,45 | 0,79 | | | | |
| | | | | | | |
| TBK1_1 | 0,52 | 0,39 | 0,43 | 0,74 | 1,18 | 0,50 |
| TBK1_2 | 0,33 | 0,29 | 1,05 | 0,77 | 1,02 | 0,71 |
| TBK1_3 | 0,79 | 0,70 | 0,93 | 1,15 | 1,42 | 0,94 |

Based on the assumption that the majority of the shRNAs should not show an effect, we used the mean of all observations as the reference point. After normalization, we analyzed each shRNA against the reference point using a one-sided, non-parametric Kolmogorov-Smirnov test. shRNAs were considered as having an impact on the read-out if they produced a signal in 1$^{st}$ quartile of all observations with a p-value lower than 0.05 in at least two out of three screens. Out of 272 shRNAs tested, 45 shRNAs had an impact on type-I interferon induction by cytosolic DNA (**Table 4.3**). These 45 shRNAs were targeting 20 genes.

**Table 4.3: The analysis of three independent screens**

| | |
|---|---|
| Genes with ≥3 successful shRNAs in ≥2 screens | 2 |
| Genes with ≥2 successful shRNAs in ≥2 screens | 10 |
| shRNAs that were successful in all 3 screens | 6 |
| shRNAs that were successful in ≥2 screens | 45 |

Out of the genes which were targeted by this 45 shRNAs, we selected those for which two or more shRNAs showed a phenotype. This led us to continue with 10 genes that had at least two shRNAs with a signal in the 1$^{st}$ quartile of all observations in at least two screens (**Table 4.4**). Exceptionally, we decided to take one additional gene which had one successful shRNA in all three screens, namely Rbms1, as it seemed to have a robust effect. Notably, Zbp1/DAI was within the candidates selected by the bioinformatics analysis. This was reassuring as Zbp1/DAI was suggested to be a DNA sensor [Takaoka *et al.* 2007].

**Table 4.4: The DNA sensor candidates after loss-of-function validation.**
We decided on 10 candidates which showed with at least 2 shRNAs a significant decrease in the IFN-β response after DNA stimulation in at least 2 screens and one additional protein (Rbms1) which showed with only one shRNA a significant reduction in all three screens.

| Name | Ref_Seq ID | Uniprot ID | Number of successful shRNA's |
|---|---|---|---|
| Themis2 | NP_001028480.1 | Q91YX0 | 5 |
| Zfp143 | NP_033307.2 | O70230 | 3 |
| Rbms2 | NP_062685.2 | Q8VC70 | 2 |
| Sbno1 | NP_001074672.1 | B2RRI2 | 2 |
| Fiz1 | NP_001103798.1 | Q9WTJ4 | 2 |
| Ecsit | NP_036159.1 | Q9QZH6 | 2 |
| Preb | NP_057912.2 | Q9WUQ2 | 2 |
| RSAD2 | NP_067359.2 | Q8CBB9 | 2 |
| Zbp1 | NP_067369.2 | A2APF7 | 2 |
| Srbd1 | NP_084409.3 | Q497V5 | 2 |
| Rbms1 | NP_001135404.1 | Q91W59 | 1 |

We assessed the knock-down efficiency for the shRNAs targetting those 10 genes (**Table 4.4**) by real-time PCR using two independent primer pairs (data not shown). **Fig. 4.18** shows four representative examples of target genes for which the respective mRNA levels in the created knockdown cell lines were measured by RT-PCR. The results show that the cell lines expressing shRNAs against Themis2, Rbms2, Zfp143, Sbno1 reduced the mRNA expression levels of the target genes by more than 75%.

Out of 28 shRNA knockdown cell lines tested, 7 cell lines failed to show an efficient knockdown for the target gene. This may have resulted from either inappropriate integration of the vector in the genome, or from the fact that these shRNAs are unable to knockdown the respective transcripts in the first place but were selected as a result of off-target effects during the initial screen.

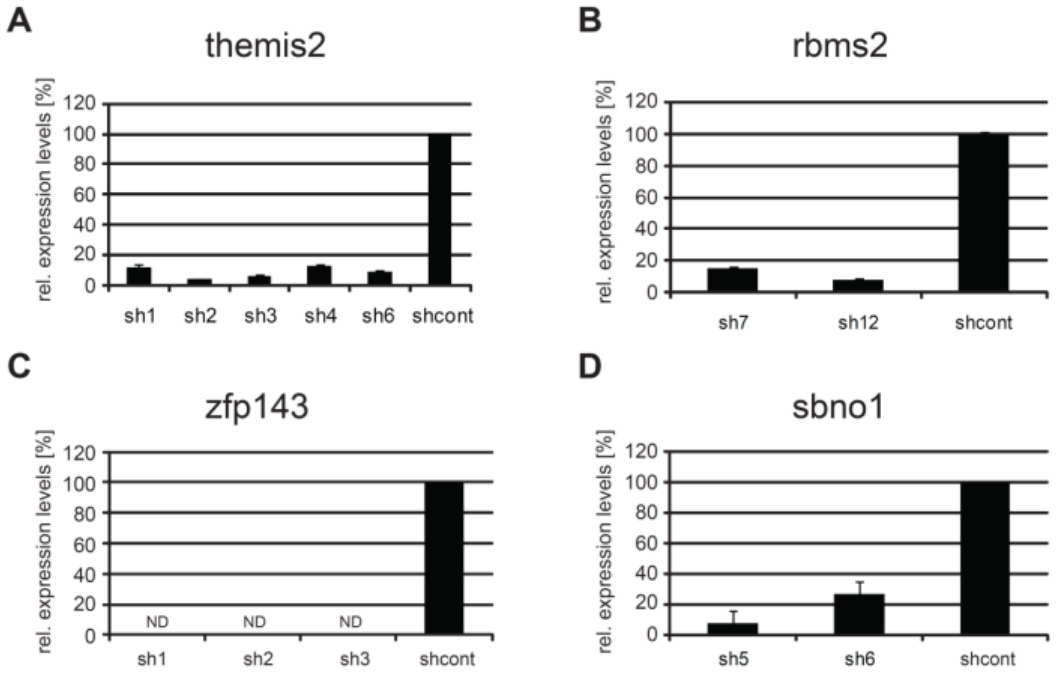



**Fig. 4.18**: Four representative examples for target gene knockdown cell lines. The transcription levels of respective genes (themis2 (A), rbms2 (B), zfp143 (C), sbno1 (D)) were analyzed by quantitative RT-PCR analysis. The levels are normalized against cyclophilin B and displayed in comparison to control knockdown cell line. Data are shown as mean ± SD of three replicates in one representative experiment out of two independent experiments using two independent primer pairs.

The 10 candidates obtained from the loss-of-function validation will undergo a secondary validation using cell lines that stably express those shRNAs. In the secondary validation, additional PAMPs will be evaluated in parallel to see whether loss of any given candidate leads to a selective impairment of DNA recognition in those cells.

# 5 Discussion

Cytoplasmic DNA triggers the production and secretion of proinflammatory cytokines and interferons. As our understanding of cytosolic DNA recognition is still limited, we aimed to characterize the cytosolic DNA response in more detail. To this end, we investigated the structural requirements for DNA recognition as well as the consequences of cytosolic DNA recognition at the level of cellular signal transduction using two synthetic DNAs (ISD and polydAdT).

Regarding the structural requirements, we tested if modifications at the 5' end lead to any changes in the DNA recognition. We showed that modification of DNA at its 5' end does not affect its innate immune potential suggesting that recognition occurs at the DNA backbone. Clearly, TLR9 does not contribute to this process, because endosomal acidification and cathepsin activity are not required for cytosolic DNA recognition. In order to estimate the minimal required length of DNA necessary to be sensed by the elusive sensor(s), we created ISDs of different lengths and found that DNA pieces shorter than 24bps fail to induce IFN-β response. Furthermore, our data show that both single- and double-stranded phosphodiester-containing ISD oligonucleotides (ISD-PO) can activate the interferon pathway. Interestingly, we were able to block the IFN signaling induced by ISD-PO with increasing amounts of ssISD-PS, similarly to CpG-recognition by TLR9.

In an attempt to characterize the pathways of cytosolic DNA recognition in RAW264.7 cells, we stimulated them with ISD and poly(dAdT) and analyzed the gene expression pattern by microarray. Different DNAs triggered a specific gene expression profile in RAW264.7 cells that was distinct from the one obtained after TLR9 dependent stimulation. This indicates that the as-yet unknown DNA sensor(s) use(s) a dedicated signaling pathway distinct from the one downstream of TLR9.

Next, we aimed at identifying the elusive IFN-inducing receptor. To this end, we purified nucleic-acid binding proteins using immobilized DNA as an affinity reagent and identified bound proteins by mass spectrometry analysis. The purifications yielded around 1600 proteins of which 1340 proteins were captured from RAW264.7 cells and 542 proteins from PBMCs. We prioritized those proteins by the means of a scoring scheme and selected a total number of 46 candidates for a loss-of function screen. To this end, we created RAW264.7 cell lines stably expressing shRNAs against the candidate proteins, challenged those cells

with DNA and measured IFN-β production. Statistical analysis revealed 45 shRNAs which led to diminished IFN-β production in those cells. These 45 shRNAs mapped to 20 genes out of which 10 genes had 2 or more shRNAs producing a signal with the aforementioned criteria. We confirmed the knockdown efficiency of 28 individual shRNAs. Eventually, secondary validation experiments will be needed to confirm the loss-of-function phenotype, investigate the gain-of-function phenotype and provide a molecular mechanism of action for selected candidates.

## 5.1 Structural requirements for cytosolic DNA recognition

In an attempt to define structural requirements of DNA, we showed that, unlike for RNA recognition, modifications at the 5'end did not affect innate immune activation. It was not excluded that modifications of 3' end may play a role in recognition.

Overall, our findings are in accordance with the notion that the distinction between self and non-self-DNA recognition derives from subcellular localization of the DNA rather than the differences in the structure [Marshak-Rothstein and Rifkin 2007]: While host DNA is usually confined to the nucleus and the mitochondria, foreign viral or bacterial DNA may occur in the endosome or the cytoplasm, depending on the route of viral entry and the place of viral replication. For example, Herpes simplex virus enters the cells via endocytosis while Vaccinia virus enters by fusion of its envelope with plasma membrane releasing the virus core into the cytosol where it also replicates. In those cases, a distinction based on structural elements is not needed as it can be made based on subcellular localization.

Additionally, it is also known that some DNA viruses like CMV and HSV are also recognized by TLR2 [Compton et al., 2003, Kurt-Jones et al. 2005]. It is assumed that this recognition involves the viral envelope glycoproteins [Rathinam and Fitzgerald 2011]. Taken together, it may be hypothesized that the cell discriminates between self and nonself agents on protein and RNA level as these structures are ubiquitous, but such a distinction may not be necessary for the DNA, as in healthy state, the presence of self DNA is limited to nucleus and mitochondria.

Our findings show that synthetic pieces of DNA shorter than 24 bps fail to induce IFN-β response. This is interesting as it suggests that DNA requires at least two helical turns to be recognized by the as-yet unknown DNA sensor. Following this observation one can speculate that DNA should have a certain length in order to allow binding of more than one

DNA sensor molecule or an additional co-receptor to form a complex on the DNA in order to activate the downstream signaling. Alternatively, one can hypothesize that DNA sensor scans along the DNA helix and for that it needs at least two helical turns. Identification of the DNA sensor will clarify these speculations.

Furthermore, our data show that both single- and double-stranded phosphodiester-containing ISD oligonucleotides (ISD-PO) can activate the interferon pathway. This is unanticipated as the original study used double-stranded ISD as an inducer [Stetson and Medzhitov 2006]. Our finding raises two possibilities: Either the as-yet unknown DNA sensor directly senses single-stranded DNA or the single-stranded DNA can become double-stranded during transfection (e.g. by hybridization with cellular mRNA), thereby rendering it immunostimulatory. The previous assumption does not exclude that the ssDNA and dsDNA may be recognized by distinct receptors, in analogy to TLR3 and TLR7. Regarding the second assumption, stimulation experiments with RNA/DNA-hybrids may give a hint if, in case of a hybridization with the cellular RNA, ssDNA becomes immunostimulatory.

Another interesting observation was made with ISD containing phosphorothioate backbones: Increasing amounts of ssISD-PS blocked IFN production induced by ISD-PO. Together with the observation that ISD-PS is capable of binding to DNA binding proteins, this experiment may suggest that ISD-PS binds to the DNA sensor as well, but possibly fails to induce a conformational change in the sensor that triggers the downstream signaling pathway (**Fig. 4.13**) This is reminiscent of the situation with TLR9 in the endosomes: TLR9 binds to and gets activated by the PO-backbone and this response can be antagonized by PS-containing oligonucleotides that lack CpG motifs [Haas et al. 2008]. It is tempting to speculate that in future ssDNA with phosphorothioate backbone can be used as a therapeutics to block aberrant cytosolic DNA recognition.

In the light of the many similarities between TLR9 signaling and signaling elicited by the as-yet unknown DNA sensor(s) (specificity for phosphodiester DNAs, utilization of similar downstream components, activation of similar cytokines and interferons), the gene signature we derived for cytosolic DNA recognition may be very helpful in distinguishing those responses – especially when TLR9 knockouts are unavailable, as in the human system. We also believe that the identity of the signature genes themselves may help better appreciate the biological response elicited by cytosolic DNA.

## 5.2 IFN-inducing cytosolic DNA sensors

Since the first reports of cytosolic DNA-dependent-IFN production, the identification of the cytosolic DNA sensor has been in the scope of many scientists [Ishii et al. 2006, Stetson and Medzhitov 2006]. Even though many proteins have been suggested as the sensor (DAI, IFI16, RNA polymerase III/RIG-I), there are many ambiguities and the ultimate prove, the IFN phenotype in the genetic deficient mice remained elusive.

In the case of DAI, the initial excitement that it may have been the long-awaited cytosolic DNA sensor was dimmed when it was published that the DAI-deficient mice failed to show any loss in the ability to sense DNA and trigger IFN [Takaoka et al., 2007, Ishii et al., 2008]. Recently, reports suggest that DAI, upon activation, recruits RIP1 and RIP3 which are the main kinases for induction of necroptosis, a programmed necrotic cell death [Rebsamen et al. 2009]. Viral proteins like M45 of mouse cytomegalovirus (MCMV), which suppresses necroptosis, have been found to block DAI-induced NF-κB activation [Rebsamen et al. 2009]. It is conceivable that DAI recognizes DNA upon viral infection and mainly induces cell death via RIP kinases [Rebsamen and Tschopp 2010]. Investigation of necroptotic effects of MCMV infections in DAI deficient cells and mice will shed light on this issue.

Regarding IFI16, the confirmation that it is a DNA sensor will only be obtained from genetically-deficient cells, generated from knockout mice [Unterholzner et al. 2010]. However, there are several murine orthologs and hence, it is not entirely clear which one would have to be knocked out to achieve a measurable effect. The closest homologue of IFI16 in mice is IFI204. The two proteins have the same domain structure but the amino acid identity is only 37%. Even though Unterholzner and colleagues show that the loss of IFI204 in mouse cell lines has similar effects as loss of IFI16 in human cells, it remains to be seen whether the IFI204 knockout mice will have a phenotype. Furthermore, IFI16 is mainly a nuclear protein and only a small fraction is present in the cytoplasm where it co-localizes with DNA. It is probable that IFI16 also enters the cytoplasm but how and where the initial binding occurs is unclear. Interestingly, IFI16 was recently shown to recognize Kaposi Sarcoma-associated herpesvirus in the nucleus and, contrary to what has initially been suggested, induces inflammasome activation [Kerur et al. 2011], but not type-I interferon induction. The importance of IFI16 and RNA polymerase III has been challenged by Sharma and colleagues which find no requirement for IFI16, RNA polymerase III/RIG-I, TLR9 and DAI in cytosolic DNA-dependent IFN induction [Sharma et al. 2011].

Of note, the majority of DNA viruses replicate in the nucleus. If the recognition of viral DNA occurs in the nucleus and no obvious structural features distinguish viral from host DNA, it is currently unclear how viral DNA would be sensed in the presence of excess host DNA in the nucleus. One possibility is that the host DNA is bound by nuclear proteins avoiding its innate immune recognition. It is also imaginable that the condensed form of genomic DNA is not immunostimulatory but in this case it is difficult to explain how the uncondensed part of the DNA (e.g. during RNA transcription) remain unrecognized.

Finally, it has also been reported that DNA recognition occurs via transcription of AT-rich DNA sequences into RNA by RNA polymerase III which is subsequently recognized by RIG-I. This is in contrast with the reports that DNA recognition is sequence independent [Stetson and Medzhitov 2006, Chiu et al. 2009, Ablasser et al. 2009, Kumar et al. 2006, Sun et al. 2006]. Of note, this mechanism of AT-rich sequence recognition is restricted to human cells as in mouse cells, redundant pathways are operational. AT-rich regions are encountered in the replication origins of viruses and eukaryotes and this may be one of the recognition patterns that the innate immune system employs. Interestingly, RNA molecules encoded by RNA polymerase III of the DNA virus Epstein Barr have been potent inducers of RIG-I dependent IFN pathway.

Taken together, two scenarios about the cytosolic DNA recognition seem possible: Either the major DNA sensor is still missing and the effects observed until today are artifacts hence the ambiguity in the data [compare Sharma et al 2010, Chiu et al. 2009 and Ablasser et al. 2009 or Takaoka et al., 2007 and Ishii et al., 2008] or the reason for wide range of observations derives from the existence of several, possibly redundant, DNA sensing receptors and mechanisms in mammals. It is possible that the majority of already identified proposed DNA sensors contribute to the innate immune response by recognizing different types of DNAs in different cell types sharing a common downstream signaling pathway. Only when every cytosolic DNA sensor is identified, we will be able to draw a comprehensive picture of innate immune recognition of nucleic acids.

## 5.3 Technical aspects of the systemic approach

In order to identify the elusive IFN-inducing cytosolic DNA receptor, we opted for nucleic acid affinity purifications followed by mass spectrometry analysis from cytoplasmic extracts of PBMCs and RAW264.7 cells. The nucleic acid purifications resulted in around 1600 proteins.

We hypothesized that the IFN-inducing cytosolic DNA receptor should be induced by IFN-β as this is the case for some other cytosolic receptors and other components of the pathway [Kang et al. 2004, Yoneyama et al, 2004, 2005]. Therefore, we treated the cells prior the affinity purification with IFN-β. But one has to bear in mind that it has not experimentally addressed whether or not the receptor is upregulated by IFN-β. It is possible that the elusive receptor is not regulated, as in the case for many TLRs. In fact, it is even plausible that the cell downregulates the cytosolic receptor as consequence of a negative feedback loop in order to switch the immune response off after a certain level of interferon has been produced. In this case, the IFN-β pretreatment would have downregulated the receptor rendering it hard to identify it by mass spectrometry. One possibility to investigate the IFN-dependent regulation is to prime the cells with IFN-β prior DNA stimulation and test whether the IFN-ß production alters compared to unprimed cells. At the same time, priming effects are complex events that may affect many components of the DNA response pathway and, hence, are difficult to interpret.

In order to identify the cytosolic DNA receptor, we aimed at identifying every cytosolic DNA binding protein using mass spectrometry. Mass spectrometry analysis is limited by the dynamic range of the instrument. Dynamic range is the ratio between the largest and smallest signal detectable by a mass spectrometer. Despite recent advancements in the field of mass spectrometry, it is still a major challenge to identify proteins present at low concentration from a mixture of proteins that are present at much higher abundance. One reason originates from the mass spectrometric algorithm where abundant peptides obscure the presence of lower abundant peptides in the ionic phase and are selected for identification. Therefore, clearly we may have missed some of the low abundant DNA binders in the cytosol during the mass spectrometry analysis and we therefore cannot know how comprehensive our proteomics dataset is. Nevertheless, the purification captured many proteins involved in cytosolic DNA recognition like TREX1, ZBP1/DAI, IFI16, AIM2 and all the members of HMGB family. This demonstrates the strength of our approach and that, by using this approach, relevant proteins can be identified.

Finally, despite the cytosolic extraction prior to the nucleic acid purifications, the list of captured proteins contained many generic nuclear DNA binders like histones, transcription factors or splicing factors. This may have been caused by the nuclear contamination of the cytoplasmic extracts, even though immunoblotting analysis of the extracts showed no specific band for the nuclear markers. Alternatively, the experimental process may have included proteins early after synthesis. In order to filter out these proteins and concentrate more on proteins with a higher probability of being the receptor, we decided to prioritize the proteins according to some predefined criteria.

Selecting around 50 candidates out of around 1,600 proteins for a functional screen was a major challenge. On the one hand, we wanted to select promising candidates and eliminate generic DNA binders; on the other hand, we aimed to stay unbiased. As a compromise, we established a scoring scheme that, although having some systematic bias as a consequence of the chosen parameters, is certainly more objective than “cherrypicking” individual proteins. The strength of a scoring scheme was to prioritize a list of proteins according to features which were defined a priori based on theoretical considerations. The weakness of such a scheme derived from the fact that certain criteria were necessarily hypothetical: For instance, there is no experimental proof that the DNA sensor is an IFN-inducible protein. Second, we do not know whether the sensor has a known established DNA-binding domain. In fact, it could well be that the sensor has escaped identification despite considerable efforts because it is a protein with unusual properties and does not have a conventional DNA binding domain.

Another important issue of the scoring scheme was the distribution of the points for each category. In order to end up with a list of proteins in which established sensors obtain relatively high scores, we decided to define the scoring scheme such that known DNA sensors would obtain high scores. For example, TREX1, an endonuclease which digests cytosolic DNA, thus preventing its recognition, obtained a total of 200 out of 200 possible points, positioned in the first place of the list. ZBP1/DAI, an alleged DNA sensor obtained 180 points ranked 24$^{th}$. It is noteworthy that IRF7, one of the master transcription factors for type I IFN, ranked with 200 points ex aequo with TREX1 in the first position. It is not surprising to find the IRF7 in cytosolic extracts as it is normally present in the cytosol and transfers to the nucleus only after being phosphorylated by TBK1. This also demonstrates the problem of sorting out sensors from transcription factors (TF). A TF involved in IFN-β signaling is very difficult to separate from a sensor which senses DNA, as they both bind DNA and likely to be regulated similarly by IFN-β. In addition, even the loss-of function experiment would look alike for a sensor and a transcription factor as both may result in

decrease of IFN-β production. Therefore, an IFN-regulated TF would score as high as the elusive DNA sensor in our scoring scheme as observed with IRF7. A way to distinguish a sensor from a TF could be to monitor the cellular localization of the respective proteins after DNA stimulation.

Interestingly, after adjusting the scoring scheme according to the known DNA sensors, IFI16, a cytosolic DNA sensor published after the establishment of the scheme, and its mouse homologue p204, ranked in the top 30 of the list demonstrating the strength of the scoring scheme.

Retrospectively, there may have been some other criteria, which could have improved the scheme. Firstly, additional points could have been distributed for proteins which are captured several times in DNA pulldowns that was carried out within this project i.e. in ISD-PO, ISD-PS, CT-DNA pulldowns. This would prioritize proteins based on the assumption that a sensor should equally recognize different DNAs regardless of specific sequence motifs. We also considered subtracting points from abundant proteins based on the assumption that, in a healthy state, a sensor should be expressed at lower levels and possibly becomes upregulated only upon infection. Given the fact that very low abundance proteins are not detected by MS analysis due to dynamic range issues, an abundant protein can be defined as a protein that is identified by mass spectrometry in many different cell lines without prior enrichment (central proteome) [Burkard et al. 2011]. But we discarded this idea when we noticed that some important innate immunity proteins like TREX1, AIM2 and HMGBs were part of the central proteome of a cell.

For the selection of the candidates, we screened the existing literature and domain composition for the first 250 proteins in each list. At this stage, we were looking for proteins with interesting domain composition like an ATP binding domain or an effector domain for initiation of downstream signaling. At the same time, we favored poorly characterized proteins and discarded well established ones like junB and IRF7. Some of the selected proteins had already a link to innate immunity although not very well established.

One such an example was the ECSIT protein. This protein has been reported to be involved in TLR dependent bacterial recognition [Moustakas et al. 2003, Kopp et al. 1999]. ECSIT is thought to act downstream of TLRs interacting with TRAF6 and activating downstream kinases to induce proinflammatory cytokines. However, these and other studies did not investigate the effect of ECSIT on IFN signaling after activation of cytosolic PRRs. We captured ECSIT exclusively in DNA pulldowns and according to gene ontology, it contains a

DNA binding domain [Ashburner et al. 2000]. It is tempting to speculate that ECSIT may be involved in DNA recognition. These observations render ECSIT an interesting candidate for further validation.

Another protein with a link to innate immunity was THEMIS2. Themis2 has been described as a scaffold protein that regulates TLR signaling and cytokine production in macrophages [Peirce et al.2010]. Like in the case of ECSIT, the effect of THEMIS2 on IFN-β production by cytosolic PRR has not been investigated. Despite the absence of any annotated DNA binding domain, we encountered THEMIS2 specifically in DNA pulldowns. The IFN inducibility and its specific occurrence in DNA pulldowns led us to include THEMIS2 in the loss-of-function validation.

SLFN (Schlafen) proteins are a family of proteins containing a DNA/RNA helicase domain and are reported to be involved in T cell development even though the mechanism is not clear [Geserick et al. 2004]. Recently, it has been reported that SLFN2 deficient mice succumb to bacterial and viral infection [Berger et al. 2010]. The report links this phenotype to the role of SLFN2 in T-cell quiescence. Although Berger and colleagues exclude the SLFN2 involvement in cytokine production by measuring the production many of them, the type I IFN levels were not tested. SLFN family members contain an RNA/DNA helicase domain and, similar to Themis2, we captured SLFN2 and SLFN5 exclusively in our DNA pulldowns. Considering that they are IFN inducible, it is possible that SLFN family members play a role in the IFN inducing cytosolic DNA recognition. Based on these reports and thoughts, we selected them for functional screen and expanded the list by other two IFN-β inducible members of the family, namely SLFN8 and SLFN9.

Finally, another candidate we selected was IFI16. This protein belongs to the same family as AIM2 and scored very high in our scheme. During the course of functional loss-of-function validation, Unterholzner and colleagues suggested IFI16 as a new DNA sensor [Unterholzner et al. 2010]. This finding underlines the validity of the entire experimental work-flow.

The loss-of-function validation resulted in 45 shRNAs giving a signal in the lower 25% of all observations. These 45 shRNAs belonged to around 20 genes. Considering that we had 46 genes tested in this assay, the number of genes with an IFN phenotype may appear to be high. On the other hand, one has to consider that these genes were chosen from a list that was enriched for proteins with the selected characteristics. Therefore, it is not unimaginable that several selected proteins are involved in the cytosolic DNA recognition.

One reason for the high number of the significant shRNAs may also derive from off-target effects that are known to be often associated with shRNA experiments. At this stage, we cannot exclude off-target effects although the algorithm applied for the design of the shRNA sequences employs strategies to reduce such effects [Reynolds et al. 2000]. In fact, we confirmed the knockdown efficiency of 28 shRNAs which showed an impact in the loss-of-function validation. Seven out of 28 shRNA did not show any reduction on the mRNA levels of respective proteins, which corresponds to 25% of selected shRNAs. One can speculate that 25% of the observed effects derive from off-target effects. Eventually, secondary validation assays like IFN-response in stable knockdown cell lines after DNA stimulation and overexpression of respective proteins will reveal more information about the role of the candidates in DNA recognition.

We carried out over six validation rounds but the assay was rather unstable. In some occasions, some of the internal control points failed to fulfill the criteria (e.g. at least 10-fold IFN-β production after DNA stimulation in comparison to untreated cells, decreased IFN production in cells transfected with shRNA2 and shRNA6 against TBK1, no morphological changes). There may be several reasons for the instability of the assay. First of all, one has to bear in mind that during one screen 292 target gene knockdown cell lines are created in 12-well plates and they were not passaged until the end of the assay. This means that the viral infection, the antibiotics selection, and the stimulation occurs in the same dish during eight days. This may cause too much stress for the created cell lines in a short time and affect the cellular response to stimulation.

Another point one has to take into consideration is the effect of shRNAs on IFN-β signaling. It is noteworthy, that in this assay we were monitoring the IFN-β response of knockdown cell lines after DNA-stimulation. The knockdown of proteins in the cells are achieved and maintained by constant production of shRNAs. It is well established that shRNAs themselves may induce in IFN-β via PRRs [Judge et al. 2005]. Although we did not observe any considerable upregulation of IFN-β after expression of shRNAs, it is imaginable that miniscule amounts of IFN-β may already have priming effects, hence result in regulation of the subsequent IFN-ß response upon DNA stimulation. As an additional complication, many of our DNA sensor candidates are IFN-inducible proteins, hence shRNA-triggered IFN induction may compromise the knockdown efficiency.

Finally, we wanted to implement two shRNA sets as positive controls. We decided for the central kinase for IFN-β pathway, TBK1, and one of the central IFN-β transcription factors,

IRF3. Unfortunately, none of the shRNAs against IRF3 showed any effect on the IRF3 expression level (data not shown). Therefore, we decided to use only TBK1 as positive control in the assays.

## 5.4 Future perspectives

We believe that we have generated a valuable list of proteins that contains promising candidates which are likely to play a role in the cytosolic DNA recognition. The top ten candidates obtained from the loss-of-function validation will undergo a secondary validation with stable knockdown cell lines and, this time, they will additionally be tested with other PAMPs in parallel to evaluate the DNA specificity (e.g. CpG, LPS, RNA). Apart from following the candidates obtained from the primary validation, one may also opt for an overexpression approach in RAW264.7 or in HEK293 cells. HEK293 cells are normally not responsive to nucleic acids. On the other hand, the transfection with TLRs or RLHs renders these cells responsive to nucleic acids (data not shown). Therefore, it is assumed that the missing component of IFN pathway in this cell line are the nucleic acid receptors even though we cannot exclude that in the case of DNA recognition, other downstream signaling components are missing as well. If the DNA sensor is the only missing component, any IFN induction upon DNA stimulation after the transfection of candidate protein would be a strong indication for a DNA sensor.

Once a candidate protein has been identified from the primary validation round, the DNA association will be confirmed and the localization before and after the DNA challenge monitored. This will further shed light on the mechanism of the DNA recognition. Finally, in order to map the sensor and downstream signaling partners, the characterization of the protein complexes formed by the protein of interest affinity purifications will be the method of choice.

Alternatively, viral replication can be used as a direct measurement of the effects of the candidate protein on viral infection. In this case, infection of cells with a DNA virus, instead of DNA transfection would be needed. In case of an inhibition of viral replication, further investigation will be necessary if the observed effects are IFN dependent.

Presumably, upon DNA recognition, distinct pathways are activated leading to cytokine production aside from type I IFNs. Therefore, it is conceivable that these pathways have also distinct receptors. For example, although the DNA viruses are mainly recognized by AIM2 for

inflammasome activation, some DNA viruses like HSV-1 have been shown to be recognized independently of AIM2 [Rathinam et al. 2010]. It will be interesting to identify additional inflammasome activating DNA receptors. Therefore, testing the effect of the selected candidates on additional read-outs (e.g. additional cytokines or cell death) may be very promising.

Until today, we still do not know how the majority of the viruses are recognized. Although it is proven that cytosolic DNA triggers innate immune response, it is not clear how a virus and bacteria is recognized in the cytosol. An interesting area of investigation will be to find out if the cytosolic receptors recognize plain DNA or DNA-viral protein complexes [Rathinam and Fitzgerald 2011].

Finally, I believe that my thesis has contributed to the characterization of the molecular mechanism of DNA recognition and has created an attractive short-list of possible DNA sensors. Identification of every receptor involved in cytosolic DNA recognition will complete our understanding of how bacteria, viruses and aberrant DNA activate or evade the innate immune system, consequently, generating novel drug targets for therapies for autoimmune diseases or against infections.

# 6 References

Ablasser, A., Bauernfeind, F., Hartmann, G., Latz, E., Fitzgerald, K.A., and Hornung, V. (2009). RIG-I-dependent sensing of poly(dA:dT) through the induction of an RNA polymerase III-transcribed RNA intermediate. Nat. Immunol. 10, 1065–1072.

Akira, S., Uematsu, S., and Takeuchi, O. (2006). Pathogen recognition and innate immunity. Cell 124: 783-801.

Alberts, B., Johnson, A., Lewis, J., Raff, M., Roberts, K., and Walter, P. (2007). Molecular Biology of the Cell, 5th edition. New York: Garland Publishing.

Alexopoulou, L., Holt, A. C., Medzhitov, R., and Flavell, R. A. (2001). Recognition of double-stranded RNA and activation of NF-κB by Toll-like receptor 3. Nature 413, 732-738.

Ashburner, M., Ball, C.A., Blake, J.A., Botstein, D., Butler, H., Cherry, J.M., Davis, A.P., Dolinski, K., Dwight, S.S., Eppig, J.T., Harris, M.A., Hill, D.P., Issel-Tarver, L., Kasarskis, A., Lewis, S., Matese, J.C., Richardson, J.E., Ringwald, M., Rubin, G.M., and Sherlock, G. (2000). Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nat. Genet. 25, 25-29.

Agrawal, S., Christodoulou, C. and Gait, M. J. (1986). Efficient methods for attaching non-radioactive labels to the 5' ends of synthetic oligodeoxyribonucleotides. Nucleic Acids Res. 14, 6227-6245.

Barbalat, R., Lau, L., Locksley, R.M., and Barton, G.M. (2009). Toll-like receptor 2 on inflammatory monocytes induces type I interferon in response to viral but not bacterial ligands. Nat. Immunol. 10, 1200–1207.

Barber, G.N. (2011). Innate immune DNA sensing pathways: STING, AIMII and the regulation of interferon production and inflammatory responses. Curr Opin Immunol. 23, 10-20.

Barton, G.M., and Kagan, J.C. (2009). A cell biological view of Toll-like receptor function: regulation through compartmentalization. Nat. Rev. Immunol. 9, 535–542.

Benko, S., Philpott, D.J., and Girardin, S.E. (2008). The microbial and danger signals that activate Nod-like receptors. Cytokine 43, 368-373.

Berger, M., Krebs, P., Crozat, K., Li, X., Croker, B.A., Siggs, O.M., Popkin, D., Du, X., Lawson, B.R., Theofilopoulos, A.N., Xia, Y., Khovananth, K., Moresco, E.M., Satoh, T., Takeuchi, O.., Akira, S., and Beutler, B. (2010). An Slfn2 mutation causes lymphoid and myeloid immunodeficiency due to loss of immune cell quiescence. Nat. Immunol. 11, 335-343.

Buerckstummer, T., Baumann, C., Bluml, S., Dixit, E., Durnberger, G., Jahn, H., Planyavsky, M., Bilban, M., Colinge, J., Bennett, K.L., and Superti-Furga, G. (2009) An orthogonal proteomic–genomic screen identifies AIM2 as a cytoplasmic DNA sensor for the inflammasome. Nat. Immunol. 10, 266-272.

Burkard, T.R., Planyavsky, M., Kaupe, I., Breitwieser, F.P., Bürckstümmer, T., Bennett, K.L., Superti-Furga, G., and Colinge, J. (2011). Initial characterization of the human central proteome. BMC Syst. Biol. 5, 17.

Chiu, Y. H., Macmillan, J. B., and Chen, Z. J. (2009). RNA polymerase III detects cytosolic DNA and induces type I interferons through the RIG-I pathway. Cell 138, 576–591.

Crow, Y.J., Hayward, B.E., Parmar, R., Robins, P., Leitch, A., Ali, M., Black, D.N., van Bokhoven, H., Brunner, H.G., Hamel, B.C., Corry, P.C., Cowan, F.M., Frints, S.G., Klepper, J., Livingston, J.H., Lynch, S.A., Massey, R.F., Meritet, J.F., Michaud, J.L., Ponsot, G., Voit, T., Lebon, P., Bonthron, D.T., Jackson, A.P., Barnes, D.E., and Lindahl, T. (2006). Mutations in the gene encoding the 30–50 DNA exonuclease TREX1 cause Aicardi-Goutieres syndrome at the AGS1 locus. Nat. Genet. 38, 917-920.

Diebold, S.S., Kaisho, T., Hemmi, H., Akira, S., and Reis E Sousa, C. (2004). Innate antiviral responses by means of TLR7-mediated recognition of single-stranded RNA. Science 303, 1529-1531.

Domon, B., and Aebersold, R. (2006). Mass spectrometry and protein analysis. Science 312, 212-217.

Ewald, S.E., Lee, B.L., Lau, L., Wickliffe, K.E., Shi, G.P., Chapman, H.A., and Barton, G.M. (2008). The ectodomain of Toll-like receptor 9 is cleaved to generate a functional receptor. Nature 456, 658-666.

Fernandes-Alnemri, T., Yu, J.W., Datta, P., Wu, J., and Alnemri, E.S. (2009). AIM2 activates the inflammasome and cell death in response to cytoplasmic DNA. Nature 458, 509-513.

Fink, S.L., Bergsbaken, T., and Cookson, B.T. (2008). Anthrax lethal toxin and Salmonella elicit the common cell death pathway of caspase-1-dependent pyroptosis via distinct mechanisms. Proc. Natl. Acad. Sci. USA 105, 4312-4317.

Foster, S.L., and Medzhitov, R. (2009). Gene-specific control of the TLR-induced inflammatory response. Clin. Immunol. 130, 7-15.

Geserick, P., Kaiser, F., Klemm, U., Kaufmann, S.H., and Zerrahn, J. (2004). Modulation of T cell development and activation by novel members of the Schlafen (slfn) gene family harbouring an RNA helicase-like motif. Int. Immunol. 16, 1535-1548.

Gilliet, M., Cao, W., and Liu, Y.J. (2008). Plasmacytoid dendritic cells: Sensing nucleic acids in viral infection and autoimmune diseases. Nat. Rev. Immunol. 8, 594–606.

Sharma, S., Deoliveira, R.B., Kalantari, P., Parroche, P., Goutagny, N., Jiang, Z., Chan, J., Bartholomeu, D.C., Lauw, F., Hall, J.P., Barber, G.N., Gazzinelli, R.T., Fitzgerald, K.A., and Golenbock. D.T. (2011). Innate Immune Recognition of an AT-Rich Stem-Loop DNA Motif in the Plasmodium falciparum Genome Immunity, Aug 3 [Epub ahead of print]

Goubau, D., Rehwinkel, J., and Reis e Sousa, C. (2010). PYHIN proteins: center stage in DNA sensing. Nat. Immunol. 11, 984-986.

Haas, T., Metzger, J., Schmitz, F., Heit, A., Muller, T., Latz, E., and Wagner, H. (2008). The DNA sugar backbone 2' deoxyribose determines toll-like receptor 9 activation. Immunity 28, 315-323.

Häcker, H., Redecke, V., Blagoev, B., Kratchmarova, I., Hsu, L.C., Wang, G.G., Kamps, M.P., Raz, E., Wagner, H., Häcker, G., Mann, M., Karin, M. (2006). Specificity in Toll-like receptor signalling through distinct effector functions of TRAF3 and TRAF6. Nature 439, 204–207.

Hayashi, F., Smith, K.D., Ozinsky, A., Hawn, T.R., Yi, E.C., Goodlett, D.R., Eng, J.K., Akira, S., Underhill, D.M., and Aderem, A. (2001). The innate immune response to bacterial flagellin is mediated by Toll-like receptor-5. Nature 410, 1099-1103.

Heil, F., Hemmi, H., Hochrein, H., Ampenberger, F., Kirschning, C., Akira, S., Lipford, G., Wagner, H., and Bauer, S. (2004). Species-specific recognition of singlestranded RNA via Toll-like receptor 7 and 8. Science 303, 1526-1529.

Hemmi, H., Takeuchi, O., Kawai, T., Kaisho, T., Sato, S., Sanjo, H., Matsumoto, M., Hoshino, K., Wagner, H., Takeda, K., Akira, S. (2000). A Toll-like receptor recognizes bacterial DNA. Nature 408, 740-745.

Hirschfeld, M., Weis, J.J., Toshchakov, V., Salkowski, C.A., Cody, M.J., Ward, D.C., Qureshi, N., Michalek, S.M., and Vogel, S.N. (2001). Signaling by Toll-like receptor 2 and 4 agonists results in differential gene expression in murine macrophages. Infect. Immun. 69, 1477-1482.

Honda, K., Takaoka, A., and Taniguchi, T. (2006). Type I interferon gene induction by the interferon regulatory factor family of transcription factors. Immunity 25, 349–360.

Honda, K., Yanai, H., Negishi, H., Asagiri, M., Sato, M., Mizutani, T., Shimada, N., Ohba, Y., Takaoka, A., Yoshida, N., and Taniguchi, T. (2005). IRF-7 is the master regulator of type-I interferon-dependent immune responses. Nature 434, 772–777.

Hornung, V., Ablasser, A., Charrel-Dennis, M., Bauernfeind, F., Horvath, G., Caffrey, D.R., Latz, E., and Fitzgerald, K.A. (2009). AIM2 recognizes cytosolic dsDNA and forms a caspase-1-activating inflammasome with ASC. Nature 458, 514-518.

Hornung, V., and Latz, E. (2010). Intracellular DNA recognition Nat. Rev. Immunol. 10, 123-130.

Hornung, V., Ellegast, J., Kim, S., Brzozka, K., Jung, A., Kato, H., Poeck, H., Akira, S., Conzelmann, K.K., Schlee, M., Endres, S., and Hartmann, G. (2006). 5’-Triphosphate RNA is the ligand for RIG-I. Science 314, 994–997.

Husebye, H., Aune, M.H., Stenvik, J., Samstad, E., Skjeldal, F., Halaas, O., Nilsen, N.J., Stenmark, H., Latz, E., Lien, E., Mollnes, T.E., Bakke, O., and Espevik, T. (2010). The Rab11a GTPase controls Toll-like receptor 4-induced activation of interferon regulatory factor-3 on phagosomes. Immunity 33, 583–596.

Irizarry, R.A., Bolstad, B.M., Collin, F., Cope, L.M., Hobbs, B., and Speed, T.P. (2003). Summaries of Affymetrix GeneChip probe level data. Nucleic Acids Res. 31, e15.

Ishii, K.J., Kawagoe, T., Koyama, S., Matsui, K., Kumar, H., Kawai, T., Uematsu, S., Takeuchi, O., Takeshita, F., Coban, C., and Akira, S. (2008). TANK-binding kinase-1 delineates innate and adaptive immune responses to DNA vaccines. Nature 451, 725–729.

Ishii, K. J., Coban, C., Kato, H., Takahashi, K., Torii, Y., Takeshita, F., Ludwig, H., Sutter, G., Suzuki, K., Hemmi, H., Sato, S., Yamamoto, M., Uematsu, S., Kawai, T., Takeuchi, O., and Akira, S. (2006). A Toll-like receptor-independent antiviral response induced by double-stranded B-form DNA. Nat Immunol. 7, 40-48.

Ishikawa, H., Ma, Z., and Barber, G.N. (2009). STING regulates intracellular DNA-mediated, type I interferon-dependent innate immunity. Nature 461, 788-792.

Ishikawa, H., and Barber, G.N. (2008). STING is an endoplasmic reticulum adaptor that facilitates innate immune signalling. Nature 455, 674–678.

Jaehning, J.A., and Roeder, R.G. (1977). Transcription of specific adenovirus genes in isolated nuclei by exogenous RNA polymerases. J. Biol. Chem. 252, 8753–8761.

Janeway, C., Murphy, K., Travers, P., and Walport, M. (2008) Janeway's Immunology. 7th edition. Garland Science, Taylor & Francis Group.

Jensen, K.E., Neal, A.L., Owens, R.E., and Warren, J. (1963). Interferon responses of chick embryo fibroblasts to nucleic acids and related compounds. Nature 200, 433–434.

Kurt-Jones, E.A., Belko, J., Yu, C., Newburger, P.E., Wang, J., Chan, M., Knipe, D.M., and Finberg, R.W., (2005). The role of toll-like receptors in herpes simplex infection in neonates. J. Infect. Dis. 191, 746–748.

Judge, A.D., Sood, V., Shaw, J.R., Fang, D., McClintock, K., and MacLachlan, I. (2005). Sequence-dependent stimulation of the mammalian innate immune response by synthetic siRNA. Nat. Biotechnol. 23, 457-62.

Kang, D.C., Gopalkrishnan, R.V., Lin, L., Randolph, A., Valerie, K., Pestka, S., and Fisher, P.B. (2004). Expression analysis and genomic characterization of human melanoma differentiation associated gene-5, mda-5: a novel type I interferon-responsive apoptosis-inducing gene. Oncogene 23, 1789–1800.

Kanneganti, T.D., Lamkanfi, M., and Nunez, G. (2007) Intracellular NOD-like receptors in host defense and disease. Immunity, 27, 549-559.

Karayel, E., Buerckstuemmer, T., Bilban, M., Duernberger, G., Weitzer, S., Martinez, J., Superti-Furga, G. (2009). TLR-independent DNA recognition pathway in murine macrophages: Ligand features and molecular signature, Eur. J. Immunol. 39, 1929-1936.

Kasitanon, N., Magder, L.S., and Petri, M. (2006). Predictors of survival in systemic lupus erythematosus. Medicine (Baltimore) 85, 147-156.

Kato, H., Takeuchi, O., Mikamo-Satoh, E., Hirai, R., Kawai, T., Matsushita, K., Hiiragi, A., Dermody, T.S., Fujita, T., and Akira, S. (2008). Length-dependent recognition of double-stranded ribonucleic acids by retinoic acid-inducible gene-I and melanoma differentiation-associated gene 5. J. Exp. Med. 205, 1601–1610.

Kawai, T. and Akira, S. (2006) TLR signaling. Cell Death Differ. 13, 816-825.

Kawai, T., and Akira, S. (2010). The role of pattern-recognition receptors in innate immunity: Update on Toll-like receptors. Nat. Immunol. 11, 373–384.

Kawai, T., Takahashi, K., Sato, S., Coban, C., Kumar, H., Kato, H., Ishii, K.J., Takeuchi, O., and Akira, S. (2005). IPS-1, an adaptor triggering RIG-I- and Mda5-mediated type I interferon induction. Nat. Immunol. 6, 981–988.

Kerur, N., Veettil, M.V., Sharma-Walia, N., Bottero, V., Sadagopan, S., Otageri, P., and Chandran, B. (2011). IFI16 acts as a nuclear pathogen sensor to induce the inflammasome in response to Kaposi Sarcoma-associated herpesvirus infection. Cell Host Microbe. 9, 363-75.

Köcher, T., and Superti-Furga, G. (2007). Mass spectrometry-based functional proteomics: from molecular machines to protein networks. Nat. Methods. 4, 807-15.

Kopp, E., Medzhitov, R., Carothers, J., Xiao, C., Douglas, I., Janeway, C.A., and Ghosh, S. (1999). ECSIT is an evolutionarily conserved intermediate in the Toll/IL-1 signal transduction pathway. Genes Dev. 13, 2059-71.

Kumar, H., Kawai, T., Kato, H., Sato, S., Takahashi, K., Coban, C., Yamamoto, M., Uematsu, S., Ishii, K.J., Takeuchi, O., and Akira, S. (2006). Essential role of IPS-1 in innate immune responses against RNA viruses. J. Exp. Med. 203, 1795–1803.

Latz, E. (2010). The inflammasomes: mechanisms of activation and function. Curr. Opin. Immunol. 22, 28-33.

Lee-Kirsch, M.A., Gong, M., Chowdhury, D., Senenko, L., Engel, K., Lee, Y.A., de Silva, U., Bailey, S.L., Witte, T., Vyse, T.J., Kere, J., Pfeiffer, C., Harvey, S., Wong, A, Koskenmies S, Hummel O, Rohde K, Schmidt RE, Dominiczak AF, Gahr M, Hollis T., Perrino, F.W., Lieberman, J., and Hübner, N. (2007). Mutations in the gene encoding the 30–50 DNA exonuclease TREX1 are associated with systemic lupus erythematosus. Nat. Genet. 39, 1065-1067.

Levy, O. (2000). Antibacterial proteins and peptides of blood: templates for novel antibacterial agents. Blood 96, 2664-2672.

Loo, Y.M., Gale, M. Jr. (2011). Immune signaling by RIG-I-like receptors. Immunity 34, 680-92.

Lund, J. M., Alexopoulou, L., Sato, A., Karow, M., Adams, N. C., Gale, N. W., Iwasaki, A. and Flavell, R. A. 2004. Recognition of single-stranded RNA viruses by Toll-like receptor 7. Proc. Natl. Acad. Sci. USA 101, 5598-5603.

Mallick, P., and Kuster, B. (2010). Proteomics: a pragmatic perspective. Nat. Biotechnol. 28, 695-709.

Marshak-Rothstein, A., and Rifkin, I.R. (2007). Immunologically active autoantigens: the role of toll-like receptors in the development of chronic inflammatory disease. Annu. Rev. Immunol. 25, 419-41.

Meylan, E., Curran, J., Hofmann, K., Moradpour, D., Binder, M., Bartenschlager, R., and Tschopp, J. (2005). Cardif is an adaptor protein in the RIG-I antiviral pathway and is targeted by hepatitis C virus. Nature 437, 1167–1172.

Moustakas, A., and Heldin, C.H. (2003). Ecsit-ement on the crossroads of Toll and BMP signal transduction. Genes Dev. 17, 2855-2859.

Muruve, D. A., Petrilli, V., Zaiss, A. K., White, L. R., Clark, S. A., Ross, P. J., Parks, R. J. and Tschopp, J., (2008). The inflammasome recognizes cytosolic microbial and host DNA and triggers an innate immune response. Nature 452, 103-107.

Napirei, M., Karsunky, H., Zevnik, B., Stephan, H., Mannherz, H.G., and Moroy, T. (2000). Features of systemic lupus erythematosus in Dnase1-deficient mice. Nat. Genet, 25, 177-181.

Oganesyan, G., Saha, S.K., Guo, B., He, J.Q., Shahangian, A., Zarnegar, B., Perry, A., and Cheng, G. (2006). Critical role of TRAF3 in the Toll-like receptor-dependent and -independent antiviral response. Nature 439, 208–211.

Philpott, D.J., and Girardin, S.E. (2010). Nod-like receptors: sentinels at host membranes. Curr. Opin. Immunol. 22, 428-34.

Palsson-McDermott, E. M. and O'Neill, L. A. (2007). Building an immune system from nine domains. Biochem. Soc. Trans. 35, 1437-1444.

Park, B., Brinkmann, M. M., Spooner, E., Lee, C. C., Kim, Y. M., and Ploegh, H. L. (2008). Proteolytic cleavage in an endolysosomal compartment is required for activation of Toll-like receptor 9. Nat Immunol. 9, 1407-1414.

Park, B.S., Song, D.H., Kim, H.M., Choi, B.S., Lee, H., and Lee, J.O. (2009). The structural basis of lipopolysaccharide recognition by the TLR4-MD-2 complex. Nature 458, 1191-1195.

Peirce, M.J., Brook, M., Morrice, N., Snelgrove, R., Begum, S., Lanfrancotti, A., Notley, C., Hussell, T., Cope, A.P., and Wait, R. (2010). Themis2/ICB1 is a signaling scaffold that selectively regulates macrophage Toll-like receptor signaling and cytokine production. PLoS One. 5, e11465.

Pichlmair, A., Reis e Sousa, C. (2007). Innate recognition of viruses. Immunity 27, 370-383.

Pichlmair, A., Schulz, O., Tan, C.P., Rehwinkel, J., Kato, H., Takeuchi, O., Akira, S., Way, M., Schiavo, G., Reis e Sousa, C. (2009). Activation of MDA5 requires higher-order RNA structures generated during virus infection. J. Virol. 83, 10761–10769.

Pichlmair, A., Schulz, O., Tan, C.P., Naslund, T.I., Liljestrom, P., Weber, F., and Reis e Sousa, C. (2006). RIG-I-mediated antiviral responses to single-stranded RNA bearing 50-phosphates. Science 314, 997–1001.

Ramakrishnan, V. (2002). Ribosome structure and the mechanism of translation. Cell 108, 557–572.

Raschke, W.C., Baird, S., Ralph, P., and Nakoinz, I. (1978). Functional macrophage cell lines transformed by Abelson leukemia virus. Cell 15, 261-267.

Rathinam, V.A., and Fitzgerald, K.A. (2011). Innate immune sensing of DNA viruses. Virology 411, 153-162.

Rathinam, V.A., Jiang, Z., Waggoner, S.N., Sharma, S., Cole, L.E., Waggoner, L., Vanaja, S.K., Monks, B.G., Ganesan, S., Latz, E., Hornung, V., Vogel, S.N., Szomolanyi-Tsuda, E., and Fitzgerald, K.A. (2010). The AIM2 inflammasome is essential for host defense against cytosolic bacteria and DNA viruses. Nat. Immunol. 11, 395-402.

Rathinam, V.A., Sharma, S., and Fitzgerald, K.A. (2010). Catenin' on to nucleic acid sensing. Nat. Immunol. 11, 466-468.

Rebsamen, M., Heinz, L.X., Meylan, E., Michallet, M.C., Schroder, K., Hofmann, K., Vazquez, J., Benedic, C.A., and Tschopp, J. (2009). DAI/ZBP1 recruits RIP1 and RIP3 through RIP homotypic interaction motifs to activate NF-kappaB. EMBO Rep. 10, 916-922.

Rebsamen, M., and Tschopp, J. (2010). Viruses under the control of RIP kinases. Cell Cycle. 9, 438-439.

Rehwinkel, J., Tan, C.P., Goubau, D., Schulz, O., Pichlmair, A., Bier, K., Robb, N., Vreede, F., Barclay, W., Fodor, E., Reis e Sousa, C. (2010). RIG-I detects viral genomic RNA during negative-strand RNA virus infection. Cell 140, 397–408.

Reynolds, A., Leake, D., Boese, Q., Scaringe, S., Marshall, W.S., and Khvorova A. (2004). Rational siRNA design for RNA interference. Nat Biotechnol. 22, 326-30.

Roberts, T.L., Idris, A., Dunn, J.A., Kelly, G.M., Burnton, C.M., Hodgson, S., Hardy, L.L., Garceau, V., Sweet, M.J., Ross, I.L., Hume, D.A., and Stacey, K.J. (2009). HIN-200 proteins regulate caspase activation in response to foreign cytoplasmic DNA. Science, 323, 1057-1060.

Rönnblom, L., Alm, G.V., and Eloranta, M.L. (2006). The type I interferon system in systemic lupus erythematosus. Arthritis Rheum. 54, 408-420.

Rotem, Z., Cox, R.A., and Isaacs, A. (1963). Inhibition of virus multiplication by foreign nucleic acid. Nature 197, 564–566.

Rothenfusser, S., Goutagny, N., DiPerna, G., Gong, M., Monks, B.G., Schoenemeyer, A., Yamamoto, M., Akira, S., and Fitzgerald, K.A. (2005). The RNA helicase Lgp2 inhibits TLR-independent sensing of viral replication by retinoic acid-inducible gene-I. J. Immunol. 175, 5260–5268.

Samarajiwa, S.A., Forster, S., Auchettl, K., and Hertzog, P.J., (2009). INTERFEROME: the database of interferon regulated genes. Nucleic Acids Research 37(Database Issue): D852-7.

Saito, T., Hirai, R., Loo, Y.M., Owen, D., Johnson, C.L., Sinha, S.C., Akira, S., Fujita, T., and Gale, M., Jr. (2007). Regulation of innate antiviral defenses through a shared repressor domain in RIG-I and LGP2. Proc. Natl. Acad. Sci. USA 104, 582–587.

Saito, T., Owen, D.M., Jiang, F., Marcotrigiano, J., and Gale, M., Jr. (2008). Innate immunity induced by composition-dependent RIG-I recognition of hepatitis C virus RNA. Nature 454, 523–527.

Sasai, M., Linehan, M.M., and Iwasaki, A. (2010). Bifurcation of Toll-like receptor 9 signaling by adaptor protein 3. Science 329, 1530–1534.

Sasai, M., Shingai, M., Funami, K., Yoneyama, M., Fujita, T., Matsumoto, M., and Seya, T. (2006). NAK-associated protein 1 participates in both the TLR3 and the cytoplasmic pathways in type I IFN induction. J. Immunol. 177, 8676–8683.

Schroder, K., and Tschopp, J. (2010) The inflammasomes. Cell 140, 821-832.

Seth, R., Sun, L., Ea, C., and Chen, Z. (2005). Identification and characterization of MAVS, a mitochondrial antiviral signaling protein that activates NF-kappaB and IRF 3. Cell 122, 669–682.

Smith, M. F. Jr., Mitchell, A., Li, G., Ding, S., Fitzmaurice, A. M., Ryan, K., Crowe, S. and Goldberg, J. B. (2003). Toll-like receptor (TLR) 2 and TLR5, but not TLR4, are required for Helicobacter pylori-induced NF-κB activation and chemokine expression by epithelial cells. J. Biol. Chem. 278, 32552-32560.

Stetson, D. B., and Medzhitov, R., Recognition of cytosolic DNA activates an IRF3-dependent innate immune response. Immunity 2006. 24: 93-103.

Stetson, D. B., Ko, J. S., Heidmann, T., and Medzhitov, R. (2008). Trex1 prevents cell-intrinsic initiation of autoimmunity. Cell 134, 587–598.

Sun, Q., Sun, L., Liu, H., Chen, X., Seth, R., Forman, J., and Chen, Z. (2006). The specific and essential role of MAVS in antiviral innate immune responses. Immunity 24, 633–642.

Sun, W., Li, Y., Chen, L., Chen, H., You, F., Zhou, X., Zhou, Y., Zhai, Z., Chen, D., Jiang, Z. (2009). ERIS, an endoplasmic reticulum IFN stimulator, activates innate immune signaling through dimerization. Proc. Natl. Acad. Sci. USA 106, 8653–8658.

Tabeta, K., Hoebe, K., Janssen, E.M., Du, X., Georgel, P., Crozat, K., Mudd, S., Mann, N., Sovath, S., Goode, J., Shamel, L., Herskovits, A.A., Portnoy, D.A., Cooke, M., Tarantino, L.M., Wiltshire, T., Steinberg, B.E., Grinstein, S., Beutler, B. (2006). The Unc93b1 mutation 3d disrupts exogenous antigen presentation and signaling via Toll-like receptors 3, 7 and 9. Nat. Immunol. 7, 156-164.

Takaoka, A., Wang, Z., Choi, M.K., Yanai, H., Negishi, H., Ban, T., Lu, Y., Miyagishi, M., Kodama, T., Honda, K., Ohba, Y., and Taniguchi, T. (2007). DAI (DLM-1/ZBP1) is a cytosolic DNA sensor and an activator of innate immune response. Nature 448, 501–505.

Takeda, K., Kaisho, T. and Akira, S. (2003). Toll-like receptors. Annu. Rev. Immunol. 21, 335-376.

Takeuchi, O., and Akira, S. (2008). MDA5/RIG-I and virus recognition. Curr. Opin. Immunol. 20, 17-22.

Trager, J., and Ward, M.M. (2001). Mortality and causes of death in systemic lupus erythematosus. Curr. Opin. Rheumatol. 13, 345-51.

Tschopp, J., and Schroder, K. (2010). NLRP3 inflammasome activation: the convergence of multiple signalling pathways on ROS production? Nat Rev Immunol, 10, 210-215.

Unterholzner, L., Keating, S.E., Baran, M., Horan, K.A., Jensen, S.B., Sharma, S., Sirois, C.M., Jin, T., Latz, E., Xiao, T.S., Fitzgerald, K.A., Paludan, S.R., and Bowie, A.G. (2010). IFI16 is an innate immune sensor for intracellular DNA. Nat. Immunol. 11, 997-1004.

Venkataraman, T., Valdes, M., Elsby, R., Kakuta, S., Caceres, G., Saijo, S., Iwakura, Y., and Barber, G.N. (2007). Loss of DExD/H box RNA helicase LGP2 manifests disparate antiviral responses. J. Immunol. 178, 6444–6455.

Wang, Z., Choi, M.K., Ban, T., Yanai, H., Negishi H., Lu, Y., Tamura, T., Takaoka, A., Nishikura, K., and Taniguchi, T. (2008). Regulation of innate immune responses by DAI (DLM-1/ZBP1) and other DNA-sensing molecules. Proc. Natl Acad. Sci. USA 105, 5477–5482.

Weitzer, S. and Martinez, J. (2007). The human RNA kinase hClp1 is active on 3' transfer RNA exons and short interfering RNAs. Nature 447, 222-226.

Werts, C., Tapping, R.I., Mathison, J.C., Chuang, T.H., Kravchenko, V., Saint Girons, I., Haake, D.A., Godowski, P.J., Hayashi, F., Ozinsky, A., Underhill, D.M., Kirschning, C.J., Wagner, H., Aderem, A., Tobias, P.S., and Ulevitch, R.J. (2001) Leptospiral lipopolysaccharide activates cells through a TLR2-dependent mechanism. Nat. Immunol. 2, 346-352.

Xu, L.G., Wang, Y.Y., Han, K.J., Li, L.Y., Zhai, Z., and Shu, H.B. (2005). VISA is an adapter protein required for virus-triggered IFN-beta signaling. Mol. Cell 19, 727–740.

Yanai, H., Ban, T., Wang, Z., Choi, M.K., Kawamura, T., Negishi, H., Nakasato, M., Lu, Y., Hangai, S., Koshiba, R., Savitsky, D., Ronfani, L., Akira, S., Bianchi, M.E., Honda, K., Tamura, T., Kodama, T., and Taniguchi, T. (2009). HMGB proteins function as universal sentinels for nucleic-acid-mediated innate immune responses. Nature 462, 99-103.

Yang, P., An, H., Liu, X., Wen, M., Zheng, Y., Rui, Y., Cao, X. (2010). The cytosolic nucleic acid sensor LRRFIP1 mediates the production of type I interferon via a beta-catenin-dependent pathway. Nat. Immunol. 11, 487-494.

Yasutomo, K., Horiuchi, T., Kagami, S., Tsukamoto, H., Hashimura, C., Urushihara, M., and Kuroda, Y. (2001). Mutation of DNASE1 in people with systemic lupus erythematosus. Nat. Genet, 28, 313-314.

Yoneyama, M., Kikuchi, M., Matsumoto, K., Imaizumi, T., Miyagishi, M., Taira, K., Foy, E., Loo, Y.M., Gale, M., Jr., Akira, S., Yonehara, S., Kato, A., and Fujita, T. (2005). Shared and unique functions of the DExD/H-box helicases RIG-I, MDA5, and LGP2 in antiviral innate immunity. J. Immunol. 175, 2851–2858.

Yoneyama, M., Kikuchi, M., Natsukawa, T., Shinobu, N., Imaizumi, T., Miyagishi, M., Taira, K., Akira, S., and Fujita, T. (2004). The RNA helicase RIG-I has an essential function in double-stranded RNA-induced innate antiviral responses. Nat. Immunol. 5, 730–737.

Yoshida, H., Okabe, Y., Kawane, K., Fukuyama, H., and Nagata, S. (2005). Lethal anemia caused by interferon-β produced in mouse embryos carrying undigested DNA. Nature Immunol. 6, 49–56.

Zhong, B., Yang, Y., Li, S., Wang, Y.Y., Li, Y., Diao, F., Lei, C., He, X., Zhang, L., Tien, P., and Shu, H.B. (2008). The adaptor protein MITA links virus-sensing receptors to IRF3 transcription factor activation. Immunity 29, 538–550.

# 7 Abbreviations

| | |
|---|---|
| AIM2 | Absent in melanoma 2 |
| ATP | Adenosine triphosphate |
| CARD | Caspase activation and recruitment domain |
| CMV | Cytomegalovirus |
| CT-DNA | Calf thymus DNA |
| DAI | DNA-dependent activator of IFN-regulatory factors |
| DC | Dendritic Cells |
| DNA | Deoxyribonucleic acid |
| DNase | Deoxyribonuclease |
| DTT | 1,4-Dithiothreitol |
| EDTA | Ethylenediaminetetraacetic acid |
| EGTA | Ethylene glycol tetraacetic acid |
| ELISA | Enzyme-linked immunosorbent assay |
| EMCV | Encephalomyocarditis Virus |
| HEPES | 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid |
| HMGB | High mobility group box 1 |
| HSV | Herpes simplex virus |
| IFI16 | Interferon gamma-inducible protein 16 |
| IFN-β | Interferon-β |
| IKK | Inhibitor of kappaB kinase |
| IPS-1 | IFN-β promoter stimulator 1 |
| IRAK | Interleukin-1 receptor-associated kinase |
| IRF3 | Interferon regulatory factor 3 |
| ISD | Interferon stimulatory DNA |
| ISD-PO | Interferon stimulatory DNA with phosphorothioate backbone |
| ISD-PS | Interferon stimulatory DNA with phosphodiester backbone |
| LPS | Lipopolysaccharide |
| MDA5 | Melanoma differentiation associated gene 5 |
| MHC | Major histocompatibility complex |
| MS | Mass spectrometry |
| MyD88 | Myeloid differentiation primary response gene (88) |
| NF-κB | Nuclear factor-κB |
| NLR | NOD-like receptor |
| PAGE | Polyacrylamide gel electrophoresis |
| PAMP | Pathogen-associated molecular pattern |
| PBMC | Peripheral blood mononuclear cells |
| PBS | Phosphate-buffered saline |
| PRR | Pattern recognition receptor |
| RIG-I | Retinoic acid-inducible gene-I |
| RLH | RIG-I-like helicase |
| RNA | Ribonucleic acid |
| SDS | Sodium dodecyl sulfate |
| STING | Stimulator of Interferon genes |
| TAK1 | TGF-beta-activated kinase 1 |
| TBK1 | TANK-binding kinase 1 |
| TF | Transcription factor |

| | |
|---|---|
| TIR | Toll-interleukin 1 receptor |
| TIRAP | TIR domain containing adaptor protein |
| TLR | Toll-like receptor |
| TNF-α | Tumor necrosis factor-alpha |
| TRAF6 | TNF receptor-associated factor 6 |
| TREX1 | Three prime repair exonuclease 1 |
| Tris | Tris(hydroxyamino)methane |
| VSV | Vesicular stomatitis virus |

# 8 Appendix

The siRNA sequences of shRNAs against DNA sensor candidates:

| Name | shRNA | siRNA sequence (5'-3') |
|---|---|---|
| Aebp1 | 1 | GAAGCCAACGTCACATCAATA |
| | 2 | GCCCGTTTCATCCGCATCTAT |
| | 3 | GGCCTTACATTTCCACTCACA |
| | 4 | GGCCACTCAGTGCAACTTCAT |
| | 5 | GCACCACAGCTACAAGGACAT |
| | 6 | GACCCCTGTCTACAGCTACTA |
| Ank3 | 1 | GCCCCAGTCGTGGGTATTTAA |
| | 2 | CCCGGAAGAGTTGGGTACAAA |
| | 3 | CAGCGCAAGCTCGAAGACTAA |
| | 4 | GAAGGACCAATATTTGATTAT |
| | 5 | GTGGATGAAATCAACCAAATA |
| | 6 | CTCGCTGAGGTGCTTCTGTAT |
| Ascc3 | 1 | GTGGTAGTTATGGGTTATCAA |
| | 2 | CAGCCACTTCTATATTAAATA |
| | 3 | CAGCCAGGACTCCATAAACAA |
| | 4 | CCCAAAGCAGTGGTTGAATAT |
| | 5 | CGGAGTACTTTCTAGCTCTTA |
| | 6 | CCCTGATTTCTCTTGGAATGA |
| Btaf1 | 1 | GTGGCAACCTTTGTATATTTA |
| | 2 | GACTGCTAATTACAGGATTAT |
| | 3 | CTGGTTAGCCTTTCTCAATAA |
| | 4 | GGAGCTGAATTTGCCTTAACA |
| | 5 | GGGGGATCATTACACTCTACA |
| | 6 | CAAGGTCAAATCCAGAATAAA |
| Cpsf7 | 1 | CGAGGACCCCTAATTTATAAA |
| | 2 | GCACCGAGACTTGCTTCATAA |
| | 3 | GGTCCTCATCTCCTCTCTTAA |
| | 4 | CCCCTGCAATTCTGTATACAT |
| | 5 | CCTCCTCGATTACCTCCTCAT |
| | 6 | CTCCACCAGATACTTACATGA |

| Name | shRNA | siRNA sequence (5'-3') |
|---|---|---|
| AI607873 | 1 | GCTCCTCACATGTCTTTAATA |
| | 2 | CCTGGCTCCATATGCAACTCT |
| | 3 | CCCAACAACATCCAGCAATCT |
| | 4 | GGGTTCCTCTTATGCCTTCAA |
| | 5 | CCCAGGTTCCTTATGTACCAT |
| | 6 | GCTCCACAGTTGTGTCCAGTA |
| Ankhd1 | 1 | CCTGAATATGTTCTCTGGTTA |
| | 2 | GGGCACCTCATATTGGAAACA |
| | 3 | GGCCTGCCGATTTCCATGTAT |
| | 4 | CTCGCTTACCTATGGCACAAT |
| | 5 | GCTACAAATCTTTGCCATTAA |
| | 6 | CACGCCTTGATGGTGAAGTAA |
| Bst2 | 1 | GAGGGAGTTCTGCTTTATTGC |
| | 2 | CGGGGTTACCTTAGTCATCCT |
| | 3 | GCTCAAAGTGTCACTGTTCCT |
| | 4 | CTGGAGAATCTGAGGATCCAA |
| | 5 | CATGGTGGTCTCCAGCCTACT |
| | 6 | CCCGTGCCCATGGATGAGATG |
| Centg3 | 1 | CCCCAACATCTACGCCATCTA |
| | 2 | GCTCTCCAACTCGGCCGCTAT |
| | 3 | CATCGAGGACTCATTTGTGAA |
| | 4 | CCAGATCCGGGAGCACGTCAT |
| | 5 | GCGGGGGGCGGCGGGCAGTTC |
| | 6 | GGCGGCCCCCCCCAGCAGTTC |
| Dna2l | 1 | CCCCTATGTTTGGGGGTATCA |
| | 2 | GACGTCTTAATGTTGCTATAA |
| | 3 | CAGGCAGCAATTAAAGATCAT |
| | 4 | GACACTTGGATAATAGATAAA |
| | 5 | GCGGAGCTATTTCAGAAGAAA |
| | 6 | GTGCCCTCACTAAATTGCTAT |

| | | |
|---|---|---|
| Ecsit | 1 | CCAGCGCATCTTCGTCCACTA |
| | 2 | GCGGGACTTGTCAGTATACAA |
| | 3 | GTGCCCATGATCAGGCAACAT |
| | 4 | CTCTCTGGCTTCGTAATAAGT |
| | 5 | GTTCACCCGATTCAAGAATAT |
| | 6 | CCCAGAAGAATGGGAGCTGTA |
| Fiz1 | 1 | CCCACTCTCCTTAGAGTAATA |
| | 2 | CTGCTCAGTATGTTGCAACGT |
| | 3 | GTGGCAAGAGTTTCCGTTACC |
| | 4 | GCTTCAAGCATAGCTTCAACT |
| | 5 | GGACGAGGGCTCTGGGAATGA |
| | 6 | GGCGCCACTTTGCTCGACACA |
| Ifi204 | 1 | TGTGTATAAGCCTATTGAAAT |
| | 2 | AATGCCAGCCCTAAGATCTGT |
| Ighmbp2 | 1 | CTGGGTTTGGTAACTATATGA |
| | 2 | GGCGGATTAATGTTGCTGTTA |
| | 3 | CCCTGAGAACTACACCCATGA |
| | 4 | CCTCGTCACTTTGCACATCCA |
| | 5 | GTGGTGGAAATAATCCTTCAA |
| | 6 | CCCCACTCTCTTTCTACAACA |
| Isg20 | 1 | CCTCAAGTTCTCCATGAATGA |
| | 2 | CGCAGTCCTGTATGACAAGTA |
| | 3 | CAGACTGAACTTCATCCTCAT |
| | 4 | CGCTGCAGCATTGTGAACATC |
| | 5 | GAGGGAGAGATCACGGACTAC |
| | 6 | GCCACTGCTCTGTGGAAGATG |
| No66 | 1 | GTCGGGTTCTTACCTTGCTAA |
| | 2 | GCCCCGAGACTTCATGGATTA |
| | 3 | GGAGGCCGATTGTTCCTGTAT |
| | 4 | CACCTCTAGTTCCGAGTTAGT |
| | 5 | GGCAACCATGTTATATGATAA |
| | 6 | GGAACCCAAGTGCTTAGAAAT |
| Obfc1 | 1 | GCTGGTTTAATCACTGTGTCT |
| | 2 | GAGGAGGCACTAAACAATAAA |
| | 3 | GCGAGAGATCTGTGCCAACAT |
| | 4 | GCAGAAGATCTACCACATCAT |
| | 5 | GTGGCTCTGATAAGCTGTACT |
| | 6 | CCAGCAGGAGTTAGAGACAGT |

| | | |
|---|---|---|
| Epsti1 | 1 | GATGTATGTCTCTCTAATTAA |
| | 2 | CTGGGGTATATGAGAAATCGT |
| | 3 | GCTGTTGCAATATGAATAGTA |
| | 4 | GGGATGGACATCAGAAGAATA |
| | 5 | GGGATGAATCTATTCGAATCA |
| | 6 | CTGCAACAGATGCGTTCTAAA |
| Ifi202b | 1 | TGAGTGATGTAACCCCAATTC |
| | 2 | GATGAAAGACTCATAGAACAT |
| | 3 | CCCCTTCCAGTGATTCATCTG |
| | 4 | GTTTCTGAGATCTACGAGGTA |
| | 5 | CCTAACCAAATTATTGAAGTG |
| | 6 | GCCTCTCCTGGACCTAACAAA |
| Irgm | 1 | CAGGGACAAGTAGTGATTAAA |
| | 2 | GCCCAAACCGTAGAGGACTAT |
| | 3 | CTTCCGTTTGTTGAGATTTCT |
| | 4 | GCTGATGACATGTGCAATTGT |
| | 5 | CGACTGATATTTGGTGTAGAT |
| | 6 | GAGAAGATCGTTGGTGATAAA |
| Mpeg1 | 1 | GCGGGTGGGGTTTATACAAAT |
| | 2 | GCCTCTGCATTTCTTCATTAA |
| | 3 | CGGGTACCACTGGATTTCAAA |
| | 4 | CCACCAACACTGTCATAGTGA |
| | 5 | CAAGCCAATATGGATGATGAT |
| | 6 | GGCACACACGTAATCACTAGT |
| Nufip1 | 1 | GCAGCCAGAATCACTTTAGAT |
| | 2 | CCAGTTCCATTGGAGAAATAT |
| | 3 | GTGTGTTCGATATATCATCAA |
| | 4 | CTCCAGATATTCGACATGAAA |
| | 5 | GTCGTTGGACTAGGAAATCAT |
| | 6 | GCTCCTTTAGTGCACATGAGA |
| Ppfibp1 | 1 | GTTTGGAGCTTGGACTTTAAT |
| | 2 | CACGCACATGTTAAAGGAAGA |
| | 3 | CCTTCACTATTGCCGCCATCT |
| | 4 | GGACCTCCTTAGAGACACAGA |
| | 5 | AAGTCCTCGATCGAGATGAAA |
| | 6 | CTGACGGCCGTAGAGAAGGAC |

| | | |
|---|---|---|
| Preb | 1 | CCAGCTAGATTGGACTATTAA |
| | 2 | CCGGGATTTCTCTAACCTCTA |
| | 3 | GCCCATGGCATTGTGGTAACA |
| | 4 | CAGCCTCTTCTAATACACCAT |
| | 5 | CCCGAAGGGGTTGAACTCAAA |
| | 6 | CTGAGAAGGGTTGTGGTCCAA |
| Pyhin1 | 1 | GCCTCCTAAGTGGATCTATAT |
| | 2 | GACAGAAGAGAAGATGTTCCA |
| | 3 | CTCAAACTATGTTGGCTGCAA |
| | 4 | CCACTTGGCTTGGGGCATATA |
| | 5 | GTGCTACTGATTAACATTAGA |
| | 6 | GCTCACAGTTTAATGGTAGAT |
| Rbms1 | 1 | GCTGGTCTAGTGCACCTTGTA |
| | 2 | GGAGCCTACTTGCCACAGTAT |
| | 3 | CAGCCGTACATTCTGCAGCAT |
| | 4 | GAGGCTGGAATGACACTCACT |
| | 5 | CTGCCAACCATATGGGAAGAT |
| | 6 | CCAGCAGTAATAACAACAGTA |
| Rsad2 | 1 | GGCAGCTGGCTTGGTATAAAT |
| | 2 | GCGGAAAGTATGTGTGGAGTA |
| | 3 | GACTCCTACCTTATCCTAGAT |
| | 4 | CAAGATCAACTCTGTCATTAA |
| | 5 | GGGTGAATACTTGGGCAAGCT |
| | 6 | GTGGATTTGATGAGAAGATGT |
| Sbno1 | 1 | CAGCCAACTATGTACTATCTA |
| | 2 | GCAATCTTAGTTAAAGAAGTA |
| | 3 | GTACACTGGCTGGATCAGTAT |
| | 4 | GCAGTTCCTGTCAAACAAGAG |
| | 5 | GCAACTCAAATTAGAAATTTA |
| | 6 | CATACTGGCGTGGCAATTGCA |
| Slfn5 | 1 | GCCAGAAAGGTCAAGAAACAA |
| | 2 | GCCCGATTCCTGGGAGATAAA |
| | 3 | CCAGCTACCTTCAGGATATAA |
| | 4 | CCCCGAATCCTATAACTTCAT |
| | 5 | CCCTGTTAAATGAAGAGTTAA |
| | 6 | CCCTGGAAGATTAGGCATCAT |

| | | |
|---|---|---|
| Prrx1 | 1 | CTGCTATGTTTAGAATCAAAT |
| | 2 | CGAGCCATGCTGGCCAATAAA |
| | 3 | GCACGTCGGGTGAACCTCACT |
| | 4 | GGTGAATGACTGGCCCACTCT |
| | 5 | CTCGTCCTGCTCCCAGACCAA |
| | 6 | CTCGACACCCTGCAGGCGAAA |
| Rbm7 | 1 | GCTCCTTGTGTGCTACTTTAA |
| | 2 | CACCCCTACCTAGCAGATAGA |
| | 3 | GGCCATACCTTTAACCAGTCT |
| | 4 | CGGCAAGCAGTGATGAACAGT |
| | 5 | GAGCCATGACTATGATAACAG |
| | 6 | GCCGAGAGGATTTCTACTATG |
| Rbms2 | 1 | CTGGGTAACCTGCCTCTTAAA |
| | 2 | CTGCTATGCATGGGGCTTACA |
| | 3 | CTGTCAGCCGTATGGCAAGAT |
| | 4 | CAATCTGCGCTAGCCCCGTAT |
| | 5 | CAGGCACAAATGGCAAAGCAA |
| | 6 | CAGAGAAGTGTGAAGCCATCA |
| Sbf1 | 1 | CTGCTCTGCTTCCAGTCGTTA |
| | 2 | GGCGGCAGTATGTTACTCGTA |
| | 3 | CCCCTAAGACTGTGGATGAGA |
| | 4 | CTTGCTGGACTCTGATTATGA |
| | 5 | GGCAGAGATCTCCTCAGTACT |
| | 6 | GTCAGCTCCATGCCACGTTAT |
| Slfn2 | 1 | CTTCCAGAGATCTGAGCTGTA |
| | 2 | GGTGGAGCAATCAAGGTTAAA |
| | 3 | CCAGATATGTTGAAGTTACAT |
| | 4 | GAAGCCTCTCTTTGTAAATGT |
| | 5 | GAAGGGGGATTAATAATCTGT |
| | 6 | CAGGACAGAATTTCAGTACGA |
| Slfn8 | 1 | GTCGAGGTCCTTCAGGACAAG |
| | 2 | CGCGTATCTCTGTGATTCTAA |
| | 3 | CAGATAAGTGTTATGATTTCT |
| | 4 | CAGTCTCTCCTATTAATTACC |
| | 5 | GCAACCACACTGCCTTTACTC |
| | 6 | CAGAGGCATCTTCCAAAGATC |

| | | |
|---|---|---|
| Slfn9 | 1 | CGCTCCAATCCCTAAATACCA |
| | 2 | CCCAGAACCATGAGTCAGAAA |
| | 3 | GAAAGGTGTCTACACACTGAA |
| | 4 | CCAAGCAACAGGAGGATAAGT |
| | 5 | CCCTTGATTTGCAGGCTTTCT |
| | 6 | GTTGACCCTGACTCTTTGAAA |
| Sp100 | 1 | CTCCCGTCTCTAACCTTCTAT |
| | 2 | CCCGAATGGGTCATCCTTAGA |
| | 3 | GGAGCACATGTGGTTAAACAA |
| | 4 | CTGAGAGAACTGATACAGAAA |
| | 5 | GGCATCTATGTGAGGAGTATA |
| | 6 | GCTTCCTGGATAAGGAGAAAT |
| Themis2 | 1 | CTGATTGAACTACCCAATGTA |
| | 2 | GACCCGGATATGGATGACCAT |
| | 3 | CAGCCGACGCTATAACCTAGT |
| | 4 | CCCAGATACTTCATGCTCTCT |
| | 5 | CTACGAGCTCTTTGGGAATGA |
| | 6 | GTGCTGAGTATGAAGCCCAAA |
| Zbp1 | 1 | GCTCCTAGACTTTAGATAGAT |
| | 2 | CCCAGGAGCTTCATTCAACAT |
| | 3 | CACGGGAATGTCATAGTAAGA |
| | 4 | GGCCAAGACATAGCTCATTCT |
| | 5 | CGGACAGACGTGGAAGATCTA |
| | 6 | CTAGCCTTGATGAAAGAATAT |
| Zcchc11 | 1 | CACGTTTAGATAGCTTATTTA |
| | 2 | CGGGGAAGTTTATCTTCATAT |
| | 3 | CGAGGAAACGTGTCCGAGTAA |
| | 4 | GGCCTTACGGTTTGCATCAAA |
| | 5 | CCACCAGTTATCCCAGTTCTA |
| | 6 | CAGAGATGTATTGTTGATAAT |
| Znfx1 | 1 | GCTCCTCGACTGTAGCAAGTA |
| | 2 | GAGCCTCGGATCGTCATTGTA |
| | 3 | GCGGAGGGCCTATATGAGTAT |
| | 4 | GCCTCTGGCATCGTGTACAAA |
| | 5 | CTTGGACCGCTTCATGAATGA |
| | 6 | CCTGCTCATCTGCTCACACAA |

| | | |
|---|---|---|
| Snrnp200 | 1 | CAGGCTGAGGTCTGGCCATTC |
| | 2 | CTGCCTACTATTACATAAACT |
| | 3 | GGGCACAGCTTACAGATAAGA |
| | 4 | GCGTGCAATCTTCGAAATTGT |
| | 5 | GGGCTATGCCTACCTATACAT |
| | 6 | GCACATACAGTCATCATTAAA |
| Srbd1 | 1 | GCCACTTTATTGAGGTGTTAT |
| | 2 | GGGGCTTGTGTAAACATAGAT |
| | 3 | GGAGAGAAAGTGGAGGTGAAA |
| | 4 | GGCTGATTCCTATCCGGTTCA |
| | 5 | GGGGTCTTTGTGGATATAGGA |
| | 6 | GGGACAGTTCTTACAGGCAAA |
| Trim56 | 1 | GAAGCAGAGCCACCTATAGGA |
| | 2 | CTGCGCCAGCTTCAGGATGCT |
| | 3 | CCTTCGAGAGGTAAACAAGGT |
| | 4 | AGAAGATGATGGGGTCTTCAT |
| | 5 | GGGCCCTGGGTTACATGGCTG |
| | 6 | GGAGCAGGTGAAAGAGAATCC |
| Zbtb7a | 1 | CATGGACTACTACCTGAAGTA |
| | 2 | GAAGGTCGCTTTCCAGTTTCT |
| | 3 | CCTGGAGTTCTTCCGCAGTAA |
| | 4 | CTGCAATGGCTTGGACTTCTA |
| | 5 | CAGCCAGTACTTCAAGAAGCT |
| | 6 | CCTACGAGTGTAACATCTGTA |
| Zfp143 | 1 | CAGGCTAATCTGAGCTACTCT |
| | 2 | GACGCCAGGGTTGGATGATTA |
| | 3 | GTCGCAATTGTAGCTCAAGAT |
| | 4 | GTCGGTCCTTTACCACATCAA |
| | 5 | GCCCTTTAAGTGTCCTATTGA |
| | 6 | GGACTCAACATGTCAACATAT |

# List of publications

**Karayel E**, Buerckstuemmer T, Bilban M, Duernberger G, Weitzer S, Martinez J, Superti-Furga G. TLR-independent DNA recognition pathway in murine macrophages: Ligand features and molecular signature, *Eur. J. Immunol.2009*. 39:1929-1936

Goncalves A, Buerckstuemmer T, Dixit E, Scheicher R, Górna MW, **Karayel E**, Sugar C, Stukalov A, Berg T, Kralovics R, Planyavsky M, Bennett KL, Colinge J, Superti-Furga G. Functional dissection of the TBK1 molecular network, PLoS ONE 2011 (*in press*)

# Curriculum vitae


| | |
|---|---|
| Name: | Evren Karayel |
| Date of Birth | 15 December 1979 |
| Nationality | Turkish |
| Address | Windmühlgasse 30/2<br>A-1060 Vienna / AUSTRIA |
| Email | ekarayel@cemm.oeaw.ac.at |


## Research – Work Experience

| | |
|---|---|
| 2007-present | **PhD Thesis,** CCHD PhD Program. Research Center for Molecular Medicine of the Austrian Academy of Sciences (CeMM) *Vienna, AUSTRIA*. Investigated viral DNA recognition in the cytoplasm and the subsequent immune response under the supervision of *Prof. Giulio Superti-Furga* |
| 04/2010 – 06/2010 | **Graduate Researcher,** University of Kyoto, *Kyoto, JAPAN*. Investigated the cellular defence proteins upon viral infection in the laboratory of *Prof. Takeshi Fujita* |
| 2005 – 2006 | **Diploma Thesis,** University of Natural Resources and Life Sciences, *Vienna, AUSTRIA* The Impact of the Interaction between IGF-II/Plasminogen and M6P/IGF2 Receptor on Skin Tumour Invasion" carried out under the supervision of *Prof. Lukas Mach* |
| 06/2004 - 08/2004 | **Undergraduate Researcher,** Institute of Food Research, *Norwich, UNITED KINGDOM*. Investigated the attachment of *Salmonella* Typhimurium to prepared vegetable tissues in the laboratory of *Prof. Tim Brocklehurst* |

## Education

| | |
|---|---|
| 2007-present | **Ph.D. Immunology,** Medical University of Vienna, Research Center for Molecular Medicine of the Austrian Academy Of Sciences (CeMM) |
| 1998 – 2006 | **M.Sc. Food and Biotechnology,** University of Natural Resources and Life Sciences*, Vienna, AUSTRIA* (combined with **B.Sc.**) |
| 1990 – 1998 | **Turkish Baccalaureate,** St. Georg's Austrian High School *Istanbul, TURKEY.* Concurrently, Austrian School Leaving Exam (Matura) passed with distinction |

## Professional Activities

| | |
|---|---|
| 2007 – 2010 | **Student Speaker** of "CCHD", PhD-Program of Medical University of Vienna |
| 2007 – 2010 | **Co-founder and co-organizer** of the international annual symposium "Bridging the Gap", |

## Scientific Meetings

| | |
|---|---|
| 2009 | Cellular and Cytokine Interactions in Health and Disease, *Lisbon, Portugal.* Poster presentation |
| 2009 | 5th YSA PhD Symposium, *Vienna, AUSTRIA.* Poster presentation |
| 2008 | TOLL2008, *Lisbon, Portugal*, Poster presentation |
| 2008 | Ubiquitin and Cellular Regulation FASEB Summer Research Conference, *Vermont, USA.* Poster presentation |
| 2008 | 4th YSA PhD Symposium, *Vienna, AUSTRIA.* Poster presentation |
| 2007 | 12th FEBS Immunology Summer School, *Hvar, Croatia*. |

## Publications

| | |
|---|---|
| 2011 | Goncalves A, Buerckstuemmer T, Dixit E, Scheicher R, Górna MW, **Karayel E**, Sugar C, Stukalov A, Berg T, Kralovics R, Planyavsky M, Bennett KL, Colinge J, Superti-Furga G. Functional dissection of the TBK1 molecular network, PLoS ONE 2011 (*in press*) |
| 2009 | **Karayel E**, Buerckstuemmer T, Bilban M, Duernberger G, Weitzer S, Martinez J, Superti-Furga G. TLR-independent DNA recognition pathway in murine macrophages: Ligand features and molecular signature, *Eur. J. Immunol.2009*. 39:1929-1936 |

## Vocational Training

| | |
|---|---|
| 2002 -2006 | **Laboratory Tutor,** "Molecular Biology", "Analytical and Physical Chemistry", "Thermodynamics", and "Energy, Momentum and Mass Transfer", University of Natural Resources and Life Sciences, *Vienna, AUSTRIA* |
| Summer 2003 | **Internship,** "Christian Doppler Laboratory" *Vienna, AUSTRIA.* Investigated the synthesis of heptoses which are common constituents of various bacterial lipopolysaccharides |
| Summer 2001 | **Internship**, "Efes Pilsen" Brewery, *Istanbul, TURKEY*. Assisted the computer controlled stages of brewing processes |
| Summer 2000 | **Internship,** "Bachmayer" Brewery, *Dorfen, GERMANY*. Actively involved in every step of brewing, from mashing to packaging |
| 02/2001 | Workshop on "Novel and Functional Food", University of Gent, *BELGIUM* |