Evolution-guided design of super restrictor antiviral proteins

Rossana Colón-Thillet

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

2019

Reading Committee:

Harmit S. Malik, Chair

Michael Emerman

Deborah Fuller

Program Authorized to Offer Degree:

Molecular and Cellular Biology

University of Washington

Abstract

Evolution-guided design of super restrictor antiviral proteins

Rossana Colón-Thillet

Chair of the Supervisory Committee:
Harmit S. Malik
Affiliate Associate Professor
Genome Sciences

Host-virus evolutionary arms-races are driven by antagonistic interactions and often manifest as recurrent amino acid changes (*i.e.,* positive selection) at their protein-protein interaction interfaces. Here, we investigated whether combinatorial mutagenesis of positions under positive selection in a host antiviral protein could enhance its restrictive properties. We tested ~800 variants of the human MxA protein, generated by combinatorial mutagenesis, for their ability to restrict Thogoto orthomyxovirus (THOV). We identified MxA 'super-restrictors' with increased binding to THOV NP target protein and 10-fold higher anti-THOV restriction relative to wild-type human MxA, the most potent naturally-occurring anti-THOV restrictor identified. However, MxA super-restrictors of THOV were impaired in their restriction of influenza A virus. Our findings thus reveal a breadth-versus-specificity tradeoff that constrains the adaptive landscape of antiviral proteins.

Table of Contents

List of figures

Acknowledgments

I would like to express my sincere gratitude to my advisors Dr. Michael Emerman and Dr. Harmit S. Malik for their continuous support, patience, motivation, and immense dedication to my training as a scientist. I will never forget our monthly meetings where we discussed ideas, results, and future directions. I left the room every time feeling fortunate to have such brilliant and enthusiastic mentors to guide me through my Ph.D. I have admired your scientific careers from the very beginning. However, today I understand that the reason behind your scientific success is an unwavering commitment to scientific integrity, and the personal and professional success of your trainees. I strive every day to emulate the example you have given me. Thank you.

I would also like to thank my thesis committee: Dr. Deborah Fuller, Dr. Jesse Bloom, and Dr. Jason Smith, for their insightful comments and encouragement. Your challenging questions undoubtedly improved my research and led me to push the limits of my knowledge. I thank my fellow labmates for stimulating discussions, motivation, and their friendship. I could not imagine a better environment to grow as a scientist. I leave my Ph.D. with friendships that hopefully will last a lifetime.

I would like to thank my parents; whose love and guidance are with me in whatever I pursue. They have shown me the power of perseverance and resilience. I am especially grateful to my sister for being the ultimate role model, and to my niece and nephews for cheering me up at every step of the way. Finally, I would like to thank my loving and supportive partner, Josef. You came into my life when I needed it the most. For this, and all you do, I will be forever grateful.

Chapter 1: Introduction

Parts of this Chapter were modified with permission from (6, 10, 17), Elsevier license numbers 4503790176086, 4503790648850, 4503791346007

Innate immunity against viral infection

Viruses are obligate parasites that invade host cells and utilize their cellular machinery to reproduce. Recent studies have revealed that vertebrate-associated viruses have existed for the entire evolutionary history of vertebrates, extending these host-virus relationships over geological time[1]. Therefore, viruses represent an ancient and ongoing selective pressure that have shaped the evolution of innate and adaptive immune systems in vertebrates[2–5]. The genetic imprints of viral infections act as a record of pathogenic challenges and host countermeasures to overcome them. Thus, the evolutionary history of immunity genes in vertebrate species can help us understand the molecular rules and evolutionary dynamics of host-virus interactions[6].

To counteract the massive challenge imposed by the pathogens, vertebrates have evolved innate and adaptive immune systems. Although there is mounting evidence of cross-talk between the two[7,8], the adaptive immune system plays a critical role in controlling viral infections and developing memory for future infections; whereas the innate immune system often serves as the first line of defense against viral infection. The innate immune system deploys cell-intrinsic effectors or restriction factors, which are germ-line encoded proteins with antiviral activity against a wide range of viral pathogens. These antiviral defenses can be deployed rapidly, often within hours of a pathogen attack, and are expressed in all cell types. The mobilization of most restriction factors requires the recognition of viral infection and a subsequent signaling cascade that leads to

the release of signaling molecules to mount an antiviral state. Other restriction factors are constitutively expressed, but their antiviral activities can be upregulated in response to infection (e.g. APOBEC family, SERINC3,5, Fv1).

The antiviral state mounted upon sensing of a viral infection is mediated by the production of interferon (IFN) and other cytokines[8]. The IFN response is induced through the recognition of pathogen associated molecular patterns (PAMPs) by germline-encoded pathogen recognition receptors (PRRs), such as Toll-like receptors (TLRs), retinoic acid-inducible gene I (RIG-I)-like

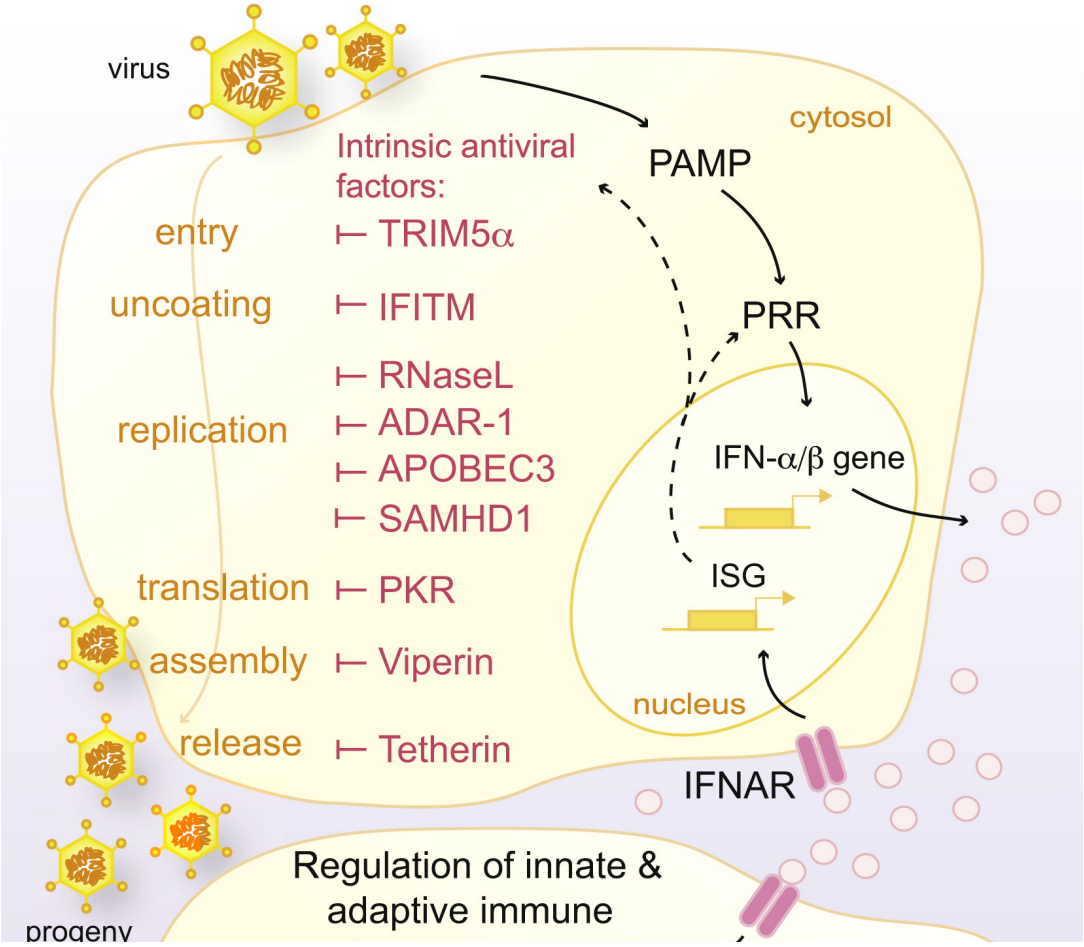


Figure 1. The cell-intrinsic response to viral infection.

A schematic representation of the IFN-mediated "antiviral state" in a virus infected cell. Cells express many intrinsic restriction factors that can block the replication of viruses at different stages of the replication cycle. The expression of these restriction factors is often regulated by the regulatory molecule IFN. IFN production is triggered by the sensing of PAMPs by PRRs. IFNs signal via the IFN receptor (IFNAR) to regulate the expression of hundreds of restriction factors.

receptors (RLRs), nucleotide-binding oligomerization domain (NOD)-like receptors (NLRs), C-type lectin receptors (CLRs), AIM-2 like receptors, OAS proteins and cGAS[9–11]. PRR are structurally diverse and recognize PAMPs through the use of distinct ligand-recognition domains. The structural diversity of PRRs is maximized by their tissue-specific expression and their localization in different cellular compartments. The differences in expression patterns reflect the two general modes of recognition by the innate immune system: cell intrinsic and cell-extrinsic. Cell-intrinsic recognition is mediated by intracellular cytosolic sensors that operate in infected cells. For example, RLRs are localized in the cytosol and act as sensors of viral replication by detecting viral RNA or DNA. Cell-extrinsic recognition is independent of viral replication and is mediated by PRRs expressed in cells specialized in pathogen detection. For instance, TLR7 and TLR9 in plasmacytoid DCs (pDCs), a subset of dendritic cells, participate in the recognition of viral components (i.e. genomic DNA and RNA) in a replication-independent way. Nucleic acids are delivered to the endosomal compartment after endocytosis, where these TLRs recognize the released nucleic acids to trigger an antiviral response.

PRRs can recognize viral structures or by-products of viral replication, such as nucleic acids or viral proteins. In the case of nucleic acids, detection requires that PRRs distinguish between self and non-self-nucleic acids to combat viral infection while preserving the integrity of the cell by avoiding a self-triggered response. For example, a major PAMP from RNA viruses is single-stranded RNA, which is also native and abundant in the host cell. RIG-I robustly discriminates between viral ssRNA from self ssRNA by recognizing a triphosphate group at the 5’ end, a structure that is lacking in host mRNAs[12]. PRRs signal through different pathways depending on their tissue-specificity and subcellular compartmentalization, but ultimately their signaling results in the activation of interferon regulatory factors (IRFs), which induce the

expression and secretion of type I IFNs in infected cells[11]. Type I IFNs have three major functions. First, they induce cell-intrinsic antiviral states in infected and neighboring cells that limits the spread of viral infection (**Figure 1**). Second, they promote antigen presentation in dendritic cells and natural killer cell functions such as cytolytic activity. Third, they activate the adaptive immune system, thus promoting the development of high-affinity antigen-specific T and B cell responses and immunological memory[7,13].

The most well-defined type I IFNs are IFNα and IFNβ. Most cell types produce IFNβ, while IFNα is produced by hematopoietic cells[14]. The receptor for IFNα and IFNβ is the IFNα receptor (IFNAR), which is widely expressed in most cell types. IFNAR binding leads to the activation of the receptor-associated protein tyrosine kinases Janus kinase 1 (JAK1) and tyrosine kinase 2 (TYK2), which phosphorylate the latent cytoplasmic transcription factors signal transducer and activator of transcription 1 (STAT1) and STAT2[15,16]. Tyrosine-phosphorylated STAT1 and STAT2 dimerize and translocate to the nucleus, where they assemble with IFN-regulatory factor 9 (IRF9) to form a complex called IFN-stimulated gene factor 3 (ISGF3). ISGF3 binds to its cognate DNA sequences, which are known as IFN-stimulated response elements (ISREs), thereby directly activating the transcription of ISGs[12,17]. ISG-encoded proteins, or restriction factors, impede pathogens by several mechanisms, including the inhibition of viral transcription, translation and replication, the degradation of viral nucleic acids and the alteration of cellular lipid metabolism[18,19].

Viral target recognition by restriction factors

Restriction factors show a wide structural and functional diversity and target almost every step of the viral replication cycle[18]. Analogous to the PRR recognition of unique viral structures, restriction factors recognize viral proteins or byproducts of viral replication (here collectively

described as viral targets). Restriction factor recognition of viral targets often requires direct binding, which leads to inhibition of viral replication. In some cases, a specific restriction factor can singularly provide protection against cross-species viral transmission, such as rhesus macaque TRIM5 (rhTRIM5) block of SIV from sooty mangabeys (SIVsm). In this study, rhTRIM5 had a potent effect during the early stages of cross-species transmission by significantly reducing viral titters and failure to induce persistent infection[20]. These suggests that adaptation in the SIVsm capsid was required for a successful jump and emergence of SIVmac from SIVsm.

Viral target recognition specificity divides restriction factors into two categories: broadly-acting or specific[21]. Broadly-acting restriction factors often target conserved viral components, such as viral genomes or membranes. An example of a membrane-targeting restriction factor is the IFITM family of proteins, which impair virion infectivity of a diverse group of viruses including retro-, orthomyxo-, flavi-, filo-, corona-, rhabdo-, bunya-, reoviruses by altering the lipid composition of the viral membrane, inhibiting membrane fusion or modifying membrane fluidity [22,23]. Another broadly acting restriction factor is the antiviral protein PKR. PKR inhibits a wide range of viruses including herpes-, orthomyxo-, retro-, flavi-, reo-, adeno-, poxviruses by recognizing viral dsRNA and inhibiting mRNA translation by eIF2a phosphorylation[24,25]. Other restriction factors can be highly specific and only target a family of viruses through recognition of a distinct family feature. An example is the restriction factor TRIM5α and the related TRIM-CypA protein that target incoming retroviral capsids and block viral replication by inducing premature uncoating, which leads to defects in viral DNA synthesis[26,27]. TRIM5α and TRIM-CypA use capsid-binding C-terminal B30.2 or cyclophilin domains, respectively, to specifically recognize the capsid protein lattice or capsid monomers of retroviruses upon cell entry and to restrict retroviral replication before reverse transcription can occur. Therefore, restriction factors tend to

be either broadly-acting factors that exploit general features of viral replication or effectors with strong pathogen preference via specific recognition of viral proteins. An atypical restriction factor that combines both target specificity and broad activity is the antiviral protein MxA (myxovirus resistance protein A). This thesis focuses on a novel methodology and dissects the molecular basis for MxA breadth and specificity for Orthomyxoviruses.

The discovery of Mx antiviral genes

The first Mx protein family member to be identified and molecularly cloned was mouse Mx1 (Myxovirus protein 1). This discovery was spurred by a chance observation made in 1961 by Jean Lindenmann, the co-discoverer of IFN with Alick Isaacs in 1957. He found that mice of a particular inbred strain (A2G) survived an otherwise lethal dose of FLUAV (Influenza A virus)[28]. Crossing of A2G mice with susceptible mice yielded F1, F2 and backcrossed offspring with percentages of susceptibility that were consistent with the presence of a single dominant allele that conferred resistance in a heritable, Mendelian fashion. At the time, it was widely accepted that IFN led to a nonselective antiviral activity. Therefore, how a single gene could mediate potent antiviral activity against a specific virus was enigmatic. Then, 22 years after the initial observation, the discovery of anti-IFN neutralizing antibodies led to a paradigm shift. Treatment of laboratory mice with the anti-IFN antibody led to a partial brake of innate resistance against mouse hepatitis virus. Surprisingly, treatment of laboratory mice with the anti-IFN antibody did not increase influenza susceptibility as it did for mouse hepatitis virus. This experiment was repeated using A2G ($Mx1^{+/+}$) mice. Treatment with anti-IFN rendered the mice as susceptible to influenza infection as genetically susceptible ones. This indicated that A2G mice encoded for a factor that was IFN-inducible and selectively inhibited influenza infection, further supporting the idea that IFN

treatment could affect different viruses distinctively[29]. Subsequent experiments with mouse-derived macrophages *in vitro* showed that IFN treatment of A2G macrophages were protected from influenza much better than macrophages from susceptible mice, while the degree of protection against other viruses was independent of the Mx1 genotype[30]. This observation eventually led to the understanding that the resistance phenotype is controlled by a single gene on chromosome 16 which is functional in wild mice but crippled in most mouse laboratory strains by large deletions or nonsense mutations[31]. The mapping to chromosome 16 was achieved by mouse-hamster somatic cell hybrids analyzed with a probe prepared from a cDNA encoding murine Mx proteins[32] .

The potent, specific and autonomous effect of Mx1 was first seen in experiments in which transfection of recombinant Mx1 into Mx1-negative mouse cells allowed for constitutive expression of the protein without IFN-induction. Infection of Mx1-transfected cells with influenza or a virus control (i.e. vesicular stomatitis virus) followed by staining of newly formed proteins or Mx1 showed an unambiguous result; Mx1-transfected cells had no expression of new viral proteins in influenza infection, while the virus could replicate normally in control cells. It is now understood that Mx1 is the main IFN-induced intracellular restriction factor against influenza and influenza-like viruses in mice[33].

The generation of an Mx1 specific monoclonal antibody (2C12) in mice led to the identification of Mx1-like proteins in IFN-treated human PBMCs[34]. This same antibody was used to isolate human MxA by immunoprecipitation[35]. Subsequently, two classes of human Mx-related cDNAs were identified that corresponded to two distinct human Mx genes that were designated *MxA* and *MxB* (now referred as *Mx1* and *Mx2*, respectively). MxA, but not MxB, was characterized as a restriction factor against influenza virus. MxB is now a known restriction factor of retro- and

herpesviruses[36–38]. Recent studies using transgenic mice carrying interferon-regulated human MxA showed that these mice are resistant to avian but not to human influenza A viruses[39]. This finding underscores the importance of MxA as a cross-species transmission barrier from the influenza avian reservoir to humans. Human influenza A viruses have adapted to overcome MxA restriction. In fact, MxA-resistance conferring mutations are present at a high frequency (>98%) in classical seasonal human isolates representing H1N1, H1N2, H2N2, and H3N2 subtypes[40]. The high frequency of MxA 'escape' variants in seasonal influenza strains attests to the continuous selection pressure for increased Mx resistance in seasonal influenza viruses and highlights the role of Mx proteins as a potent cross-species barrier for avian influenza strains.

Evolutionary history of Mx proteins

The importance of Mx proteins as antiviral factors has been experimentally demonstrated extensively. The evolutionary history of Mx protein in mammals reflects dynamic, lineage-specific changes in copy number and gene conversion, characteristic of many antiviral gene families[6,41]. Phylogenetic analyses of mammalian Mx paralogs from at least one representative mammalian order revealed shared synteny of the Mx locus through terrestrial vertebrates. It is thought that MxA and MxB genes diverged at or just prior the origin of the eutherian mammal lineage. A lack of Mx proteins in the marsupial genomes sequenced to date suggest that Mx gene (s) were lost from the common marsupial ancestor. Therefore, most eutherian mammal genomes encode two Mx genes: MX1 and MX2. In humans and other primates, the gene products derived from MX1 and MX2 are named MxA and MxB, respectively. Interestingly, the two rodent Mx genes group together in a branch that is most closely related to Mx1 of other mammals. This suggests that Mx2 in rodents was derived from an ancestral Mx1-like gene in a gene conversion event and that the ancestral Mx2 was lost during evolution.

Phylogenetic analyses suggest that the lineage-specific loss of Mx proteins is not unusual. Both MxA and MxB proteins are lost in toothed whales (including orcas and dolphins). This lack of functional Mx genes is thought to have detrimental impact in the immune response against viral infection in toothed whales[42], evidenced by the high mortality rates due to CMV outbreaks in dolphins[43]. The loss of Mx proteins in whales could be attributed to the extinction of Mx-targeted pathogens, relaxing the selective pressure to maintain Mx activity. A similar loss of constraint in armadillo, felids and squirrels could explain the loss or pseudogenization of the MxB gene in these lineages[41]. The dynamism of the evolutionary history of the Mx gene family in mammals highlights the strong diversifying selection that shapes the functional repertoires and specificities of antiviral genes.

Evolutionary studies that aim to dissect the relationship between Mx proteins of different organisms have provided clues about the antiviral functional diversity and the emergence of specialized functions in different subclasses of Mx proteins[44,45]. An example is the length variability of the N-terminal region that precedes the GTPase domain in Mx2 proteins in mammals. This N-terminal segment in human MxB encodes for a functional nuclear localization signal (NLS), which is absent in human MxA. Previous studies have shown that human MxA and both murine Mx1 and Mx2 proteins have broad and potent activity against a diverse range of RNA and DNA viruses. Initial studies suggested that, in contrast to MxA, MxB had no antiviral activity but instead regulates the nucleo-cytoplasmic transport of cellular factors[46–48]. It is now understood that MxB is a potent restriction factor of lentiviruses and herpesviruses[49]. The N-terminal NLS in MxB is necessary and sufficient for its antiviral activity. An engineered chimeric MxB/MxA proteins carrying the first 85 N-terminal amino acids of MxB was sufficient to confer anti-herpesvirus activity to human MxA. The chimeric protein also gained anti-HIV-1 activity. Thus, the highly

divergent N-terminal segment acts as a module for antiviral activity against herpes- and lentiviruses[50]. As we will discuss in the following chapters, MxA follows a similar pattern. The C-terminal L4 loop of the protein acts a modular target recognition domain for Orthomyxoviruses. The breadth of antiviral activity in Mx proteins could be then attributed to the physical separation of antiviral specificity domains that act as modules to recognize distinct viral targets. The residues in MxB that act as molecular determinants for anti-lentivirus activity do not overlap with those under positive selection. Therefore, lentiviruses did not drive the rapid evolution observed in the N-terminal tail of MxB. It is possible that this selection is the result of MxB having a long-standing and important role in the interferon response to viral infection against a broader range of pathogens, including herpesviruses. [41]

Subcellular localization of Mx proteins and antiviral profiles

Innovation during the evolutionary history of Mx proteins can also be appreciated in the different subcellular localization of Mx orthologs and paralogs[51]. Different Mx proteins associate with distinct intracellular compartments in the cytoplasm and the nucleus. However, the localization of Mx proteins does not always match with the site of replication of the viruses it restricts. For instance, human MxA is a cytoplasmic protein that restricts the replication and secondary transcription of FLUAV, a virus that replicates in the nucleus. Nonetheless, conferring nuclear localization to human MxA via an NLS increases its anti-FLUAV restriction activity. In general, human MxA seems to restrict viruses irrespectively of their replication site.

MxA retains the lipid binding capacity characteristic of dynamin proteins. Purified MxA protein oligomerizes in ring-like structures around negatively charged liposomes and transforms them into tubes. The role of membrane binding for MxA action is unclear, but it is possible that

this close association to membranes puts MxA in close proximity to viruses that depend on intracellular membranes to replicate such as bunyaviruses, positive-strand RNA viruses, poxviruses, and African swine fever virus (ASFV). Murine Mx proteins follow a more obvious pattern, as nuclear rodent Mx1 proteins inhibit viruses that have a nuclear replication phase (influenza and influenza-like viruses), but not viruses that replicate exclusively in the cytoplasm, whereas cytoplasmic rodent Mx2 proteins, that exhibit about 70% amino acid sequence identity to human MxA, inhibit viruses that replicate in the cytoplasm, such as vesicular stomatitis virus (VSV) and bunyaviruses[52–54].

MxA mechanism of action

The human *Mx1* gene encodes for a 662-amino acid (76kDA) protein that shares many properties with the dynamin superfamily of large GTPases, a group of mechanoenzymes that promote membrane fission through a contractile motion or 'power-stroke'[55,56]. MxA consists an N-terminal GTPase domain that binds and hydrolyzes GTP, a middle domain that mediates self-assembly and a carboxy-terminal GTPase effector domain[57]. Analogous to dynamins, MxA has the ability to self-assemble into highly ordered oligomers and form ring-like structures around liposomes, inducing liposome tubulation[58].

This first observations of MxA ring-like structures were obtained from electron microscopy studies of purified full-length MxA[47]. Since then, these structures have been further characterized by size exclusion chromatography and sedimentation assays[59]. At low protein concentrations and physiological salt concentrations, Mx proteins form tetramers in solution. At higher protein concentrations (>1.5 mg/ml), these tetramers oligomerize further into large filaments and rings. Addition of the non-hydrolysable GTP analog, GTPγS, changes the ring structure, leading to the

appearance of spirals and stacks of rings[60]. In the presence of liposomes, MxA assembles into more uniform rings and induces liposome tubulation[61]. It is unclear if the MxA association and rearrangement of lipids plays a role in antiviral activity. The resolution of the MxA crystal structure revealed the inter- and intra- molecular interactions that mediate ring formation[62–64]. MxA oligomerization is mediated by three different interfaces (stalk-stalk, BSE-stalk, GTPase-GTPase) and the loop L4 region (**Figure 2**). These interfaces generate a crisscross interaction pattern between the stalk domains of different MxA molecules that results in ring formation. In these rings, the stalk domain is oriented inwards, while the GTPase domains are located in the outside of the ring. Furthermore, contacts between the BSE and the stalk of the neighboring parallel MxA monomer also mediate oligomer formation (**Figure 2** see orientation of oligomer rungs). This BSE-stalk interaction is crucial for the transmission of conformational changes triggered by GTPase activity from neighboring MxA molecules, and is required for antiviral activity[62,65].

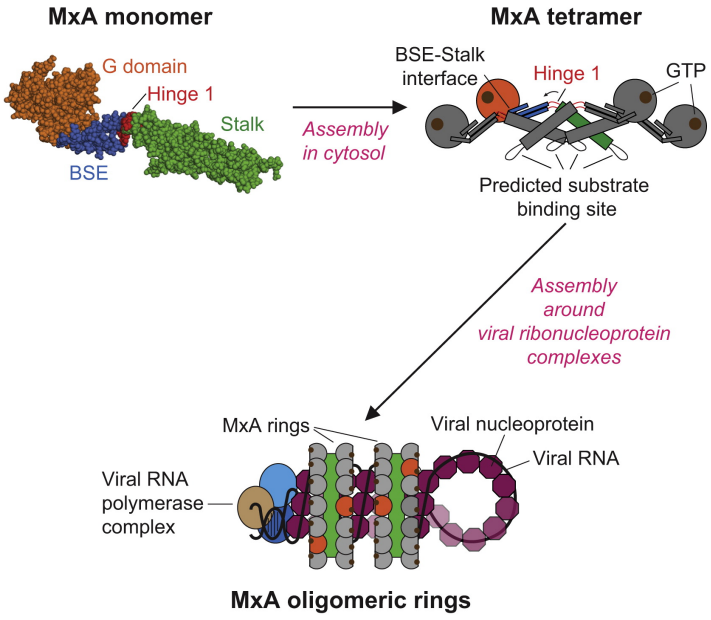


Figure 2. Inter- and intra- molecular interactions for ring formation.

(left) MxA monomer with domains colored. (right) Assembly of MxA monomers into tetramers is mediated by inter molecule stalk-stalk interactions, BSE-stalk interactions and interactions in the L4 loops (not shown) in the C-terminus. (bottom) Proposed assembly based on structural modeling and biochemical analysis around vRNPs. GTPase-GTPase interactions between rungs drive the mechanical force to constrict the target, a mechanism analogous to membrane scission by dynamins.

The details of how GTPase activity, oligomerization and lipid rearrangement contribute to antiviral activity have remained elusive. MxA appears to be a fully functional contractile machine like dynamin; its power stroke represents the mechanism through which stored chemical energy is used to perform mechanical work[66]. MxA has been proposed to assemble around viral nucleoproteins[29]. Therefore, it is possible that the contractile activity of MxA leads to mechanical disruption or entrapment of the target. However, it is also possible that Mx proteins could exert antiviral activity by perturbing viral membrane remodeling or viral trafficking. Thus, although the structural analyses nicely elucidate the nature of the MxA antiviral complex, they still have left unanswered the important question of how this complex might engage with its numerous viral targets.

Mx proteins combine breadth and specificity

MxA targets a wide range of DNA and RNA viruses including orthomyxo-, bunya-, rhabdo-, paramyxo-, picorna-, toga-, reo-, hepadna-, asfar-, and poxvirus[67]. MxA expression is induced by type I (α/β) and type III (λ) interferons, but not type II (γ) interferons or other cytokines[68]. The breadth MxA activity would suggest that MxA recognizes a ubiquitous target in all the viruses it targets. However, a number of studies that aimed to identify the viral targets of MxA suggested that MxA interacts with highly divergent proteins across a diversity of viral families. For instance, multiple lines of evidence suggest that MxA targets the viral ribonucleoprotein complexes (vRNPs) through direct interaction with the nucleoprotein (NP) that

is the main component of the nucleocapsid in othromyxoviruses, a family of viruses that includes influenza (IAV) and Thogoto virus (THOV)[40,69,70]. Similarly, MxA inhibits the replication of bunyaviruses by sequestering the viral nucleocapsid (N) in perinuclear compartments[54]. It has been suggested that MxA specificity relies on subcellular compartmentalization. Members of the orthomyxovirus family replicate in the nucleus, while bunyaviruses replicate in the cytoplasm. However, in both viruses, the nucleoprotein supports the formation of vRNPs. In addition, both bunyavirus and orthomyxovirus nucleoprotein, support roles beyond RNP structure in replication and transcription. Therefore, one model was that MxA might target proteins with similar functions in distinct viral life cycles. However, MxA antiviral activity extends to DNA viruses, which makes the existence of a common viral target difficult. MxA inhibits the hepadnavirus hepatitis B virus (HBV) via the abundantly-expressed hepatitis B core antigen protein (HBcAg)[71]. MxA has also been reported to restrict large double stranded DNA viruses including the orthopoxvirus monkeypox and the Asfarvirus African swine fever virus (ASFV); for these viruses MxA targets are unknown[72].

The paradox of MxA broadly-acting yet specific antiviral activity highlights the complexity of MxA viral target recognition. A combination of structural and biochemical studies had revealed that MxA forms higher order ring-like oligomers that act as contractile machines fueled by GTP hydrolysis and inter-ring interactions. GTPase activity and oligomer formation are required for the antiviral activity against most restricted viruses. Despite this knowledge, for decades, mutagenesis and deletion efforts had failed to reveal the viral target recognition domain(s) of the antiviral protein MxA because disruption of any one of the domains of MxA eliminated viral restriction[21]. This highlights the need for alternative approaches to dissect host-virus molecular dynamics.

MxA susceptibility determinants in Orthomyxoviruses

MxA antiviral activity has been most intensively studied against orthomyxoviruses, which are a family of RNA viruses that include seven genera: Influenza virus A, B, C, D, Isavirus, Thogotovirus, and Quaranjavirus. The virion contains 6-8 segments of linear negative-sense single-stranded RNA[73]. In contrast to other RNA viruses, orthomyxoviruses replicate in the nucleus. The reference virus for the antiviral activity of MxA is influenza virus. This stems from the fact that a resistance phenotype led to the discovery of Mx proteins and that influenza is an important seasonal human pathogen[29]. Occasionally, influenza strains from the avian reservoir infect humans causing outbreaks with a high mortality rate[74]. MxA has also been described to inhibit the replication of other members of the Orthomyxovirus family, including Thogotovirus (THOV), a tick-borne virus that infects rats and domestic animals in a natural setting. In livestock, THOV causes severe disease, including febrile illness and abortions in sheep. In humans, THOV is known to cause febrile illness and encephalitis (THOV is classified as a Biosafety level 3 virus in the US). THOV shows a wide geographical range and antibodies against THOV have been identified in humans living in parts of Europe, Asia, and Africa[75].

The basis for MxA restriction of Orthomyxoviruses is the interaction with vRNPs through contacts with NP. In THOV, the result from the MxA-NP interaction leads to a block of vRNP nuclear import, a step that is essential for viral replication. In contrast, influenza vRNPs translocate into the nucleus where primary transcription takes place. Influenza restriction by MxA occurs at an ill-defined step after viral primary transcription but before viral protein expression. These different mechanisms could be due to specific modes of interactions with NP. The genetic determinants for MxA susceptibility for both THOV and influenza have been mapped to NP. The most compelling evidence for this hypothesis comes from detailed mapping and characterization

of the MxA-resistance mutations present in pandemic strains of influenza A/BrevigMission/1/1918 (1918) and A/Hamburg/4/2009 (pH1N1)[40]. Although the resistance-associated amino acids differ in the two strains, both form a surface-exposed cluster in NP. Swapping of the resistant-associated amino acid cluster into the MxA-sensitive avian influenza virus A/Thailand/1(KAN-1)/04 (H5N1) resulted in a gain of MxA resistance. Indeed, the acquisition of this cluster is required for adaptation to human cells as observed in the avian influenza A/Shanghai/1/2013 (H7N9) strain that emerged in 2013 causing more than 300 hundred cases with often fatal outcomes[76]. In human-adapted strains, the MxA-resistance conferring mutations are highly conserved underscoring the role of MxA as a potent cross-species barrier from avian zoonoses[40]. A recent study comprehensively mapped the full set of sites in NP that affect MxA sensitivity through a deep mutational scanning approach. This systematic approach revealed that most sites with the largest effect on MxA resistance already encode for amino acids that confer resistance, and that these amino acids are conserved across human and avian influenza strains[77]. This gain of function also has fitness consequences; it is coupled to a loss in replication efficiency due to failure to traffic vRNPs into the nuclear compartment. These findings highlight the potent barrier for cross-species transmission that MxA imposes since even MxA 'escape' variants incur a fitness loss[78].

The physical interaction of MxA with THOV NP has been demonstrated by *in vitro* co-sedimentation assays, electron microscopy and co-immunoprecipitation experiments. However, there is little evidence for a physical interaction between IAV NP and MxA. Certainly, the fact that mutations in surface exposed residues are responsible for MxA resistance or sensitivity suggests that a physical interaction takes place. A recent study demonstrated that dimeric forms of MxA interact with ectopically expressed NP from resistant or sensitive influenza strains, albeit to a lesser extent to the latter, by co-immunoprecipitation experiments[79]. This finding contends the

expectation that MxA resistant strains will escape MxA interaction. In co-immunoprecipitation experiments of IAV-infected cells, the interaction with NP is undetectable regardless of oligomeric state.

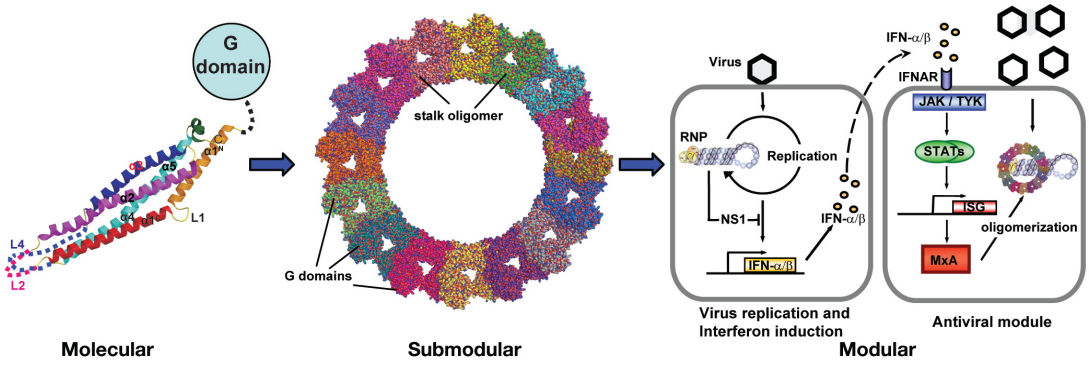


Figure 3. Functional organization of the antiviral module containing MxA.

(left) Monomeric MA, crystal structure of the MxA stalk. (middle) Oligomeric MxA ring assembled via stalk interaction between MxA monomers combined with the structure of the globular G domain of dynamin. (right) Proposed mechanism of action of the antiviral module. Viral sensing leads to activation of IFN expression. The viral protein NS1 suppresses the immune response by inhibiting the expression of IFN. IFN sensing through IFNAR and the JAK/STAT pathways in neighboring cells leads to the expression of interferon stimulated genes (ISGs), including MxA. Upon infection, cytoplasmic MxA might recognize the incoming viral RNP structures, and start to self-assemble into rings resulting in a stable oligomeric complex that blocks viral RNP function.

Host-virus arms races

Consistent with their important role in viral defense and constant need for adaptation, genes encoding for restriction factors (i.e. innate immunity genes) are some of the most rapidly evolving genes in mammals[80–83]. The signatures of rapid evolution can often be detected in gene regions corresponding to physical interaction interfaces between host and viral proteins. The recurrent selection of amino acid changes that confer a fitness advantage leads to a much greater likelihood that a nucleotide substitution that changes an amino acid (called a nonsynonymous substitution) will be selected for than will a silent (or synonymous) substitution. Thus, we can identify

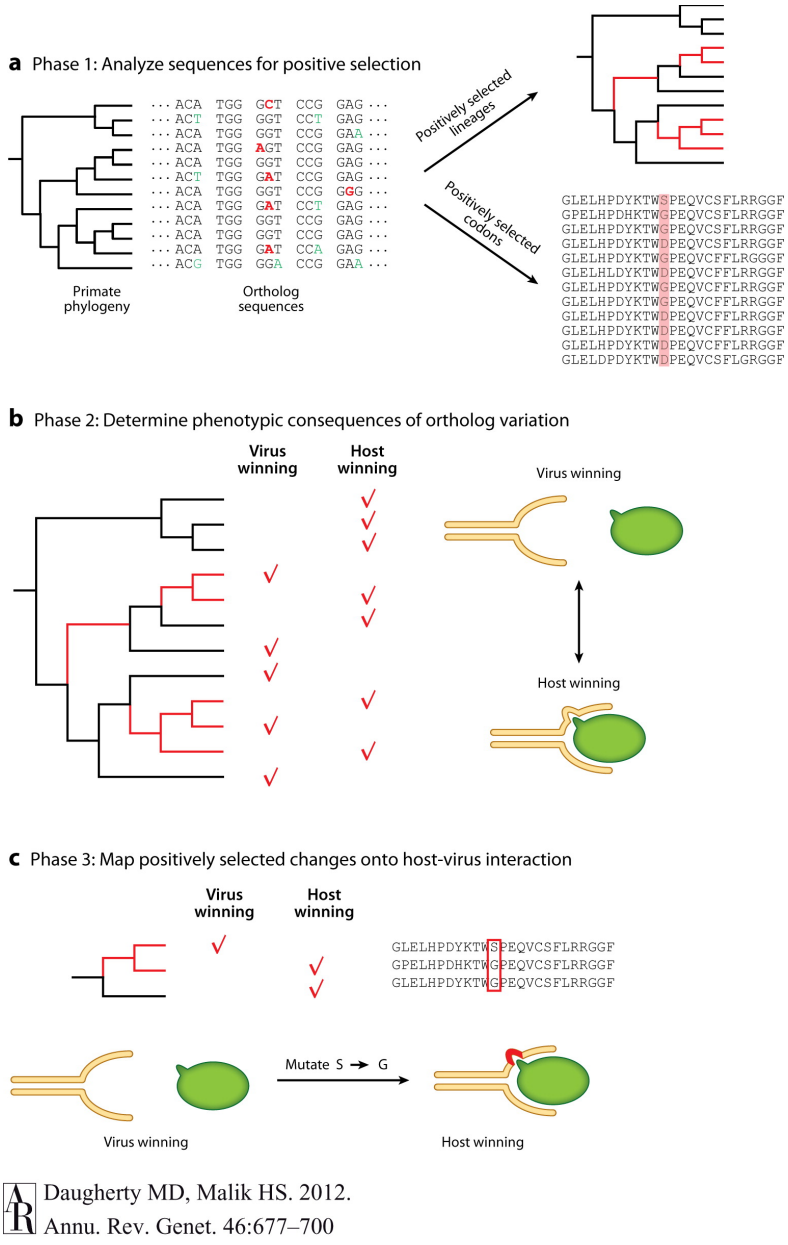
evolutionary signatures of rapid evolution, or positive selection, using maximum likelihood statistical methods that identify genes with an accelerated rate of nonsynonymous (dN) over synonymous (dS) substitutions[84,85]. A gene, or individual codon, is said to be evolving under positive selection when this ratio, dN/dS, statistically exceeds 1. Signatures of rapid evolution can be used to identify antiviral specificity domains and facilitate the functional dissection of host-virus interactions (**Figure 4**).

The basis for the genetic conflict is a protein-protein interaction in which binding affinity is constantly under selection. Restriction factors often directly bind to viral targets creating evolutionary dynamic host-pathogen interfaces[5]. We termed this an 'evolutionary arms race' in which both sides of the conflict are evolving rapidly to maintain the status quo. An example of a well-characterized host-virus arms race is SAMHD1 and lentiviral Vpr/Vpx proteins. SAMHD1 functions as a deoxynucleotide triphosphohydrolase, reducing cytoplasmic dNTP concentrations and thereby inhibiting reverse transcription of retroviruses. The retroviral accessory protein Vpx or Vpr antagonizes this host antiviral effect by binding and recruiting SAMHD1 to the Cullin4/DCAF1 ubiquitin ligase complex, targeting SAMHD1 for proteasomal degradation. Single changes at positively selected sites in SAMHD1 result in a dramatic switch between resistance and susceptibility to Vpx-mediated degradation. The episodes of positive selection in SAMHD1 coincide with the acquisition of SAMHD1-degrading capabilities in lentiviruses. This highlights how the functional consequences of positive selection shaped the evolution of lentiviruses. Phylogenetic tracing of the origin of the host-virus arms race between SAMHD1 and Vpr revealed that the ability to lentiviruses to degrade SAMHD1 was a newly evolved function of the Vpr gene, which arose in addition of Vpr's existing function to cause cell-cycle arrest. In other viruses, the Vpr gene has been duplicated and gave rise to Vpx, which has subfunctionalized to

antagonize SAMHD1, whereas the cell-cycle arrest function is performed by Vpr. The SAMHD1-Vpx/Vpr host-virus arms race exemplifies the power of evolution-guided approaches to identify critical residues that affect molecular recognition on both sides of the genetic conflict.

Evolution-guided discovery of antiviral specificity domains

Evolution guided-functional analyses begin by identifying a known or predicted virus-host conflict. The first phase in this methodology is to analyze sequences for positive selection (**Figure 4a**). This is achieved by assembling orthologous sequences across a phylogenetic range of species. This range of species divergence needs to be large enough for enough statistical power and avoid stochasticity-driven false positives, but not too large that dS will be underestimated as a consequence of saturation[6]. In primates, it has been determined that a 20-species data set composed of 8 hominoids, 8 Old World monkeys, and 4 New World monkeys allows for an adequate characterization of positive selection[86]. The second step is to determine the phenotypic consequences of ortholog variation by evaluating the antiviral activity of the restriction factor with a viral infectivity assay or replication competent virus (**Figure 4b**). This step will identify active and inactive orthologs against the virus being tested. The third phase involves the characterization and validation of the host-virus interface (**Figure 4c**). This can be achieved by mapping the differences in positively selected sites between active and inactive orthologs against the pathogen being assayed. The expectation is that by swapping the amino acid from active variants into inactive ones the phenotype of restriction activity will be transferred.



Daugherty MD, Malik HS. 2012.
Annu. Rev. Genet. 46:677–700

Figure 4. Evolution-guided functional analysis of host-virus arms races.

(a) Phase 1: Analysis of sequences for positive selection. Ortholog DNA sequences from the primate phylogeny are aligned and analyzed for dN/dS rates using maximum likelihood analysis. This methodology results in the identification of lineage and codon-specific positive selection. (b) Phase 3: Ortholog sequences are tested for their ability to restrict a virus that is known or predicted to be in a genetic conflict with the gene product in question. (c) Phase 3: The potential interfaces are mapped by swapping positive selected sites that differ from active to inactive orthologs. In this example, the active (winning) version of the host protein contains a G, whereas the inactive (losing) contains a S. Validation of the interface is achieved by swapping the S to a G in the inactive variant and evaluating for a gain of function.

We can utilize signatures of rapid evolution to make predictions about which amino acids in antiviral proteins that mediate the binding to viral targets. This evolution-guided approach has revealed that rapidly evolving sites in antiviral proteins often dictate the outcome of host-pathogen interactions. The species-specificity in TRIM5α discussed earlier can be attributed to amino acid changes in a cluster of five positively selected sites in the B30.2 domain. Swapping this patch of residues between the rhesus and human orthologs conferred the anti-HIV-1 capsid specificity of the rhesus TRIM5α protein onto the otherwise non-restrictive human protein. Remarkably, a single residue out of the five positively selected residues can explain many of the species specificity differences that underlie why the rhesus, but not the human, TRIM5α could protect against HIV-1 infection[87,88]. This and other studies of primate restriction factors have revealed that single amino acid changes at host-virus interfaces can govern host-virus interactions[89–92]. Thus, identification of positively selected sites can guide the identification of viral target recognition domains and dissect the molecular rules of target specificity.

The L4 loop is the antiviral specificity domain against Orthomyxoviruses

To understand how MxA might engage with its numerous viral targets, a previous study in our lab took an evolution-guided approach that leverages the evolutionary history of host-virus interactions[93]. By identifying rapidly evolving residues in MxA, potential viral contacts were characterized by functional assays involving primate MxA orthologs with different antiviral potency against Orthomyxoviruses. The study found that, although the majority of the MxA gene has evolved under purifying selection to maintain structural and enzymatic properties, a few codons (3.7%) have evolved under positive selection. This finding indicates that positive selection has acted across MxA in response to antagonistic evolution with viruses.

The majority of positively selected sites were located in the MxA stalk domain. Within the stalk domain, the unstructured loop L4 contains 18.6% of sites evolving under positive selection. The enrichment of positively selected sites in L4, suggested that this loop has been subject to recurrent and strong positive selection during primate evolution. To evaluate the functional consequences of rapid evolution in MxA, 16 primate species spanning hominoids, old world moneys and new world monkeys were evaluated for their antiviral activity against THOV. Only human, chimpanzee, gorilla and orangutan reduced THOV replication by 4 to 20-fold. This species specificity shows how genetic divergence of primate MxA orthologs bears dramatic functional consequences for MxA antiviral activity against THOV.

Chimeras between the active human and the inactive African Green Monkey MxA revealed that the L4 loop is the key determinant of MxA antiviral specificity against THOV. Introduction of the human MxA L4 loop into the African Green Monkey MxA resulted in a dose-dependent rescue of activity against THOV. Human MxA and African Green Monkey MxA differ by four amino acids in the L4 loop, two (561 and 566) of which are evolving under positive selection. Surprisingly, a single amino acid confers MxA antiviral specificity for THOV. Since the serine residue at 566 is shared between human MxA and other inactive primate orthologs, the analysis was focused in residue 561. A change from phenylalanine present in humans to the valine in African Green Monkey (F561V) in the Human MxA protein ablated its activity. The reciprocal change in the African Green Monkey MxA protein conferred a gain of function. Indeed, when a panel of amino acids was tested at residue 561, only mutants with aromatic residues phenylalanine, tyrosine and tryptophan restricted THOV. These findings highlight a biochemical requirement for an aromatic amino acid at residue 561. Similar specificity requirements were found for the H5N1 avian influenza strains. Taken together, these studies reveal that the L4 loop is the viral recognition

domain for Orthomyxoviruses and that single residues in the L4 loop can dictate the outcome of the MxA-NP interaction. These studies led to the hypothesis that MxA antiviral breadth relies on specificities encoded by single residues. In this way, residue 561 could be specific towards Orthomyxoviruses, while non-561 residues could dictate specificity towards other viruses.

In my thesis, I take a step back from this hypothesis to comprehensively characterize all rapidly evolving sites in the L4 loop and evaluate their role in Orthomyxovirus recognition. Traditional 'deep mutational scanning' creates single amino acid mutant variants of individual proteins to screen for 'loss-of-function' mutations[94]. However, these methods are limited to single or a couple of mutations and cannot address the role of epistasis due to an exponential increase in complexity[95]. Using evolution-guided approaches, I have reduced the dimensionality of sequence space to be mutagenized by focusing on MxA residues most likely to lead to changes in antiviral activity against THOV. Through this approach, I have identified MxA variants with enhanced antiviral activity that could guide the design of potent therapeutic antivirals for important human pathogens such as influenza viruses.

Chapter 2: Methods

NNS library construction

The L4 loop five-point mutant library was constructed using oligonucleotide-directed mutagenesis of the rapidly evolving sites in MxA. To mutate the five rapidly evolving sites in the L4 loop (positions 540, 561, 564, 566 and 567), two mutagenic oligonucleotides (one sense, one antisense) were synthesized (IDT) that contain sequence complementarity to 70 base pairs (bp) in the region encoding for the rapidly evolving residues. For the targeted positions, the oligonucleotides contain NNS codons, in which N is a mixture of A, T, C and G, and S is a mixture of G and C. This biased randomization results in 32 possible codons with all 20 amino acids sampled –a significant decrease in library complexity and incidence of stop codons without the loss of amino acid complexity. One round of PCR was carried out with either the sense or antisense oligonucleotide and a flanking antisense or antisense oligonucleotide. A second round of PCR using a combination of first round products and both flanking primers produced the full-length double stranded product. The full-length PCR product was purified, digested and ligated into the pQXCIP retroviral expression vector. The ligation was purified and eluted in 10-ul dH_2O, which was used to transform ElectroMAX™ DH10B™ cells (ThermoFisher). 836 colonies were randomly selected and overnight expanded in LB liquid medium containing 100ug/ml of ampicillin.

Sequencing

The library variants and all other MxA clones in this study were sequenced in full-length using standard Sanger sequencing with a set of four primers: F1: 5'-CCG CTG GTG CTG AAA CTG AAG AAA C-3', R1: 5'-ACC ACC AGA TCA GGC TTC GTC AAG A TTC-3', F2: 5'-ACA

GTA TCG TGG TAG AGA GCT GCC-3' and R2: 5'-AGC ATG GCC CAG GAG GTG GAC CCC G-3'.

Plasmids

MxA L4 loop variants were cloned into the NotI/EcoRI sites of the retroviral expression vector pQXCIP-3x flag. Point mutants were generated using the Q5® site-directed mutagenesis kit (New England Biolabs).

Minireplicon assays

The THOV minireplicon assay was performed in 293T cells in a 96-well format. Briefly, 4.0 ng each of PB2, PB1 and PA, 1.0 ng of NP, all in pCAGGS expression vector, as well as 20 ng pHH21-vNP-FF-Luc (firefly luciferase), 50 ng of pRL-SV40-R*luc* (Promega) and varying input or a fixed amount of 200 ng MxA plasmids were co-transfected into 293T cells using the reagent TransIT-LT1 (Mirus Bio). At 24h post-transfection, luciferase activity was measured using the Dual-Glo system (Promega). Expression of MxA proteins was detected by Western blot analysis using mouse ANTI-FLAG® monoclonal antibody F4042 (Sigma) and rabbit anti-β-actin polyclonal antibody AB-8227 (Abcam). Corresponding data from minireplicon assays and Western blots are derived from samples generated from a single master mix. All experiments were done in triplicate.

Logo plots

Difference logo plots comparing super-restrictors, restrictors or non-restrictor classes to the full library. For each class of clones, we constructed an amino acid position frequency matrix, and

generated difference logo plots using the DiffLogo R package[96]. Amino acids shown above the y=0 line are enriched in each class of clones compared to their background frequency in the whole library, and those below the y=0 line are depleted. The total height of each stack of letters represents how different the two classes of clones are from one another, and the height of each amino acid letter reflects how much it contributes to the overall difference in frequencies.

Co-immunoprecipitation

24 h post-transfection, cells were infected with THOV (MOI 10). 24 h post-infection, co-immunoprecipitation analysis was performed. The cells were lysed in 50 mM Tris (pH 8.0), 150 mM NaCl, 1 mM EDTA, 0.5% Nonidet P-40 and incubated with anti-FLAG-M2 affinity gel (Sigma-Aldrich) for 2 h at 4 °C. After extensive washing, the precipitates as well as the whole-cell lysates (WCL) were subjected to standard Western blot analysis using antibodies against the MxA protein and THOV NP.

Statistical Analysis

Data analyses were done using GraphPad Prism 7.0 software. All data are shown as mean ± SEM. Statistical analysis was performed using unpaired t -tests with two tailed, 95% confidence. P values less than 0.05 were considered statistically significant.

Chapter 3: Evolution-guided design of super restrictor antiviral protein reveals a breadth-versus -specificity tradeoff

Introduction

The innate arm of mammalian immunity encodes dozens of antiviral proteins that act cell-autonomously to block viral replication[12]. As a response to the pathogenic pressure exerted by viruses, there is selection for amino acid-altering mutations in antiviral genes that confer increased ability to inhibit viruses. This accelerated evolution, or positive selection, can be detected, in a gene or specific codons, by measuring the rate of amino acid-altering changes (nonsynonymous, dN) over silent changes (synonymous, dS)[6]. Many antiviral proteins described to date exhibit signatures of rapid evolution, an indicator of their direct interaction to rapidly evolving pathogens[18]. Functional studies based on positive selection analysis have shown that that residues under positive selection are found at the protein-protein interface between viral and host factors and that changes to these residues determines viral target specificity[91,97,98].

MxA is a dynamin-like GTPase with potent restrictive antiviral activity against a wide spectrum of RNA and DNA viruses[67]. Evolution-guided functional studies, along with biochemical characterization show that MxA restricts members of the Orthomyxoviridae virus family via direct binding of the disordered Loop L4 region to the viral nucleoprotein (NP)[67]. Five sites in the Loop L4 have been recurrently selected throughout primate evolution (540, 561, 564, 566, 567). Remarkably, a single rapidly evolving amino acid (561) in this cluster solely determines antiviral specificity against Thogoto virus (THOV) and influenza virus (IAV)[45,97]. However, since the MxA target interface its likely much bigger than a single residue, it is possible that other rapidly evolving sites in the L4 loop play a role in mediating specificity.

Previous approaches to examine the role of positive selection on antiviral proteins have looked backwards in evolutionary time to uncover the rules that govern host-pathogen interactions[6]. In this study, we use evolutionary insights from positive selection as a guide to design versions of antiviral proteins through a novel forward combinatorial screen to test the hypothesis that unconstrained amino acid diversity at viral binding sites can reveal antiviral protein variants with increased restrictive potency and modified specificities. We generated novel combinations of amino acids at the rapidly evolving sites in the L4 loop of MxA and evaluated their antiviral activity against known targets. In doing so, we identified human MxA variants with increased antiviral potency against THOV relative to the versions found in nature that we term 'super restrictors', or evolution-guided variants of a natural antiviral protein with increased potency. We find that super restriction to THOV leads to a partial loss of function against IAV. Thus, our data suggest a model on which antiviral specificity domains (i.e. L4 loop) populate a functional sequence space to target multiple viruses simultaneously.

Results

Variation in rapidly evolving sites determines potency and specificity

To evaluate if amino acid variation in the rapidly evolving sites could generate variants with increased antiviral activity relative to the versions found in nature ("super restrictor factors"), we designed a gene library that encodes for variants of human MxA containing random combinations of all amino acids in the five rapidly evolving sites in the L4 loop region in MxA (see Methods, **Figure 5**). The library was generated with a set of mutagenic primers that introduce a combination of NNS codons in the rapidly evolving sites. We randomly selected 600 clones from the library and evaluated the antiviral activity of each clone individually in triplicates using the

THOV minigenome assay, as previously described[99]. We selected THOV to evaluate the antiviral activity of MxA variants because of its robust dynamic range. In this assay, the viral polymerase components (PB2, PB1, PA) and NP are delivered *in trans* to drive the expression of a firefly luciferase from a viral RNA minigenome in antisense orientation that is flanked by viral 5' and 3' UTRs. Firefly luciferase expression is detected 24h-post transfection and normalized to renilla

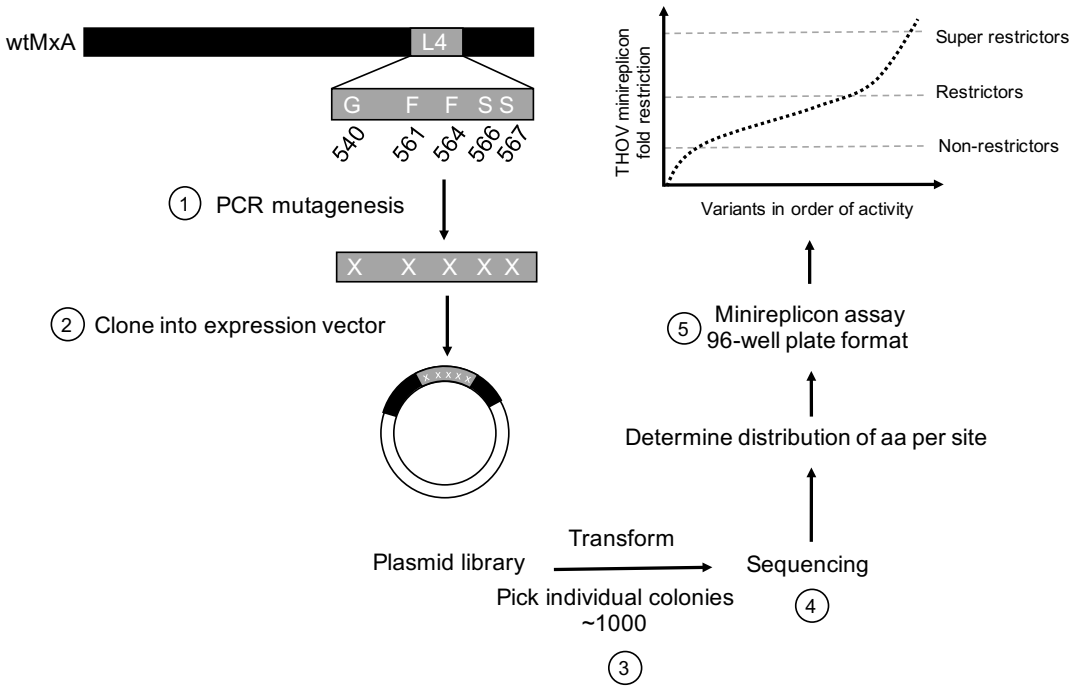


Figure 5. Schematic of combinatorial mutagenesis screen.

The mutant library was generated using oligonucleotide-directed mutagenesis of the rapidly-evolving sites in the Loop L4 (see methods). The resulting gene library was cloned into the expression vector pQXCIP and transformed in bacterial cells. Six hundred individual colonies were randomly selected, plasmids were extracted and sequenced. The antiviral activity of each variant was evaluated on the THOV minigenome assay (see Methods).

luciferase levels, serving as a transfection control. Luciferase expression in the absence of MxA is normalized to 100%. Overexpression of human MxA reduces luciferase expression in a dose-responsive manner. The GTPase-deficient and antivirally inactive mutant, T103A, serves as the negative control. We considered the average and standard deviation of nearly a hundred replicates of wild-type human MxA and catalytic mutant T103A (Supplemental Figure 1, Mean = 3.97%, SD= 2.9; input of MxA plasmid: 50ng) to set our active cutoff at 10-fold reduction of luciferase

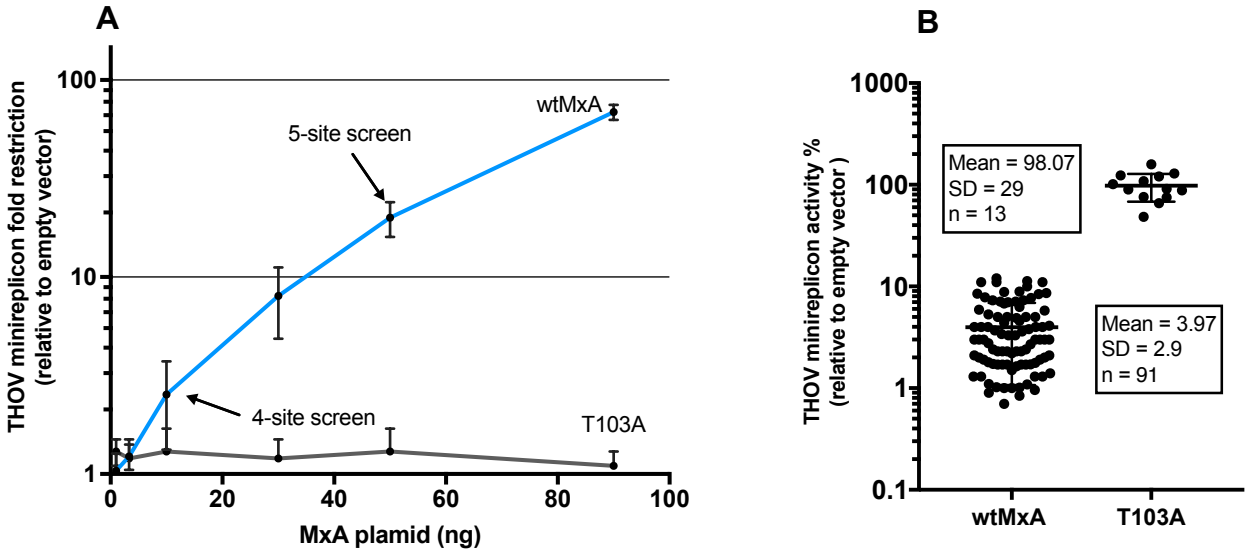


Figure 6. Restriction sensitivity range of the minireplicon assay.

(A) Dose-responsiveness of the antiviral activity of wtMxA and the catalytic mutant MxA-T103A was determined by co-transfecting increasing amounts of MxA expression constructs with the THOV minigenome into HEK-293T cells. The data is shown in average fold restriction relative to empty vector. (B) Threshold determination. wtMxA replicates for threshold determination. 91 distinct clones of wtMxA and MxA-T103A were tested in triplicates in the minireplicon assay. Error bars represent the ±SEM.

expression. By these parameters, about 5% of the variants tested are THOV restrictors (**Figure 7A**). In the restrictor group, all of the active variants encode an aromatic amino acid (Phe, Tyr, Trp) at residue 561. These findings are congruent with a previous study that determined the necessity of an aromatic residue at site 561 for antiviral activity against THOV. To evaluate if all aromatic amino acids produced variants with similar antiviral activity potency, we grouped the variants by their amino acid at 561 (**Figure 7B**). Whereas Phe and Tyr variants have comparable antiviral activity levels, Trp variants are mostly non-restrictors. This data suggests that an aromatic residue at position 561 is necessary, but not sufficient to confer antiviral activity against THOV. Interestingly, mutants with an aromatic amino acid at residue 561 that lost activity against the THOV minireplicon, still retained the ability to inhibit IAV in a minigenome with polymerase genes of the strain A/Vietnam/1203/2004 (H5N1) at levels comparable to the antiviral activity of

wild-type MxA (**Figure 7D**). Thus, we conclude that changes at non-561 sites can have target-specific negative epistatic effects on antiviral activity.

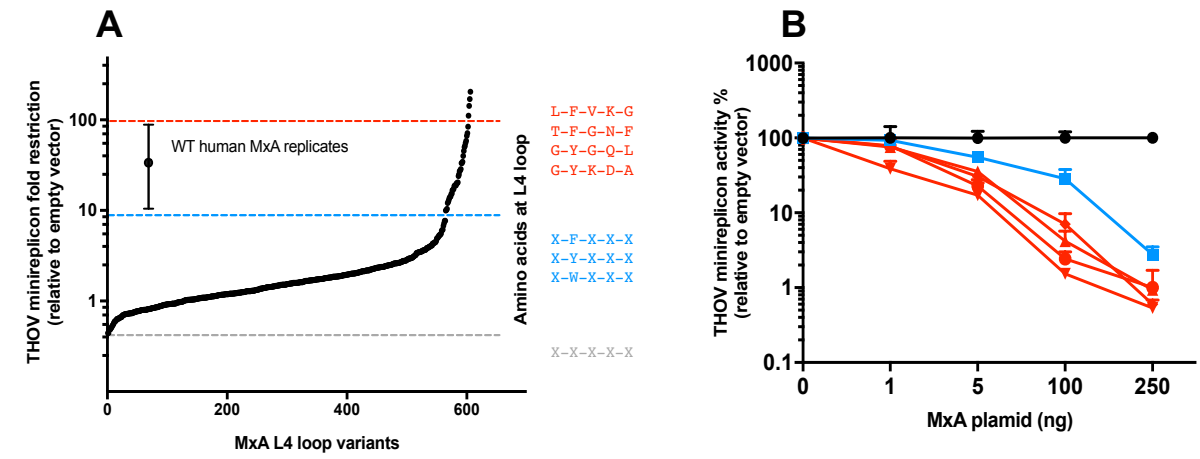
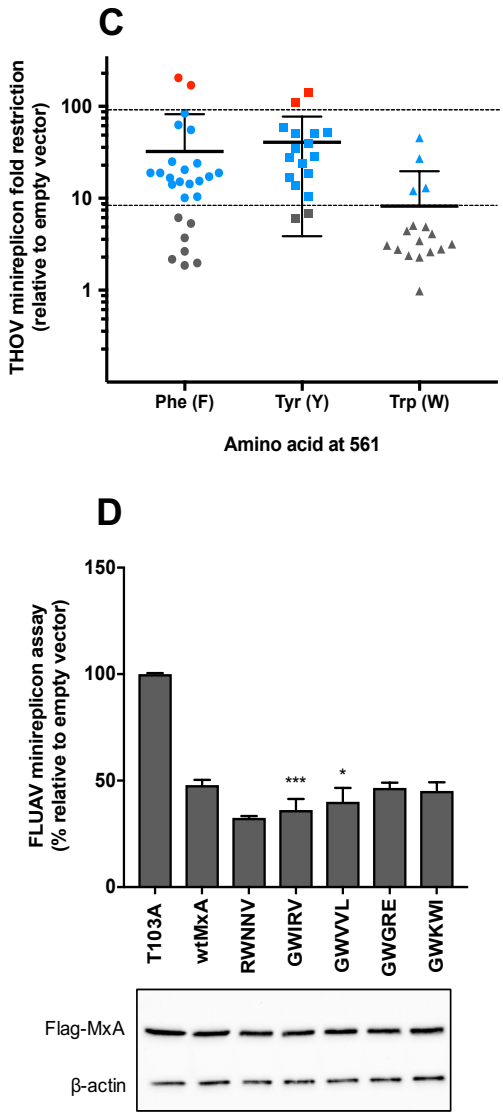


Figure 7. Mutagenesis of rapidly evolving sites in the MxA L4 loop alters antiviral potency and specificity.



(**A**) Functional consequences of amino acid variation at rapidly evolving sites in L4 loop. Six hundred MxA variants were evaluated for their ability to decrease polymerase activity in a THOV minigenome system. Briefly, HEK 293T cells were co-transfected with expression plasmids encoding for the MxA variants (50ng of expression plasmid) and the THOV minireplicon system, including a reporter construct encoding firefly luciferase under the control of the THOV viral promoter, in a 96-well plate format. After 24 h, firefly luciferase activity was determined and normalized to the constitutively co-expressed Renilla luciferase (transfection efficiency control). Results are presented relative to the activity in the absence of MxA, the empty vector control, and as the mean average of fold restriction in three independent experiments. The variants are shown in increased order of antiviral activity. The activity of 91 replicates of wild-type MxA was used to establish a 10-fold cutoff for restrictive variants (Supplement 3). The right panel shows the amino acid consensus sequence of each category. The dashed lines represent category thresholds (red = super restrictors, blue = restrictors, and grey = non-restrictors).

(B) Variants LFVKG, GYGQL, GYKDA have increase antiviral activity compared to wild-type at a range of MxA input plasmid. Dose-responsiveness of the antiviral activity of wild-type MxA and L4 loop super restrictor mutants was determined by co-transfecting increasing amounts of MxA expression constructs with the THOV minigenome into HEK-293T cells. MxA antiviral activity is measured as described in Figure 7A. The data points represent the mean average of three independent experiments. Flag-tagged MxA and β-actin protein levels were detected in THOV minigenome assay lysates by Western blot.

(C) An aromatic amino acid is necessary but not sufficient to confer antiviral activity against THOV. Activity distribution of variants categorized by aromatic amino acid Phe, Tyr or Trp at site 561. The dashed lines represent the distribution of wild-type MxA replicates. The points colored in blue represent variants with wt-like activity, grey represents non-restrictors and red colored dots represent super restrictor variants. Each data point is the mean average of three biological replicates.

(D) THOV non-restrictors retain activity against IAV in the VN/04 minireplicon system. MxA variants (300ng) and the components of the replicon system, including a reporter construct encoding for firefly luciferase under the IAV promoter and *Renilla* luciferase to monitor transfection efficiency, were co-transfected in 293T cells. Firefly luciferase activity was measured in the cell lysates at 24 h post-transfection and normalized to the expression of the *Renilla* luciferase. Error bars represent standard deviation of three biological replicates. *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$ (t-test) for values compared to wild-type MxA.

In contrast, to the more common loss of function mutations in the screen, we also found more rare examples of "super restrictors", i.e. variants with greater activity than wild-type MxA; 0.6% of variants had a 2-fold to a 5-fold increase in antiviral activity relative to wild-type MxA (**Figure 7A and 7B**). All of these variants encode for an aromatic amino acid (Phe or Tyr) at site 561 and a combination of other amino acids at the remaining sites. We termed these variants 'super restrictors' LFVKG, GYKDA, TGNFG and GYGQL, based on the amino acid identity at the five rapidly evolving sites. We confirmed the enhanced antiviral activity of super restrictor variants by testing their dose-responsiveness in the THOV minireplicon assay (**Figure 7C**). These findings suggest that variation in the rapidly evolving sites can, not only determine antiviral activity, but also has the capacity to modulate the antiviral potency of MxA via positive epistasis. Interestingly, mutants with an aromatic amino acid at residue 561 that lost activity against the THOV minireplicon, still retained the ability to inhibit IAV in a minigenome with polymerase genes of the strain A/Vietnam/1203/2004 (H5N1) at levels comparable to the antiviral activity of wild-type MxA (**Figure 7D**). Thus, we conclude that changes at non-561 sites can have target-specific negative epistatic effects on antiviral activity.

Genetic basis determination of super restriction phenotype

We were interested in uncovering the evolutionary paths that can lead to super restriction from the parental WT human MxA L4 loop sequence. Despite their enhanced antiviral potency, we do not observe super restrictors in nature. It is possible that to achieve super restriction, multiple changes, and therefore intermediates, are necessary. In the analogy of a fitness valley, the current MxA sequence could exist in a local optimum. Any non-synonymous mutations could push MxA away from the local optimum and render the protein less restrictive. If MxA its under constant selective pressure to maintain activity, then these intermediates will not be viable. An alternate explanation is that super restrictors are not naturally occurring because there is no selection for increased antiviral activity. To explore the former, we set to trace the evolutionary paths that lead to super restriction. All super restrictors encode for the necessary aromatic residue at site 561

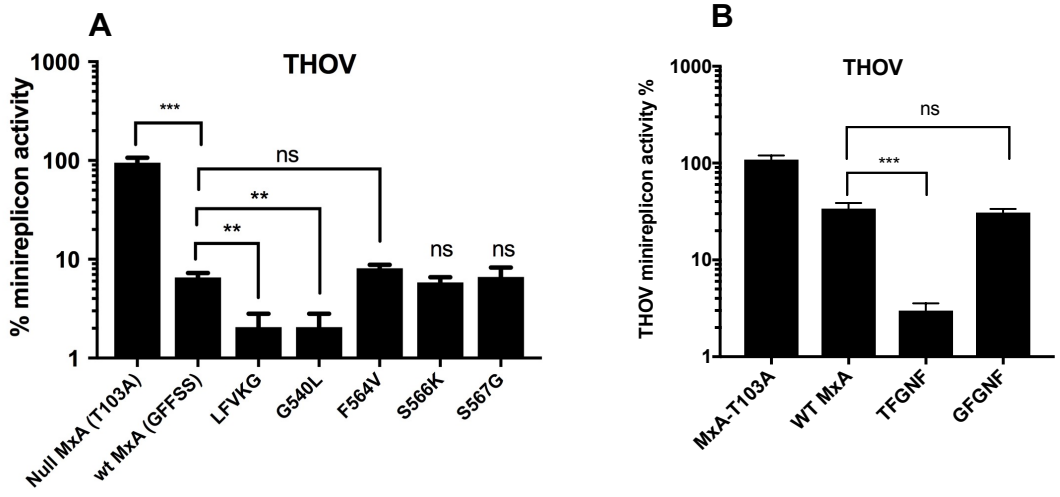


Figure 8. Site 540 largely determines super restriction.

(A) Single point mutants in wtMxA background reveal role of position 540. Mutations in LFVKG were individually introduced in the wtMxA background and tested for their ability to inhibit THOV minigenome activity. Error bars represent standard deviation of three biological replicates. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001 (t-test) for values compared to wild-type MxA.

(B) Reversion to the wild-type amino acid at 540 in super restrictor TFGNF results in a loss-of-function. The mutation T540G was introduced in the super restrictor variant TFGNF and its activity evaluated in the THOV minigenome. Error bars represent standard deviation of three biological replicates. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001, ****p ≤ 0.0001 (t-test) for values compared to wild-type MxA.

(**Figure 7A and 7C**) and the remaining sites contain a variety of amino acids with diverse biochemical properties. To determine the minimum number of changes required for super restriction and map the genetic basis for super restriction, we generated single amino acid swaps between WT MxA (GFFSS) and the super restrictor LFVKG. Since site 561 encoded for phenylalanine in both variants, we generated four variants using site-directed mutagenesis: G540L, F564V, S566K and S567G. The restrictive activity of these variants was evaluated using the THOV minigenome, as described before. The mutant G540L completely recapitulated the phenotype of the super restrictor LFVKG, whereas the mutants F564V, S566K and S567G resulted in an antiviral activity similar to WT levels (**Figure 8A**). To investigate if in other super restrictors site 540 played a role in antiviral activity, we reverted the 540 position to glycine, the wild-type amino acid, in variant TFGNF. A change from a threonine to a glycine, was sufficient to decrease the antiviral activity to wild-type levels (**Figure 8B**). Taken together, this data suggests that site 540 has a modulatory role in antiviral activity and is a significant determinant for super restriction.

Modified 561F screen reveals potent super restrictor variants

Since our initial screen identified super restrictors at a low frequency because site 561 was randomized, we designed a 4-site screen in which site 561 is fixed to the parental wild type amino acid, phenylalanine (Phe) in order to better explore the super restrictor landscape. For this iteration, we randomly selected 168 variants and assessed their antiviral activity individually in triplicates using the THOV minigenome assay. We found that 12.8% of the variants in this screen have increased antiviral activity compared to wild type MxA, an increase of more than 20-fold from our previous screen (**Figure 9A and 9B**). The super restrictors identified in this iteration range from 2-fold to a 10-fold increase in antiviral activity relative to wild-type MxA. We confirmed the enhanced antiviral activity phenotype of these variants by assessing the dose-

responsiveness of the THOV minireplicon to increasing concentrations of MxA encoding plasmid (**Figure 10A**). Super restrictors QFAYS, VFRSV and TFAMC are the most potent variants identified from the two screens combined. Taken together, these results suggest that once the requirement for an aromatic amino acid at 561 is met, changes at the other rapidly evolving sites have a higher probability of increasing the antiviral potency of MxA. These results provide evidence that suggests that non-561 residues have a modulatory role in determining antiviral potency.

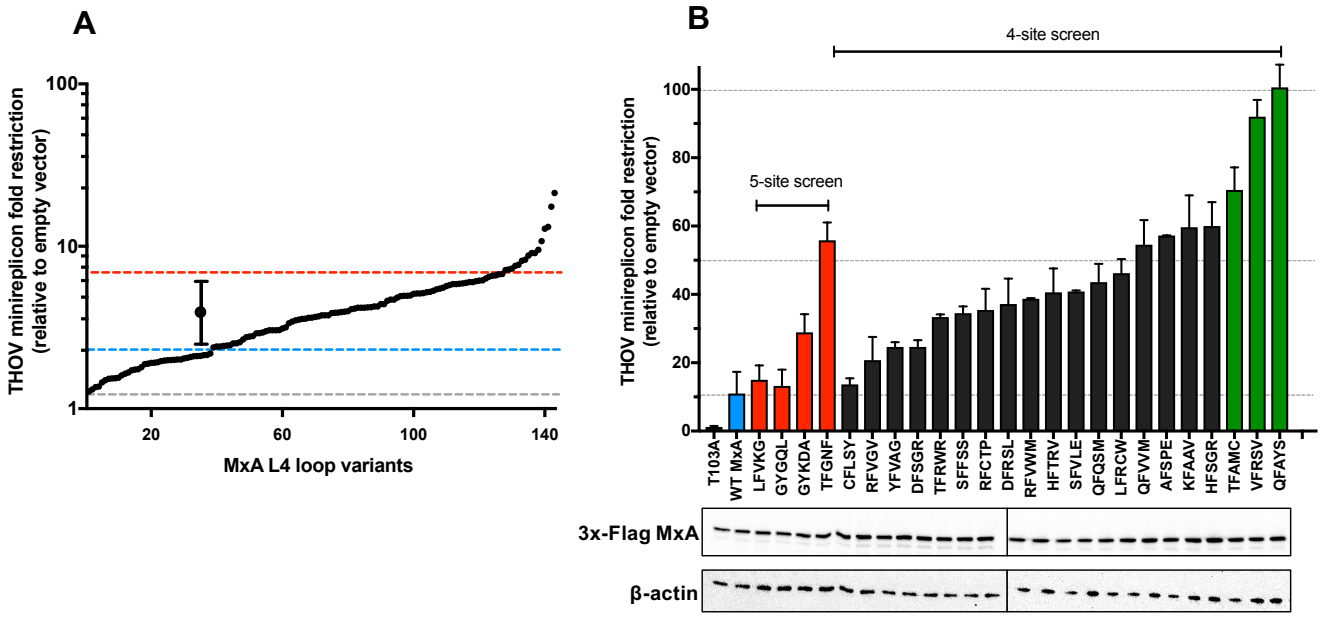


Figure 9. Modified 561F screen reveals potent THOV super restrictor variants.

(A) Functional distribution of MxA L4 loop variants (561F). The 561F MxA L4 loop library was constructed as described in Figure 1A. The antiviral activity against of 141 variants was tested in the THOV minigenome assay, as described in Figure 1B. In this iteration of the screen the amount of MxA expression plasmid was lowered to 30ng to avoid saturation and maximize the dynamic range. Results are presented in fold restriction relative to the activity in the absence of MxA, the empty vector control, and as the mean average of three independent experiments. The variants are shown in increased order of antiviral activity. The dashed lines represent the upper limit for non-restrictors (grey), restrictors (blue) and super restrictors (red).
(B) Most potent super restrictors show an 8- to 10-fold improvement in antiviral activity relative to wtMxA against THOV in minigenome assay. Antiviral activity of super restrictors is shown in fold restriction of normalized luciferase activity of the THOV minigenome relative to an empty vector control. Top restrictors from both screens were tested in parallel in the THOV minigenome assay, as described in Figure 1B. The error bars represent the standard deviation from the mean (± SEM) in three independent experiments.

The target for MxA is a multimer of the NP protein of orthomyxoviruses. We have previously shown that the L4 loop in MxA influences binding to NP. We tested if increased affinity or stabilization of the MxA-NP interaction could explain the super restriction phenotype by monitoring the association of NP with Flag-tagged Mx super restrictors in THOV-infected cells. Super restrictors QFAYS, VFRSV and TFAMC interact more potently with THOV NP when compared to wtMxA (**Figure 10B**). Furthermore, the non-restrictor PFFSS has lost the association with THOV NP. This indicates that the mechanism for super restriction involves greater avidity or a stabilization of MxA interaction with the THOV NP

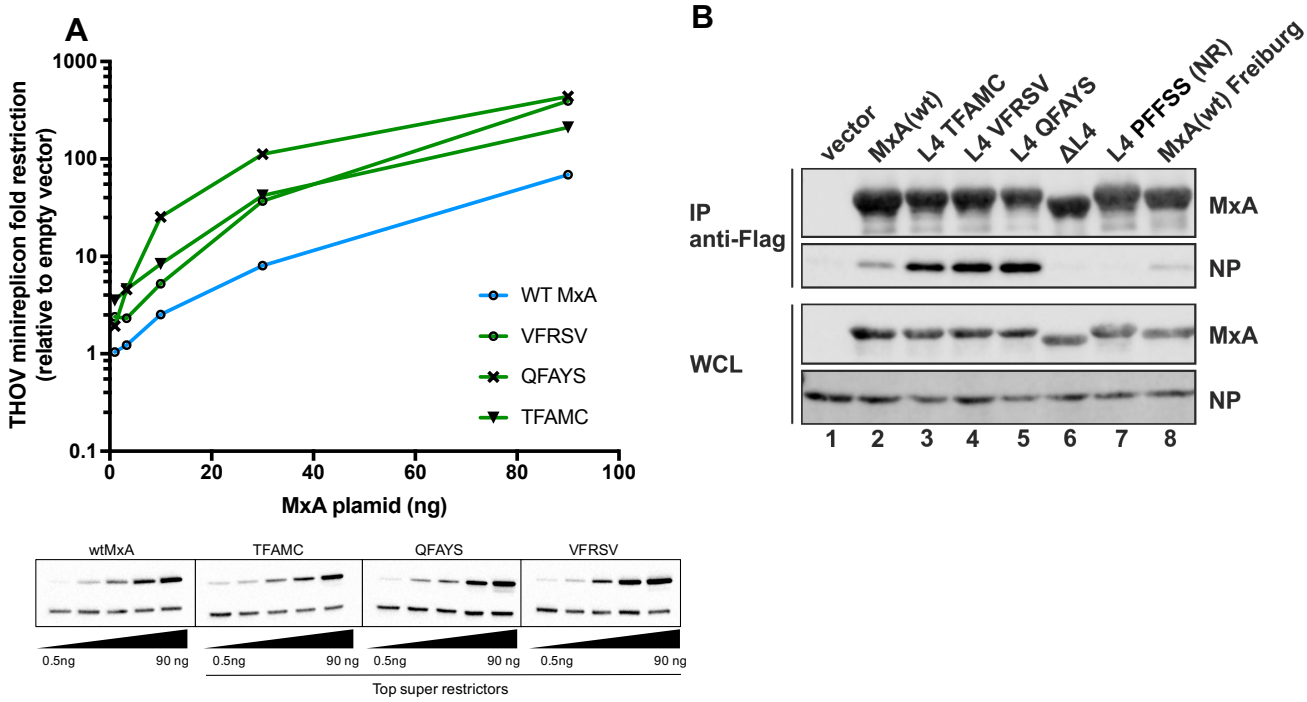


Figure 10. Super restrictors potently interact with THOV NP.

(A) Super restrictors have comparable expression to WT human MxA. Dose-responsiveness of the antiviral activity of wild-type MxA and L4 loop super restrictor mutants was determined by co-transfecting increasing amounts of MxA expression constructs with the THOV minigenome into HEK-293T cells. MxA antiviral activity is measured as described in Figure 1B. The data points represent the mean average of three independent experiments. Flag-tagged MxA and β-actin protein levels were detected in THOV minigenome assay lysates by Western blot.

(B)Super restrictors potently interact with THOV NP. Co-immunoprecipitation of viral NP with wtMxA and L4 loop mutanys TFAMC, VFRSV and QFAYS was performed with cell lysates of THOV infected 293Ts transfected with the Flag-tagged Mx expression constructs. Immunoprecipitation was performed using anti-Flag affinity gel and the precipitated proteins were detected by Western blot with viral NP and MxA-specific antibodies.

Comprehensive library analysis for amino acid preferences

To investigate if super restrictors, non-restrictors or restrictors have preferred amino acids in any of the sites mutagenized, we performed differential logo plot analyses that compare each one of the subsets of variants identified by their functionality in the THOV minigenome assay to the starting library. We find that when comparing preferences in the super restrictor vs. the wt-like restrictor pool, glutamate is preferred at position 540 in super restrictors. However, we also find glutamate as a preferred amino acid in the non-restrictor to restrictor pool. This indicates that for super restriction a single amino acid is not sufficient, rather super restrictors have to encode for a combination of amino acids at the rapidly evolving sites.

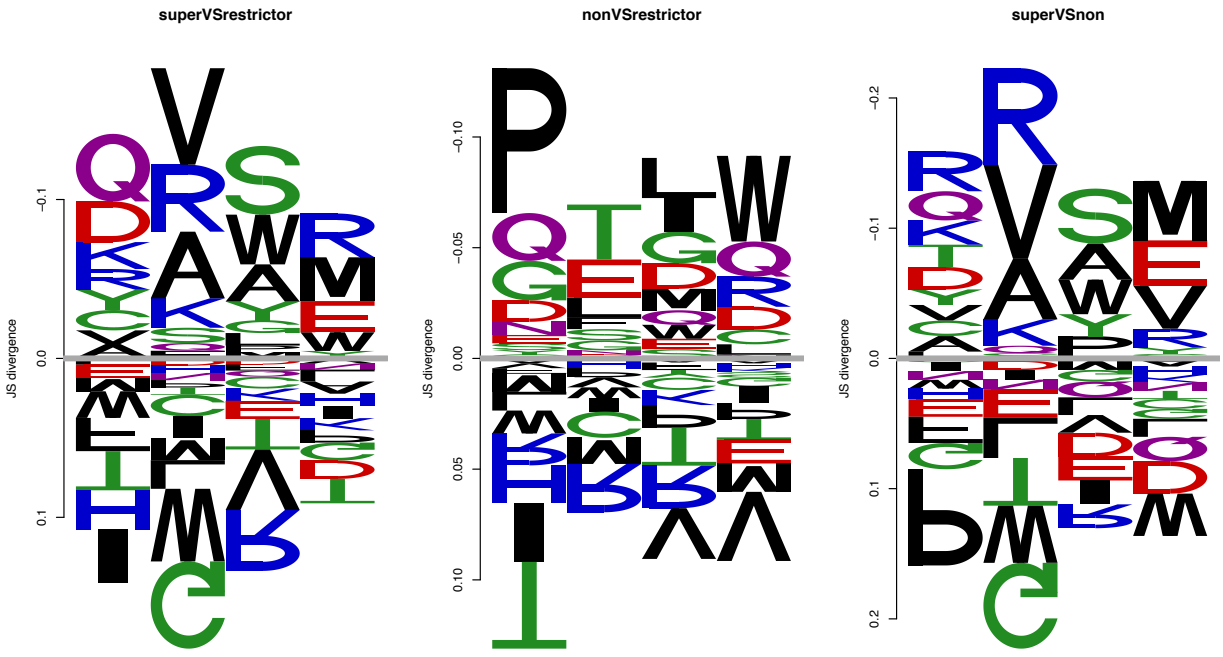


Figure 11. Glutamine at site 540 is enriched in super restrictor pool.

Difference logo plots comparing super-restrictors, restrictors or non-restrictor classes to the full library. For each class of clones, we constructed an amino acid position frequency matrix, and generated difference logo plots using the DiffLogo R package. Amino acids shown above the y=0 line are enriched in each class of clones compared to their background frequency in the whole library, and those below the y=0 line are depleted. The total height of each stack of letters represents reflects how different the two classes of clones are from one another, and the height of each amino acid letter reflects how much it contributes to the overall difference in frequencies.

Super restriction is dictated by interactions between multiple rapidly evolving sites

MxA variants with Glutamine (Q) at position 540 are enriched in our 4-site super-restrictor pool (Figure 11); three of the ten most-potent super-restrictors include the variants QFAYS, QFQSM and QFVVM. We tested the role of amino acid 540 by making revertant mutations in each of these three super-restrictors, swapping the glutamate (Q) residue at position 540 to glycine (G) present in wtMxA. For all three variants, this reversion leads to drop in restriction activity, in one case to below wtMxA levels (**Figure 12A**). Thus, Q540 is necessary for the super-restriction we observed these three super-restrictors. However, Q540 is not sufficient to explain the degree of super-restriction since QFAYS, QFQSM and QFVVM are more potent at restricting THOV than the single Q540 change in wtMxA, which leads to only a 2-fold improvement in antiviral activity. These results imply that position 540 as well as other rapidly evolving positions in L4 together mediate the super-restriction phenotype. Furthermore, we also mutated the Glycine at residue 540 (G540) to every other amino acid in the wtMxA backbone (**Figure 12B**). We found that a single G540A substitution in wtMxA increases THOV restriction 5-fold, while a G540Q mutation increased restriction by 2-fold. However, further reinforcing the contribution of other residues, none of the variants at 540 in the wtMxA background achieved the level of super-restriction in the most potent variants (**Figure 12B**). Moreover, variants QFQSM and QFVSM are both super-restrictors, whereas QFLSM is a non-restrictor, despite differing only at position 564 (**Figure 13A**) Thus, the super-restriction phenotype is due to contributions of changes at multiple amino acids under positive selection.

Molecular basis of MxA super restriction

We next tested whether or not super-restriction could simply be the additive effect of individual amino acid changes at each of the rapidly evolving positions in L4 by swapping in a

'preferred' alanine residue at position 540 (**Figure 13B**) into three super-restrictor combinations – HFSGR, TFAMC and VFRSV – that each carry an 'unpreferred' residue at position 540 (**Figure 12B).** If super-restriction were truly additive, the X540A swap should further increase super-restriction by each of these variants. Contrary to this prediction, we find that there is no change in super-restriction for the HFSGR and TFAMC variants (**Figure 13B**). Furthermore, an alanine swap into VFSRV dramatically lowered restriction activity by 10-fold (**Figure 13B**), even though the same swap into wtMxA (GFFSS) increased restriction by five-fold (**Figure 13B**). Thus, these results show that the paths to super-restriction are not straightforward and fitness peaks cannot be

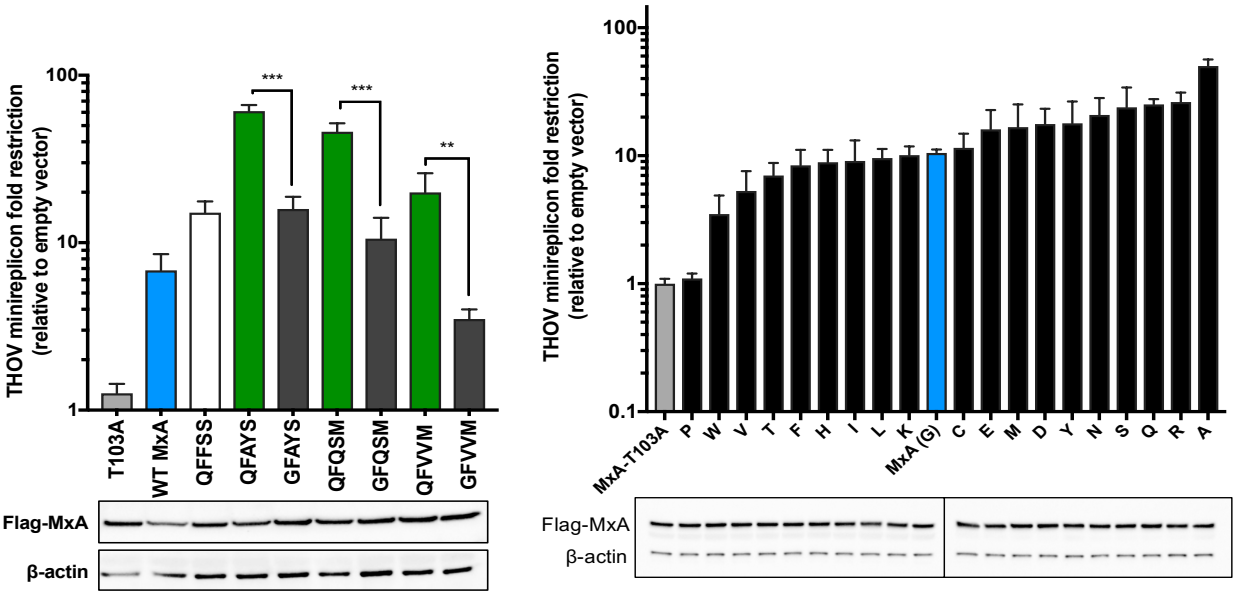


Figure 12. Multiple sites mediate super restriction phenotype.

(A) Super restriction is achieved through epistatic interactions between rapidly evolving sites. Super restrictors encoding for Glutamine (Q) at site 540 were modified to encode for the wild-type amino acid in human MxA, Glycine (G) using site-directed mutagenesis. The Q540G revertant mutants (dark grey) were evaluated for their ability to restrict luciferase activity in the THOV minigenome assay (MxA plasmid input = 30ng). The activity of reciprocal G540Q mutation in the wild-type background was also assessed (white bar).

(B) Mutation at site 540 enhances antiviral activity but not to super restrictor levels. A panel with all amino acids at position 540 was generated by site-directed mutagenesis (G540X mutants). The antiviral activity of the G540X panel was assessed in the THOV minigenome assay, as described in Figure 1B, with 30ng of MxA expression plasmid. The wild-type is shown in blue and the T103A MxA negative control in grey. The data is shown in fold restriction relative to empty vector as the mean average. Error bars represent the ±SEM. Flag-tagged MxA and β-actin protein levels were detected in THOV minigenome assay lysates by Western blot.

simply ascended by additive effects of individual mutations, but instead reinforce the utility of our combinatorial mutation approach on the L4 residues rather than deep-mutational scanning of each individual L4 residue.

L4 loop-mediated super restriction is virus specific

The fact that MxA restricts a number of different viruses simultaneously led us to investigate next how gaining super restriction to one pathogen (i.e. THOV) can affect the antiviral activity against another virus. It is possible that human MxA exist as a generalist solution to restrict a large number of viruses. Specialization to restrict one virus very potently might render the protein

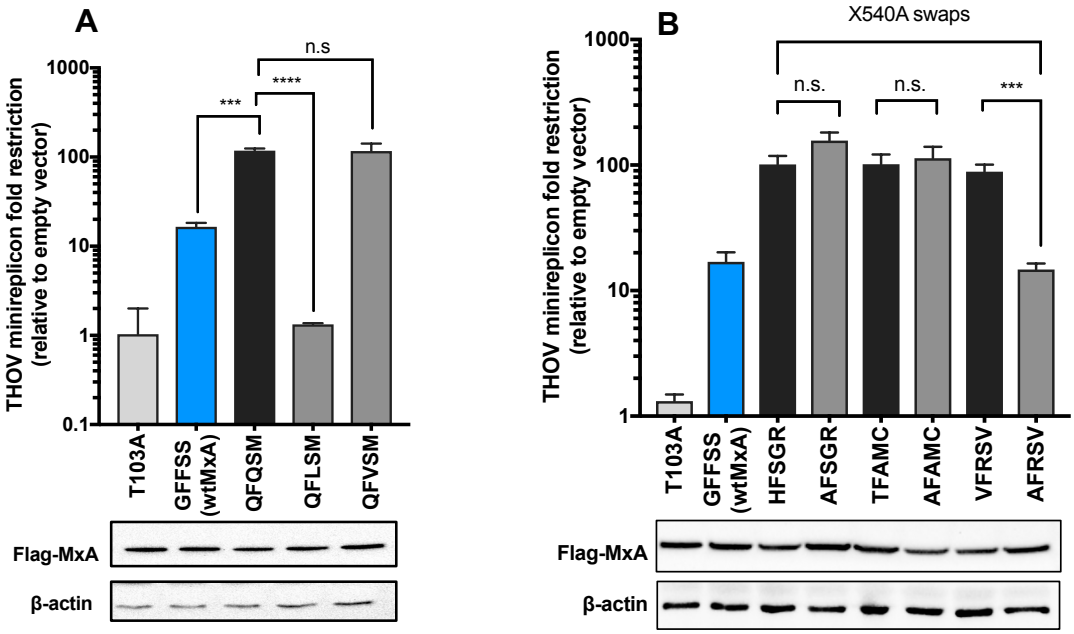


Figure 13. Super restriction is achieved through multiple amino acid changes in L4 loop.

(A) Super restriction is achieved through contributions of changes at multiple sites in L4 loop. Super restrictor QFQSM was modified at site 564 to encode for lysine or valine. The activity of the mutants was assessed using the THOV minigenome, as described before. The data is shown in fold restriction relative to empty vector as the mean average. Error bars represent the ±SEM. Flag-tagged MxA and □-actin protein levels were detected in THOV minigenome lysates by Western blot.
(B) Super restriction conferring amino acids from different backgrounds do not have an additive effect in potency. Unexpected super restrictors HFSGR, TFAMC and VFRSV (black bars) were modified to encode for alanine at residue 540 using site-directed mutagenesis (dark grey bars). The activity of the mutants was assessed using the THOV minigenome, as described before. The data is shown in fold restriction relative to empty vector as the mean average. Error bars represent the ±SEM. Flag-tagged MxA and β-actin protein levels were detected in THOV minigenome lysates by Western blot.

inactive against other viruses. To test this hypothesis, we tested the restrictive capacity of all the super restrictors identified against the H5N1 minigenome (**Figure 14**). As an additional control, we included human MxA with a nuclear localization signal, which has been reported to increase antiviral activity of human MxA from 10-fold to a 100-fold. Therefore, this assay has the dynamic range to show enhancement of restriction activity. We found that THOV super restrictors have reduced antiviral activity against H5N1. Therefore, optimization to restrict THOV, impairs the ability to restrict IAV in a minigenome assay.

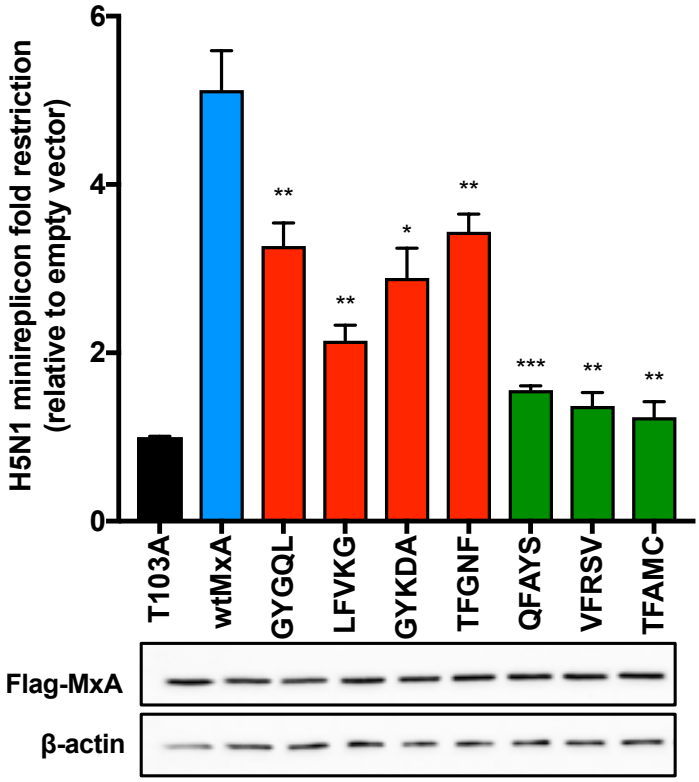


Figure 14. Super restriction is virus-specific.

Super restriction is virus-specific. 293T cells were co-transfected with expression plamids for the MxA variants (300ng) and the H5N1 minigenome system VN/04, including a reporter construct encoding firefly luciferase under the control of the viral promoter. After 24h, firefly luciferase activity was determined and normalized to the activity of a constitutively co-expressed Renilla luciferase. Results are presented relative to the activity in the absence of MxA, the empty vector control and as the mean average of three independent experiments. Error bars represent the ± standard deviation from the mean. *p ≤ 0.05, **p ≤ 0.01, ***p ≤ 0.001 (t-test) for values compared to wtMxA.

Discussion

Viral target recognition domains in antiviral proteins are characterized by a strong signature of positive selection. The functional consequences of amino acid variation at viral target recognition domains has been characterized for more than a dozen primate antiviral proteins, including the IFN-induced innate immunity restriction factor MxA. In primate MxA, multiple and distant clusters of rapidly evolving sites, or hotspots, suggest that these regions could be acting as modular target recognition domains. This model provides an explanation for the broad, yet specific nature of MxA antiviral activity. Indeed, the finding that single residue changes within hotspots can dictate specificity further supports for this model. Here, we evaluated how epistatic interactions in clusters of rapidly evolving sites in the L4 Loop of MxA can determine antiviral specificity and potency. To comprehensively evaluate the role of all rapidly evolving sites at viral recognition surfaces, we have conceptualized and implemented a combinatorial mutagenesis screen, in which sites that are rapidly evolving in antiviral specificity domains are simultaneously randomized. Through this evolution-guided combinatorial mutagenesis approach, we have determined that the rapidly evolving sites in antiviral interfaces play a crucial role modulating the antiviral potency and specificity of MxA.

MxA breadth is at odds with the observed highly specific targeting of unique viral structures. Our study suggests that human MxA exists as a generalist sequence that can target multiple viruses simultaneously. Through variation in the L4 loop, we discovered virus-specific restrictor variants. For example, a handful of variants had lost activity against THOV, but retained the capacity to inhibit IAV. This is evidence that target recognition is achieved by specific combinations of amino acids and epistatic interactions in the L4 loop. These epistatic interactions in the L4 loop make more likely that adaptation to a newly encountered pathogen could

compromise the pre-existing substrate specificity or antiviral range. Thus, if the L4 loop mediates antiviral specificity against multiple targets, the sequence space where MxA is able to target multiple targets at once effectively could be severely constrained. If optimization to restrict one target comes with a tradeoff in specificity, more potent MxA variants could not be sampled by natural selection. Even if the mutations confer a short-term fitness advantage, their restricted antiviral range will result in a loss at the population level over evolutionary time. The THOV super restrictors identified in this study support this model. These variants had more antiviral potency against THOV, while their activity against IAV was weakened. The loss of antiviral breadth can explain why super restrictor variants are not naturally occurring.

This study revealed that rapidly evolving residues can have a modulatory effect on antiviral activity. From the five rapidly evolving residues in L4, mutations in site 540 had a strong effect in potency. Although most amino acids at this site did not have an effect, some others were sufficient to drastically alter antiviral potency. To investigate if amino acid changes at residue 540 could alter the properties of the Loop L4 region, we modeled the MxA structure using the Robetta software (http://robetta.bakerlab.org/results.jsp?id=81973). In the model structure, residues 542-580 are disordered. The Loop L4 is mostly disordered as is comprised by residues 532-572. Site 540, however, is part of a helical structure that starts in residue 537 and ends in site 554. This suggests that, from all rapidly evolving sites, changes at site 540 could have the greatest impact in structural integrity and, therefore, in antiviral activity.

The molecular details of MxA restriction of orthomyxoviruses are not fully understood. However, oligomerization and GTPase activity are indispensable for antiviral activity. Also, multiple lines of evidence suggest that MxA restriction of orthomyxoviruses is mediated through interactions with NP via the L4 loop. Since the super restrictors identified in this study are virus-

specific, we propose that this enhanced antiviral activity is a product of target affinity and not an overall change in the intrinsic properties of MxA. For instance, previous studies have identified variants with increased antiviral activity against both THOV and IAV. This enhancement is the result of a subcellular re-localization to the nucleus or a change in MxA-MxA interaction dynamics[69,100].

Our findings suggest that wtMxA provides a generalist solution to restrict both THOV and IAV and perhaps a large number of other viruses; specialization to restrict one virus might render the protein ineffective against other viruses. Our findings are highly reminiscent of an artificial selection experiment that made the GroEL chaperone highly specific for one substrate but at the expense of its substrate-binding breadth[101]. Similar breadth-specificity tradeoffs have also been observed in cytochrome p450 detoxification genes that protect herbivorous insects from plant counterdefenses and in plant disease resistance (R) genes encoded by the Leucine Rich Repeat (LRR) gene family[102,103]. Our findings suggest that, under threat by multiple viruses, antiviral genes such as MxA appear to be under evolutionary pressure to encode less potent, but more broadly active alleles rather than specialist sequences with the most potent antiviral activity.

Our strategy for the discovery of super-restriction factors preserves protein domains subject to purifying selection and samples mutations at residues already highlighted by positive selection analyses as recurrent targets of adaptation. This focus on only a smaller subset of residues, in this case L4, allows us to more fully explore outcomes of combinatorial mutagenesis by relaxing the constraints imposed by epistasis and historical contingency[94]. Using this strategy, we have been able to identify MxA variants that are 10-fold better restrictors than wtMxA, itself the most potent MxA orthologs known against THOV[21]. Although MxA is unusual in its antiviral breadth, many antiviral proteins evolve under similar selective regimes to MxA, in which they undergo recurrent

positive selection in response to viral pathogens[6]. In many of these cases (*e.g.,* TRIM5, MxB), positive selection appears enriched in unstructured loops like MxA L4[104,105]. We speculate that unstructured loops like L4 represent an ensemble of many conformations that provide MxA with structural and evolutionary flexibility to adapt to binding distinct viral targets, whereas super-restricting variants may be less structurally flexible, trading off increased restriction of certain viral targets with decreased antiviral range. Thus, unexpectedly, we find that breadth-specificity tradeoffs constrain the adaptive landscape of antiviral proteins.

Chapter 4: Molecular basis for MxA super restriction against THOV and IAV

Introduction

MxA is a dynamin-like GTPase with antiviral activity against Orthomyxoviruses. Closely resembling canonical dynamins, MxA oligomerizes into higher order ring-like structures[65]. Both GTPase activity and oligomer formation are required for antiviral activity[79]. MxA is comprised of an N-terminal globular GTPase-containing head (G domain) and a C-terminal stalk, which are connected by a hinge-like bundle-signaling element (BSE)[67]. The L4 loop (L4), which protrudes from the MxA C-terminal stalk, is a hotspot for recurrent positive selection in primates[45,97]. In my thesis, I show that combinatorial mutagenesis of rapidly evolving sites can modify potency and specificity towards THOV and IAV.

Previous studies have identified MxA variants with enhanced antiviral activity against THOV and IAV[69,100]. This enhancing effect is attributed to re-localizing MxA to the nucleus or by an allelic variation in humans, V470G, in the stalk helix $\alpha 2^{S}$. However, these mutations seem to alter intrinsic properties in MxA and are not a feature of target recognition specificity. In contrast to our super restrictors, both of these mutations result in enhanced antiviral activity against both IAV and THOV. The localization into the nucleus results in a block of viral primary transcription, whereas the point mutant V470G shows a weaker MxA-MxA interaction. Given their distinct phenotypic effects, these two mutations are probably acting through different pathways. The specific mechanisms by which these mutations could be enhancing antiviral activity are not fully understood. Since we observe enhancement of antiviral activity in a target-dependent manner, we predict that L4 loop mutations are enhancing antiviral activity by a distinct mechanism. If the previously described mutations are acting through a different mechanism then the super restriction L4 loop mutations, the combination of the L4 loop super restrictor combinations and these other

mutations will have a synergistic effect on antiviral potency. Here, we investigated the molecular mechanisms that give rise to MxA variants with increased antiviral potency.

Results and Discussion

Comparison of known variants with enhanced antiviral activity

To evaluate if combining these mutations have a synergistic effect in enhancing antiviral activity, we added either a nuclear localization signal (NLS) or the V470G mutation in super restrictors YFVAG, RFVGV and TFRWR (Figure 15). The ability of these mutants to inhibit polymerase activity was tested in the THOV minigenome. Only the addition of a NLS-tag has a synergistic effect on antiviral activity, in particular for mutants RFVGV and YFVAG. This suggests that their super restriction is independent on localization and that the mechanism mediating enhanced antiviral activity is distinct from nuclear localization. The combination of L4 mutations and the V470G mutation did not result greater antiviral activity. This suggests that

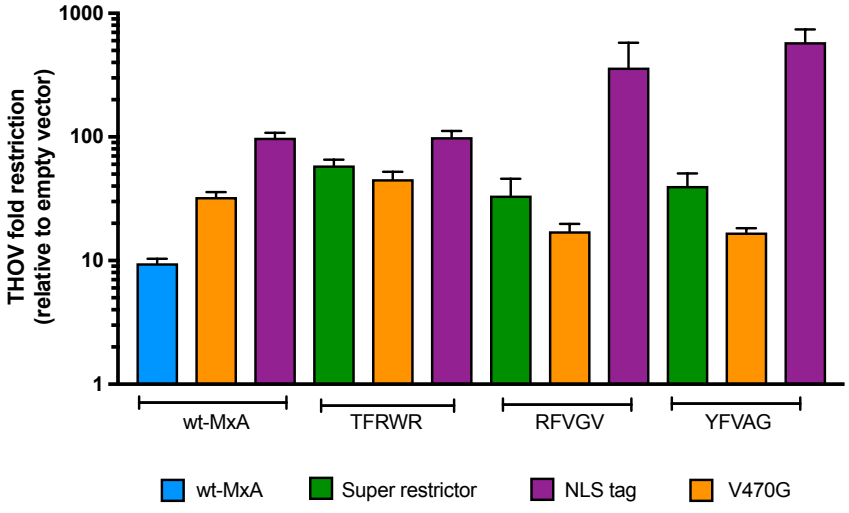


Figure 15. Additivity of restriction-enhancing mutations.

Super restrictor variants were modified by either adding the mutation V470G or an NLS tag. The restriction activity of the mutants was tested using the THOV minireplicon assay. The data is shown in fold-restriction relative to empty vector.

mutation V470G and the L4 loop mutations in the super restrictors are working through the same pathway.

Antiviral activity is stabilized in the nucleus

Computational models based on thermodynamic parameters suggest that a large number of amino acids substitutions could exponentially decrease the probability that such a mutant protein will retain its wild-type structure. Most of the variants in this study are four or more amino acids away from the parental sequence. Thus, it is possible that these variants have significant changes in structure and, therefore, stability. Since the target of MxA antiviral activity is the nucleoprotein (NP), we hypothesized that if less stable variants are re-localized to the nucleus, where NP is highly

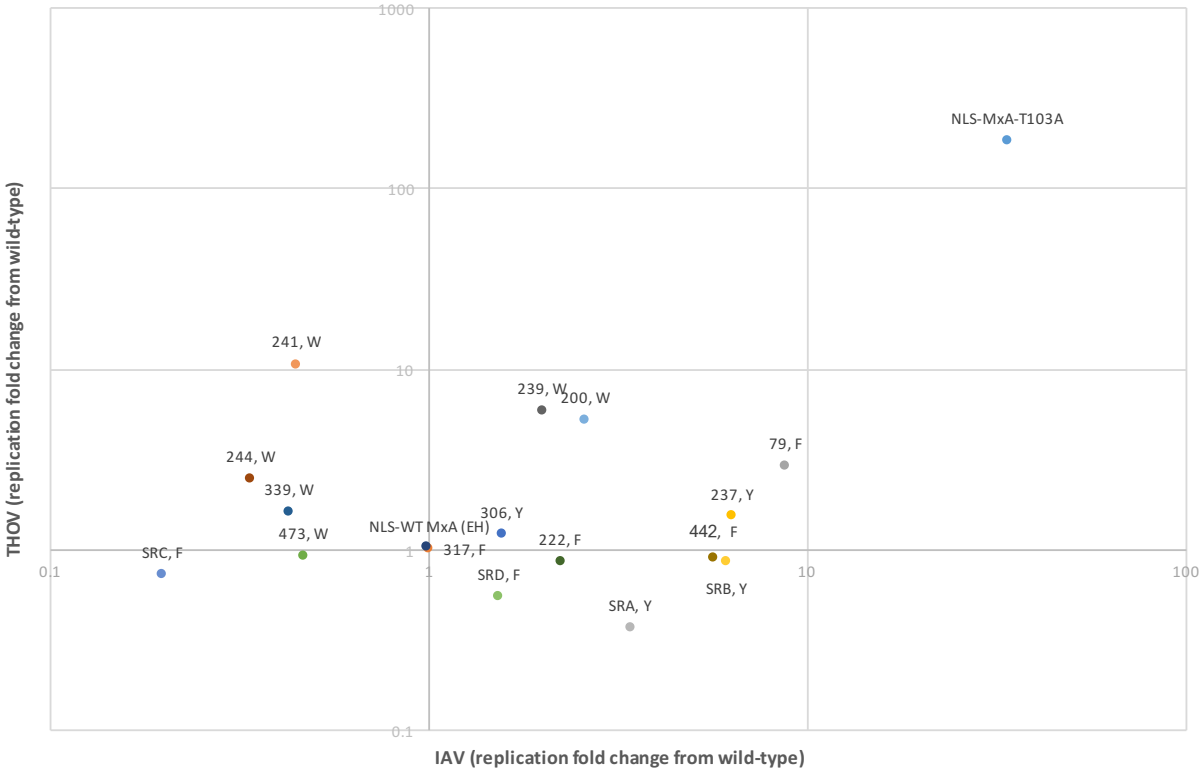


Figure 16. Fold change effect of NLS tag in L4 loop variants.

Fold change in restrictive activity from cytoplasm to nuclear re-localization normalized to wt-MxA. Each data point represents the fold change in activity for each variant pair (cytoplasmic vs. nuclear). The number next to each point represents the variant number. The letter represents the amino acid at site 561. THOV non-restrictor (F561W) have increased gain of restriction compared to wtMxA.

concentrated, they will have an increase in antiviral activity greater than the fold change observed when wild-type MxA is located in the nuclear compartment. To test this hypothesis, we selected representative variants from the non-restrictors, restrictors and super restrictors from Chapter 3 and added the nuclear localization signal (NLS) of the simian virus 40 (SV40) large T antigen to each one of these MxA variants. We then evaluated the antiviral activity of these mutants using the THOV minigenome assay. NLS tagged super restrictors and variants with wt-like antiviral activity had an enhancement effect of 10 to 20 –fold. Interestingly, non-restrictors had an increase in antiviral activity of over 60-fold in some cases (Figure 16). Taken together, this indicates that by re-localizing non-restrictors into the nucleus, antiviral activity is improved beyond that observed in wild-type MxA. This suggests that non-restrictors are misfolding or are not stable enough to target NP in the cytoplasm.

Chapter 5: Perspectives and future directions

Selective pressure exerted by viral pathogens can lead to relentless genetic innovation in mammalian innate immunity genes, whose protein products (restriction factors or antiviral proteins) act as the first line of defense against invading viral pathogens[5,6,18]. Studies of virus-host interactions have provided insight into the complex genetic adaptations in host genomes in response to rapidly evolving viruses. Well-studied examples of genetic innovation are the gene duplication and subsequent sub-functionalization of APOBEC3 proteins, the gene fusions in TRIM proteins, the epistatic binding regulations in PKR, and rapid evolution in most restriction factors described to date[6,106]. Despite all the innovative solutions to counteract viruses, host genomes might nevertheless seem at a disadvantage; viruses have a higher rate of evolution due to short generation times, high mutation rates and large population sizes. For instance, in a single round of RNA virus replication, it has been estimated that every point mutation is generated; this represents a remarkable array of genetic variation that for selection to act upon.

To counterbalance the seemingly insurmountable challenge imposed by viruses, the vertebrate immune system encodes for dozens of restriction factors that employ diverse mechanisms to halt viral replication and prevent cross-species transmission. In a somewhat redundant fashion, some viruses are targeted by multiple restriction factors that can act in a concerted fashion at different stages in their life cycle. This multifaceted attack could contribute to the host temporarily winning a host-virus conflict because adaptation to evade multiple restriction factors could compromise viral fitness. On the host side, there are multiple examples of restriction factors that counteract multiple viruses. In fact, several restriction factors have this capability that is exerted in two general approaches; restriction factors can be broad, and target a common feature in multiple viruses, or specific, and target a specific virus family[21].

The combination of breadth and specificity possess a unique challenge in host adaptation since preference for one target could lead to the loss of a pre-existing specificity against another. To offset this possibility, different functions in restriction factors could be segregated to different protein domains. This separation of functionalities could guarantee that adaptation to one virus would not affect activity against another. However, if multiple viruses are targeted by the same domain, the sequence space in which multiple viruses could be targeted is limited especially since most amino acid residues in restriction factors are highly conserved, with only a small percentage evolving rapidly. In this thesis, I explored the role of rapidly evolving amino acid residues at antiviral specificity domains with the goal of understanding how rapidly evolving sites can coordinate specificity and potency against multiple viruses.

I focused on the antiviral protein MxA, whose broad antiviral activity suggests that it has been involved in multiple genetic conflicts over evolutionary time. Indeed, MxA displays multiple hotpots of positive selection scattered throughout the protein. The loop L4 bears the most striking signature of positive selection with 18.6% of codons evolving under positive selection (average dN/dS value = 5.08)[93]. Functional experiments revealed that the loop L4 also acts as the antiviral specificity domain against Orthomyxoviruses THOV and IAV. Variation of amino acids in rapidly evolving sites explains the differences in antiviral activity among primate MxA orthologs against both viruses. This antiviral specificity and binding to NP is governed by a single amino acid F561 in loop L4. Moreover, grafting of the antivirally active (against THOV and IAV) human MxA loop L4 into the mouse Mx1 conferred both a gain of antiviral activity and binding to the THOV NP. Thus, the loop L4 is a key determinant of MxA interaction with NP proteins from *Orthomyxoviruses*[97].

The observation that a single residue determines the specificity against *Orthomyxoviruses* is intriguing since the MxA-NP surface is probably larger than a single residue. The determination of the role of site 561 was made by swapping amino acids from active and inactive MxA primate orthologs in which the rest of the rapidly evolving residues are nearly identical. Thus, the role of other rapidly evolving sites in loop L4 was not evaluated. In this thesis, we explored the role of all rapidly evolving sites in the loop L4 through a combinatorial mutagenesis screen and found that site 561 is indeed a strong specificity determinant element; all restrictors encode for an aromatic amino acid at residue 561. However, we find that non-561 sites play important modulatory roles in antiviral activity. Indeed, whereas my initial screen found that less than 5% of variants were capable of restricting THOV, a subsequent screen in which site 561 was fixed to the required phenylalanine resulted in a THOV in ~65% of the variants tested. Nevertheless, our finding that 35% of the variants were still worse than human MxA at restricting THOV despite possessing F561 reveals the potential for negative epistasis from the other rapidly evolving L4 residues. Thus, we found that simultaneous variation of amino acids at other rapidly evolving sites could result in either loss of restriction, gain of restriction (or super-restriction) and a specificity tradeoff.

My findings reveal that enhanced antiviral activity is genetically determined not just via a single residue but instead through a combination of amino acids, with site 540 being a strong determinant for super-restriction. I found evidence for epistatic interactions between rapidly evolving sites to coordinate potency. Only site 540, distal from the other four rapidly evolving sites, was highly sensitive to changes away from the wild type amino acid, glycine. This indicates that most rapidly evolving sites are tolerant of mutations. Interestingly, mutagenesis at conserved residues in loop L4 markedly negatively affect antiviral activity against THOV. The mutation tolerance of positively selected residues suggests that evolution has honed in on positions that alter

specificity and potency, but not at other sites that have a high probability of disrupting the binding to viral targets, or otherwise affect general MxA properties like oligomerization.

Wild type human MxA has the ability to restrict THOV and IAV. The increased restrictive capacity of anti-THOV MxA super-restrictors is coupled with increased avidity to THOV NP. However, the anti-THOV super-restrictors identified in this study, which are derived from human MxA, have lost or greatly reduced their ability to restrict IAV. Therefore, it is possible that by gaining avidity to THOV NP, the super restrictors have lost the ability to bind to IAV NP. Although a convincing MxA-IAV interaction remains to be demonstrated, it is presumed that direct physical binding is the basis for restriction as human-adapted IAV strains are capable of escape the antiviral effect of MxA through mutations in surface-exposed residues in NP[40,76,77]. My findings reveal that the acquisition of super restriction through mutations that result in a strong pathogen preference come at the cost of antiviral breadth. This specificity tradeoff could explain why super restrictor sequences are naturally rare. Even if super restrictors confer a short-term fitness advantage, these alleles will be lost in a population over evolutionary time as hosts encounter multiple viruses.

Our screen focused on THOV, which is highly sensitive to human MxA activity. However, IAV is an interesting target for MxA combinatorial screens. Human MxA restricts THOV 10-100 times better than the most MxA-sensitive influenza strain[107]. This indicates that MxA antiviral activity can be improved upon significantly, both for MxA-sensitive and MxA-resistant IAV strains. THOV and IAV are both targeted by the L4 loop, with site 561 having a necessary biochemical requirement for viral target recognition. Therefore, we propose that through combinatorial mutagenesis we can identify super restrictor variants of MxA-sensitive IAV strains. Since we observed a specificity tradeoff in THOV super restrictors, it is possible that IAV super restrictors also show a loss of THOV activity, supporting the specificity versus breadth tradeoff

argument. Screening combinatorial mutants using MxA-resistant strains determine if changes in rapidly evolving sites in the loop L4 are sufficient to re-establish viral target recognition. Thus, combinatorial mutants could be functionally assayed against multiple IAV strains to better understand how MxA develops specificity to closely related strains.

We hypothesize that MxA has been selected to act as a generalist sequence to target multiple viruses simultaneously. It is possible that breadth is a result of a frequency of viral infections that have imposed a selective pressure for breadth rather than specificity. The adaptive cost of selection for specificity to one virus will result in loss of some preexisting specificity to another making the occurrence of super restrictors rare, and generalist sequences the solution for adaptation to breadth. We propose that our evolution-guided approach can discover novel super-restrictor versions of other Mx proteins that target important human pathogens for potential therapeutic applications.

A potential target for combinatorial mutagenesis is the broadly-acting antiviral protein bat Mx1, recently shown to have antiviral activity against significant human viral pathogens such as Ebola and influenza viruses. Bat Mx1 is evolving under positive selection and this amino acid diversity that might explain species-specific antiviral activities of bat Mx1 orthologues (Figure 18). In bat Mx1, the L4 loop harbors most of the sites under positive selection and it is possible

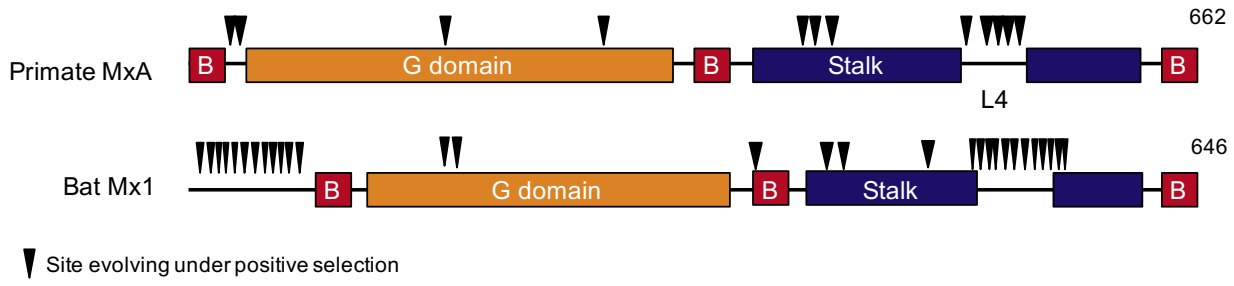


Figure 17. Positive selection in bat and primate Mx1 proteins.

Illustration of primary structure of bat Mx1 and primate MxA: The unstructured regions (black), the bundle signaling elements (B, red), the G domain (orange), the stalk (blue), and loop 4 (L4).

that it might play a similar role in target recognition as primate MxA. Therefore, we propose that combinatorial mutagenesis in the L4 loop of bat Mx1 could reveal more potent anti-ebola or anti-influenza variants for potential therapeutic use.

Although Mx proteins are unusual in its antiviral breadth, many antiviral proteins evolve under similar selective regimes to Mx, in which they undergo recurrent positive selection in response to viral pathogens[6]. In many of these cases (*e.g.,* TRIM5, MxB), positive selection appears enriched in unstructured loops like MxA L4[26,104,108]. We speculate that unstructured loops like L4 represent an ensemble of many conformations that provide MxA with structural and evolutionary flexibility to adapt to binding distinct viral targets, whereas super-restricting variants may be less structurally flexible, trading off increased restriction of certain viral targets with decreased antiviral range. Thus, breadth-specificity tradeoffs constrain the adaptive landscape of antiviral proteins. We propose that our evolution-guided approach can increase the potency of known antiviral genes to target viruses that are currently poorly targeted because of evolutionary restraints acting on antiviral genes like MxA to maintain a broad range of antiviral specificities.

Appendix 1: *Flow cytometry-based combinatorial screen with MxA-sensitive or –resistant GFP IAV strains*

Introduction

The antiviral activity of MxA relies in the recognition of the viral NP in IAV. The origin of NP (human or avian) is a major determinant for MxA sensitivity[70]. By interacting with NP, MxA is thought to prevent the nuclear entry of the influenza RNA genome into the nucleus for transcription. IAVs must therefore evolve and maintain human MxA resistance to successfully infect human populations. Indeed, MxA resistance mutations in NP are very rare in avian IAVs (>1%) whereas they are found at very high frequencies (>98%) in human IAV isolates[40]. MxA resistant mutations are also present in the NP of pandemic viruses that emerged from bird-to-human zoonotic transmissions (e.g., 1918, 2009 IAVs and the 2013 IAV outbreak in China). Taken together, this suggests that MxA is a barrier to cross-species infection.

Previous studies have shown that MxA uses the L4 loop to target the viral ribonucleoprotein complexes present in IAVs, which consist of genomic RNA coated by the viral nucleoprotein (NP) and the associated polymerase complex[109]. Variation in primate MxA proteins also results in dramatic variation in their ability to bind IAV NP and mediate restriction of IAVs. Indeed, single amino acid changes in L4 were able to confer antiviral activity onto previously ineffective MxA variants. Furthermore, we find that single amino acid substitutions in rapidly evolving sites at the L4 loop can confer antiviral activity against IAV. Interestingly, the L4 loop is enriched in positive selected sites. In this series of experiments, we aimed to determine the functional effect of amino acid variation at positive selected sites in L4 by developing a high-throughput screen using GFP-expressing influenza viruses with different degrees of MxA sensitivity (Figure 18). The following experiments describe pilot experiments to develop tools for high-throughput screening of the MxA L4 loop library.

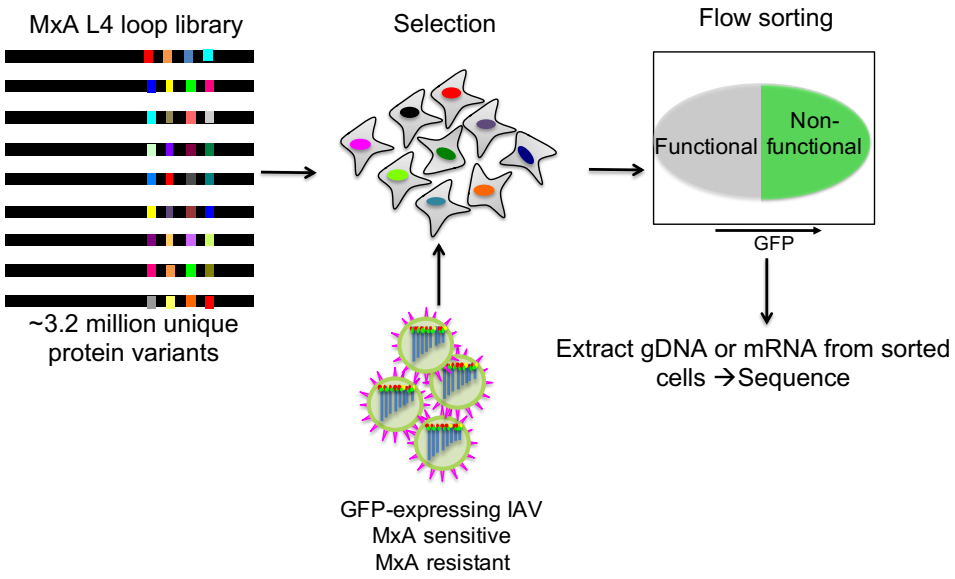


Figure 18. High-throughput functional screen of MxA L4 loop variants.

A mammalian cell line will be infected with a virus library encoding for MxA L4 loop variants. The infection will be performed at a MOI that restricts the expression of MxA to one per cell. This cell library will be infected with GFP-expressing viruses that are sensitive or resistant to wtMxA. After 24hrs of incubation, the infected cells will be sorted based on GFP expression. gDNA or mRNA will be extracted from the GFP-negative and GFP-positive cells and sequenced for depletion or enrichment of MxA L4 loop variants.

Results and discussion

Generation of MxA-sensitive or resistant GFP viruses

In order to evaluate the functionality of variants in the MxA L4 loop library, we built MxA -sensitive or -resistant viruses in the same genetic background (H1N1/A/WSN/1933). To achieve the different degrees of MxA sensitivity or resistance, we modified the nucleoprotein in a GFP-expressing WSN33 influenza virus[77]. Previous mapping of MxA-sensitizing mutations identified three critical mutations in avian NP; I/V100R, P283L and Y313F that underlie sensitivity to MxA. Individually, these mutations provide partial sensitivity to MxA, but the full effect can be appreciated when all three residues are mutated to the sensitizing residues[40]. We attempted to recapitulate these phenotypes by engineering the NP MxA-sensitizing mutations in the

H1N1/WSN33 strain. In its wild type form, WSN33 encodes for one of the sensitizing mutations, 283L (1S), and are partially resistant to MxA (Figure 17A). When we mutate position 283 in the WSN33 strain to a proline (0S), full MxA resistance is gained (Figure 17B). However, when we introduce two sensitizing mutations, either 100R/283L or 283L/313F, we do not observe an increase in sensitivity to MxA activity (Figure 17C and D). Therefore, the lack of dynamic range between WT WSN33 and 1S/2S mutants is not wide enough to use it as a selective tool for the high-throughput screen. We were not successful in building a 3S virus due to a severe fitness effect when these three mutations are combined in the GFP viruses. Recent studies have showed that MxA-sensitizing mutations lead to a defect in vRNP translocation into the nucleus, which may explain our the loss of fitness of the 1S and 2S mutants and the failure to produce 3S viruses [110].

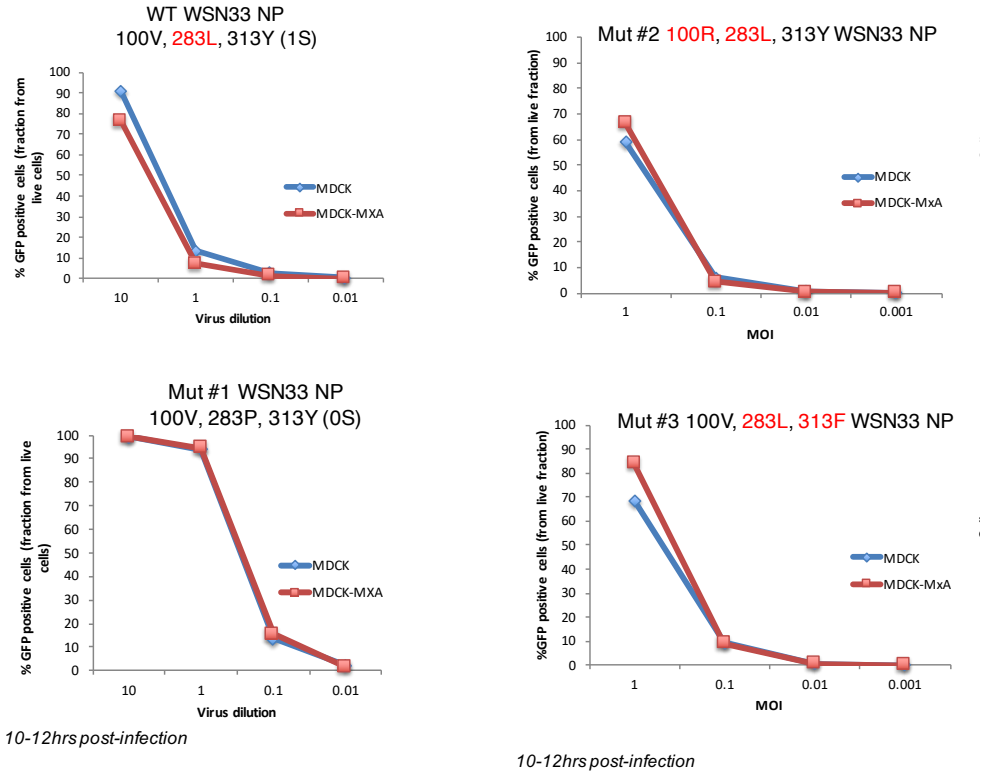


Figure 19. MxA sensitivity of GFP-expressing WSN33 virus harboring mutations in NP.

MDCK-MxA or MDCK cells were infected with serial viral dilutions of GFP-expressing WSN33 virus encoding for mutations in NP. After 12 hours of infection, cells were fixed and GFP expression levels were quantified through flow cytometry.

Assay development to detect the effect on output virus

The seemingly lack of restriction by MxA of the NP mutant viruses could be due to MxA blocking a step after primary transcription of all virus segments, including the GFP reporter segment. If MxA restricts a later step after the transcription of the GFP segment, we would be unable to detect a significant decrease in GFP levels due to MxA action. However, if a later stage is targeted, we should be able to detect a stronger effect at the level of output virus. To test this hypothesis, we designed an assay to detect viruses in the supernatant of infected cells that have been co-transfected with plasmids encoding for HA (the segment replaced by GFP and provided in trans during virus production) and MxA.

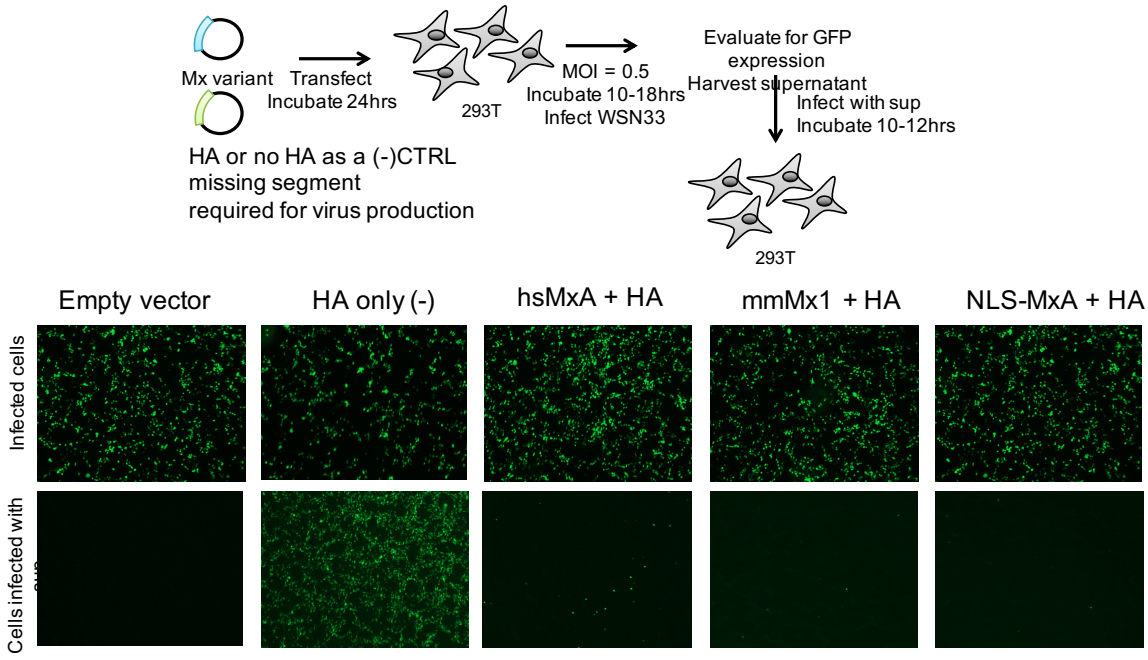


Figure 20. Assay to detect production of infectious GFP viruses in MxA-expressing cells.

293T cells are co-transfected with Mx and HA encoding plasmids (No HA is added to the transfection as a negative control and HA only is added as a positive control). Transfected cells are then incubated for 24 hours to allow for protein expression. Cells are then infected and incubated for 10-18 hours. Supernatants are harvested and cleared using a syringe filter. The cleared supernatants are used to infect 293T cells to detect viral production.

Co-transfecting HA and subsequently infecting with GFP viruses, allows for replication and production of infectious viruses (see second column in image panel in figure 17). However, if MxA is blocking viral replication, we will not detect viruses in produced in HA expressing cells. We observe that human MxA has a strong effect in output virus as the cells co-transfected with HA and wtMxA (hsMxA) do not produce infectious viruses. This suggests that MxA blocks a step after primary transcription that severely affects virus assembly and budding. Although the dynamic range in this assay is wide, the assay is difficult to perform in a high-throughput format. Taken together, these pilots experiments show that introducing MxA-sensitizing mutations in WSN33 NP has fitness cost that leads to lower titers and, in some cases, failure to form infectious particles (Figure 20). An alternative to this approach will be to utilize IAV strains that have naturally developed MxA-sensitivity and exhibit robust replication in tissue culture. Recent H5N1-Luc-expressing viruses could be a suitable option as these strains are sensitive to MxA restriction and encode for a reporter to allow for viral replication surveillance[111]. Our experiments detecting produced virus in MxA-expressing cells indicate that WSN33 is sensitive to MxA restriction, but that the restrictive effects of MxA are observed at later stages (i.e. viral assembly or budding), rather than at primary transcription. These findings are congruent with earlier studies that determined the MxA does not inhibit primary transcription, but rather a later step in the viral life cycle[52]. Our results highlight the need for a robust IAV infectivity assay that captures the step in the viral life cycle that MxA affects in a platform that couples genotype to phenotype for high-throughput screenings.

Bibliography

1. Shi, M. *et al.* The evolutionary history of vertebrate RNA viruses. *Nature* **556,** 197–202 (2018).
2. Emerman, M. & Malik, H. S. Paleovirology - Modern consequences of ancient viruses. *PLoS Biology* **8,** (2010).
3. Katzourakis, A. Editorial overview: Paleovirology: the genomic fossil record, and consequences of ancient viral infections. *Current Opinion in Virology* **25,** ix–xi (2017).
4. Holmes, E. C. The evolution of endogenous viral elements. *Cell Host and Microbe* **10,** 368–377 (2011).
5. Duggal, N. K. & Emerman, M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nat. Rev. Immunol.* **12,** 687–95 (2012).
6. Daugherty, M. D. & Malik, H. S. Rules of Engagement: Molecular Insights from Host-Virus Arms Races. *Annu. Rev. Genet.* **46,** 677–700 (2012).
7. Le Bon, A. & Tough, D. Links between innate and adaptive immunity via type I interferon. *Curr. Opin. Immunol.* **14,** 432–436 (2002).
8. Iwasaki, A. & Medzhitov, R. Control of adaptive immunity by the innate immune system. *Nature Immunology* **16,** 343–353 (2015).
9. Damania, B. & Blackbourn, D. J. Innate barriers to viral infection. in *Future Microbiology* **7,** 815–822 (2012).
10. Kawai, T. & Akira, S. Innate immune recognition of viral infection. *Nat. Immunol.* **7,** 131–137 (2006).
11. Riera Romo, M., Pérez-Martínez, D. & Castillo Ferrer, C. Innate immunity in vertebrates: An overview. *Immunology* **148,** 125–139 (2016).

12. Schoggins, J. W. & Rice, C. M. Interferon-stimulated genes and their antiviral effector functions. *Curr. Opin. Virol.* **1,** 519–525 (2011).

13. Theofilopoulos, A. N., Kono, D. H., Beutler, B. & Baccala, R. Intracellular Nucleic Acid Sensors and Autoimmunity. *J. Interf. Cytokine Res.* **31,** 867–886 (2011).

14. Ivashkiv, L. B. & Donlin, L. T. Regulation of type I interferon responses. *Nat. Rev. Immunol.* **14,** 36–49 (2013).

15. Levy, D. & Darnell, J. J. STATs: transcriptional control and biological impact. *Nat. Rev. Mol. Cell Biol.* **3,** 651–662 (2002).

16. Stark, G. & Darnell, J. J. The JAK-STAT pathway at twenty. *Immunity* **36,** 503–514 (2012).

17. MacMicking, J. D. Interferon-inducible effector mechanisms in cell-autonomous immunity. *Nature Reviews Immunology* **12,** 367–382 (2012).

18. Kluge, S. F., Sauter, D. & Kirchhoff, F. SnapShot: Antiviral Restriction Factors. *Cell* **163,** 774–774e1 (2015).

19. Malim, M. H. & Bieniasz, P. D. HIV restriction factors and mechanisms of evasion. *Cold Spring Harb. Perspect. Med.* **2,** (2012).

20. Hefti, H. P. *et al.* Human MxA protein protects mice lacking a functional alpha/beta interferon system against La crosse virus and other lethal viral infections. *J. Virol.* **73,** 6984–6991 (1999).

21. Mitchell, P. S., Emerman, M. & Malik, H. S. An evolutionary perspective on the broad antiviral specificity of MxA. *Current Opinion in Microbiology* **16,** 493–499 (2013).

22. Bailey, C. C., Zhong, G., Huang, I.-C. & Farzan, M. IFITM-Family Proteins: The Cell's First Line of Antiviral Defense. *Annu. Rev. Virol.* **1,** 261–283 (2014).

23. Yu, J. *et al.* IFITM Proteins Restrict HIV-1 Infection by Antagonizing the Envelope Glycoprotein. *Cell Rep.* **13,** 145–156 (2015).

24. Bierle, C. J., Semmens, K. M. & Geballe, A. P. Double-stranded RNA binding by the human cytomegalovirus PKR antagonist TRS1. *Virology* **442,** 28–37 (2013).

25. Dauber, B. & Wolff, T. Activation of the antiviral kinase PKR and viral countermeasures. *Viruses* **1,** 523–544 (2009).

26. Grütter, M. G. & Luban, J. TRIM5 structure, HIV-1 capsid recognition, and innate immune signaling. *Curr. Opin. Virol.* **2,** 142–150 (2012).

27. Pertel, T. *et al.* TRIM5 is an innate immune sensor for the retrovirus capsid lattice. *Nature* **472,** 361–365 (2011).

28. Lindenmann, J. Resistance of mice to mouse-adapted influenza A virus. *Virology* **16,** 203–204 (1962).

29. Haller, O., Arnheiter, H., Pavlovic, J. & Staeheli, P. The Discovery of the Antiviral Resistance Gene Mx: A Story of Great Ideas, Great Failures, and Some Success. *Annu. Rev. Virol.* 1–19 (2018). doi:10.1146/annurev-virology-092917-043525

30. Haller, O., Arnheiter, H., Lindenmann, J. & Gresser, I. Host gene influences sensitivity to interferon action selectively for influenza virus [10]. *Nature* **283,** 660–662 (1980).

31. Staeheli, P. *et al.* Interferon-regulated influenza virus resistance gene Mx is localized on mouse chromosome 16. *J. Virol.* **58,** 967–9 (1986).

32. Staeheli, P., Haller, O., Boll, W., Lindenmann, J. & Weissmann, C. Mx protein: Constitutive expression in 3T3 cells transformed with cloned Mx cDNA confers selective resistance to influenza virus. *Cell* **44,** 147–158 (1986).

33. Haller, O., Arnheiter, H. & Lindenmann, J. Genetically determined resistance to infection

by hepatotropic influenza A virus in mice: effect of immunosuppression. *Infect Immun* **13,** 844–854 (1976).

34. Haller, O., Arnheiter, H., Pavlovic, J. & Staeheli, P. The Discovery of the Antiviral Resistance Gene Mx: A Story of Great Ideas, Great Failures, and Some Success. *Annu. Rev. Virol.* 1–19 (2018). doi:10.1146/annurev-virology-092917-043525

35. Staeheli, P. & Haller, O. Interferon-induced human protein with homology to protein Mx of influenza virus-resistant mice. *Mol. Cell. Biol.* **5,** 2150–3 (1985).

36. Liu, Z. *et al.* The interferon-inducible MxB protein inhibits HIV-1 infection. *Cell Host Microbe* **14,** 398–410 (2013).

37. Crameri, M. *et al.* MxB is an interferon-induced restriction factor of human herpesviruses. *Nat. Commun.* **9,** (2018).

38. Haller, O. Dynamins are forever: MxB inhibits HIV-1. *Cell Host and Microbe* **14,** 371–373 (2013).

39. Deeg, C. M. *et al.* In vivo evasion of MxA by avian influenza viruses requires human signature in the viral nucleoprotein. *J. Exp. Med.* **214,** 1239–1248 (2017).

40. Mänz, B. *et al.* Pandemic Influenza A Viruses Escape from Restriction by Human MxA through Adaptive Mutations in the Nucleoprotein. *PLoS Pathog.* **9,** (2013).

41. Mitchell, P. S., Young, J. M., Emerman, M. & Malik, H. S. Evolutionary Analyses Suggest a Function of MxB Immunity Proteins Beyond Lentivirus Restriction. *PLoS Pathog.* **11,** (2015).

42. Braun, B. a., Marcovitz, A., Camp, J. G., Jia, R. & Bejerano, G. Mx1 and Mx2 key antiviral proteins are surprisingly lost in toothed whales. *Proc. Natl. Acad. Sci.* **112,** 8036–8040 (2015).

43. Beineke, A., Siebert, U., Wohlsein, P. & Baumgärtner, W. Immunology of whales and dolphins. *Veterinary Immunology and Immunopathology* **133,** 81–94 (2010).

44. Fuchs, J. *et al.* Evolution and Antiviral Specificities of Interferon-Induced Mx Proteins of Bats against Ebola, Influenza, and Other RNA Viruses. *J. Virol.* **91,** e00361-17 (2017).

45. Mitchell, P. S., Emerman, M. & Malik, H. S. An evolutionary perspective on the broad antiviral specificity of MxA. *Current Opinion in Microbiology* **16,** 493–499 (2013).

46. King, M. C., Raposo, G. & Lemmon, M. A. Inhibition of nuclear import and cell-cycle progression by mutated forms of the dynamin-like GTPase MxB. *Proc. Natl. Acad. Sci. U. S. A.* **101,** 8957–62 (2004).

47. Melen, K. *et al.* Interferon-induced Mx proteins form oligomers and contain a putative leucine zipper. *J. Biol. Chem.* (1992).

48. Melén, K. *et al.* Human MxB protein, an interferon-α-inducible GTPase, contains a nuclear targeting signal and is localized in the heterochromatin region beneath the nuclear envelope. *J. Biol. Chem.* **271,** 23478–23486 (1996).

49. Staeheli, P. & Haller, O. Human MX2/MxB: a potent interferon-induced post-entry inhibitor of herpesviruses and HIV-1. *J. Virol.* JVI.00709-18 (2018). doi:10.1128/JVI.00709-18

50. Goujon, C. *et al.* Transfer of the Amino-Terminal Nuclear Envelope Targeting Domain of Human MX2 Converts MX1 into an HIV-1 Resistance Factor. *J. Virol.* **88,** 9017–9026 (2014).

51. Haller, O., Staeheli, P., Schwemmle, M. & Kochs, G. Mx GTPases: Dynamin-like antiviral machines of innate immunity. *Trends in Microbiology* **23,** 154–163 (2015).

52. Pavlovic, J., Haller, O. & Staeheli, P. Human and mouse Mx proteins inhibit different

steps of the influenza virus multiplication cycle. *J. Virol.* **66,** 2564–9 (1992).

53. Haller, O., Frese, M., Rost, D., Nuttall, P. A. & Kochs, G. Tick-borne thogoto virus infection in mice is inhibited by the orthomyxovirus resistance gene product Mx1. *J Virol* **69,** 2596–2601 (1995).

54. Frese, M., Kochs, G., Feldmann, H., Hertkorn, C. & Haller, O. Inhibition of bunyaviruses, phleboviruses, and hantaviruses by human MxA protein. *J. Virol.* **70,** 915–23 (1996).

55. Haller, O. & Kochs, G. Interferon-induced Mx proteins: Dynamin-like GTPases with antiviral activity. *Traffic* (2002). doi:10.1034/j.1600-0854.2002.31003.x

56. Praefcke, G. J. K. & McMahon, H. T. The dynamin superfamily: universal membrane tubulation and fission molecules? *Nat. Rev. Mol. Cell Biol.* **5,** 133–47 (2004).

57. Haller, O., Stertz, S. & Kochs, G. The Mx GTPase family of interferon-induced antiviral proteins. *Microbes Infect.* **9,** 1636–1643 (2007).

58. Von Der Malsburg, A., Abutbul-Ionita, I., Haller, O., Kochs, G. & Danino, D. Stalk domain of the dynamin-like MxA GTPase protein mediates membrane binding and liposome tubulation via the unstructured L4 loop. *J. Biol. Chem.* **286,** 37858–37865 (2011).

59. Kochs, G., Haener, M., Aebi, U. & Haller, O. Self-assembly of human MxA GTPase into highly ordered dynamin-like oligomers. *J. Biol. Chem.* **277,** 14172–14176 (2002).

60. Nakayama, M. *et al.* Structure of mouse Mx1 protein: Molecular assembly and GTP-dependent conformational change. *J. Biol. Chem.* **268,** 15033–15038 (1993).

61. Accola, M. A., Huang, B., Masri, A. Al & McNiven, M. A. The antiviral dynamin family member, MxA, tubulates lipids and localizes to the smooth endoplasmic reticulum. *J. Biol. Chem.* **277,** 21829–21835 (2002).

62. Gao, S. *et al.* Structure of Myxovirus Resistance Protein A Reveals Intra- and Intermolecular Domain Interactions Required for the Antiviral Function. *Immunity* **35,** 514–525 (2011).

63. Gao, S. *et al.* Structural basis of oligomerization in the stalk region of dynamin-like MxA. *Nature* **465,** 502–506 (2010).

64. Nakade, K., Handa, H. & Nagata, K. Promoter structure of the MxA gene that confers resistance to influenza virus. *FEBS Lett.* **418,** 315–318 (1997).

65. Daumke, O., Gao, S., von der Malsburg, A., Haller, O. & Kochs, G. Structure of the MxA stalk elucidates the assembly of ring-like units of an antiviral module. *Small GTPases* **1,** 62–64 (2010).

66. Rennie, M. L., McKelvie, S. A., Bulloch, E. M. M. & Kingston, R. L. Transient dimerization of human MxA promotes GTP hydrolysis, resulting in a mechanical power stroke. *Structure* **22,** 1433–1445 (2014).

67. Verhelst, J., Hulpiau, P. & Saelens, X. Mx proteins: antiviral gatekeepers that restrain the uninvited. *Microbiol. Mol. Biol. Rev.* **77,** 551–66 (2013).

68. Holzinger, D. *et al.* Induction of MxA gene expression by influenza A virus requires type I or type III interferon signaling. *J. Virol.* **81,** 7776–7785 (2007).

69. Turan, K. *et al.* Nuclear MxA proteins form a complex with influenza virus NP and inhibit the transcription of the engineered influenza virus genome. *Nucleic Acids Res.* **32,** 643–652 (2004).

70. Zimmermann, P., Manz, B., Haller, O., Schwemmle, M. & Kochs, G. The viral nucleoprotein determines Mx sensitivity of influenza A viruses. *J Virol* **85,** 8133–8140 (2011).

71. Gordien, E. *et al.* Inhibition of hepatitis B virus replication by the interferon-inducible MxA protein. *J. Virol.* **75,** 2684–91 (2001).

72. Netherton, C. L. *et al.* Inhibition of a large double-stranded DNA virus by MxA protein. *J. Virol.* **83,** 2310–20 (2009).

73. King, A. M. Q., Adams, M. J., Carstens, E. B. & Lefkowitz, E. J. *Virus Taxonomy: classification and nomenclature of viruses. Elsevier Academic Press* (2012). doi:ISBN 9780123846846

74. Mostafa, A., Abdelwhab, E. M., Mettenleiter, T. C. & Pleschka, S. Zoonotic potential of influenza A viruses: A comprehensive overview. *Viruses* **10,** (2018).

75. Kosoy, O. I. *et al.* Novel Thogotovirus associated with febrile illness and death, united states, 2014. *Emerg. Infect. Dis.* **21,** 760–764 (2015).

76. Riegger, D. *et al.* The Nucleoprotein of Newly Emerged H7N9 Influenza A Virus Harbors a Unique Motif Conferring Resistance to Antiviral Human MxA. *J. Virol.* **89,** 2241–2252 (2015).

77. Ashenberg, O., Padmakumar, J., Doud, M. B. & Bloom, J. D. Deep mutational scanning identifies sites in influenza nucleoprotein that affect viral inhibition by MxA. *PLoS Pathog.* **13,** (2017).

78. Gotz, V. *et al.* Influenza A viruses escape from MxA restriction at the expense of efficient nuclear vRNP import. *Sci Rep* **6,** 23138 (2016).

79. Nigg, P. E. & Pavlovic, J. Oligomerization and GTP-binding requirements of MxA for viral target recognition and antiviral activity against influenza A virus. *J. Biol. Chem.* **290,** 29893–29906 (2015).

80. Gibbs, R. A. *et al.* Evolutionary and biomedical insights from the rhesus macaque

genome. *Science* **316,** 222–234 (2007).

81. Kosiol, C. *et al.* Patterns of positive selection in six mammalian genomes. *PLoS Genet.* **4,** (2008).

82. Lindblad-Toh, K. *et al.* A high-resolution map of human evolutionary constraint using 29 mammals. *Nature* **478,** 476–482 (2011).

83. and Analysis Consortium, T. C. S. Initial sequence of the chimpanzee genome and comparison with the human genome. *Nature* **437,** 69–87 (2005).

84. Kosakovsky Pond, S. L., Frost, S. D. W. & Muse, S. V. HyPhy: Hypothesis testing using phylogenies. *Bioinformatics* **21,** 676–679 (2005).

85. Yang, Z. PAML 4: Phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* **24,** 1586–1591 (2007).

86. McBee, R. M., Rozmiarek, S. A., Meyerson, N. R., Rowley, P. A. & Sawyer, S. L. The effect of species representation on the detection of positive selection in primate gene data sets. *Mol. Biol. Evol.* **32,** 1091–1096 (2015).

87. Sawyer, S. L., Wu, L. I., Emerman, M. & Malik, H. S. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. *Proc. Natl. Acad. Sci. U. S. A.* **102,** 2832–7 (2005).

88. Stremlau, M., Perron, M., Welikala, S. & Sodroski, J. Species-Specific Variation in the B30 . 2 (SPRY) Domain of TRIM5 ␣ Determines the Potency of Human Immunodeficiency Virus Restriction. *Society* **79,** 3139–3145 (2005).

89. Carpentier, K. S., Esparo, N. M., Child, S. J. & Geballe, A. P. A Single Amino Acid Dictates Protein Kinase R Susceptibility to Unrelated Viral Antagonists. *PLoS Pathog.* **12,** (2016).

90. Yap, M. W., Nisole, S. & Stoye, J. P. A single amino acid change in the SPRY domain of human Trim5alpha leads to HIV-1 restriction. *Curr. Biol.* **15,** 73–8 (2005).

91. Compton, A. A., Hirsch, V. M. & Emerman, M. The host restriction factor APOBEC3G and retroviral Vif protein coevolve due to ongoing genetic conflict. *Cell Host Microbe* **11,** 91–98 (2012).

92. Price, A. J. *et al.* Active site remodeling switches HIV specificity of antiretroviral TRIMCyp. *Nat. Struct. Mol. Biol.* **16,** 1036–1042 (2009).

93. Mitchell, P. S. *et al.* Evolution-guided identification of antiviral specificity determinants in the broadly acting interferon-induced innate immunity factor MxA. *Cell Host Microbe* **12,** 598–604 (2012).

94. Fowler, D. M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat. Methods* **11,** 801–7 (2014).

95. Romero, P. A. & Arnold, F. H. Exploring protein fitness landscapes by directed evolution. *Nat. Rev. Mol. Cell Biol.* **10,** 866–76 (2009).

96. Nettling, M. *et al.* DiffLogo: A comparative visualization of sequence motifs. *BMC Bioinformatics* **16,** (2015).

97. Mitchell, P. S. *et al.* Evolution-guided identification of antiviral specificity determinants in the broadly acting interferon-induced innate immunity factor MxA. *Cell Host Microbe* **12,** 598–604 (2012).

98. Lim, E. S. *et al.* The ability of primate lentiviruses to degrade the monocyte restriction factor SAMHD1 preceded the birth of the viral accessory protein Vpx. *Cell Host Microbe* **11,** 194–204 (2012).

99. Weber, F., Haller, O. & Kochs, G. MxA GTPase blocks reporter gene expression of

reconstituted Thogoto virus ribonucleoprotein complexes. *J. Virol.* **74,** 560–563 (2000).

100. Graf, L. *et al.* Effects of allelic variations in the human myxovirus resistance protein A on its antiviral activity. *J. Biol. Chem.* **293,** 3056–3072 (2018).

101. Wang, J. D., Herman, C., Tipton, K. A., Gross, C. A. & Weissman, J. S. Directed evolution of substrate-optimized GroEL/S chaperonins. *Cell* (2002). doi:10.1016/S0092-8674(02)01198-4

102. Li, X., Baudry, J., Berenbaum, M. R. & Schuler, M. A. Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. *Proc. Natl. Acad. Sci.* (2004). doi:10.1073/pnas.0308691101

103. Farnham, G. & Baulcombe, D. C. Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. *Proc. Natl. Acad. Sci.* (2006). doi:10.1073/pnas.0605777103

104. Biris, N. *et al.* Structure of the rhesus monkey TRIM5 PRYSPRY domain, the HIV capsid recognition module. *Proc. Natl. Acad. Sci.* **109,** 13278–13283 (2012).

105. Fribourgh, J. L. *et al.* Structural insight into HIV-1 restriction by MxB. *Cell Host Microbe* (2014). doi:10.1016/j.chom.2014.09.021

106. Duggal, N. K. & Emerman, M. Evolutionary conflicts between viruses and restriction factors shape immunity. *Nature Reviews Immunology* **12,** 687–695 (2012).

107. Deeg, C. M. *et al.* In vivo evasion of MxA by avian influenza viruses requires human signature in the viral nucleoprotein. *J. Exp. Med.* **214,** 1239–1248 (2017).

108. Alvarez, F. J. D. *et al.* CryoEM structure of MxB reveals a novel oligomerization interface critical for HIV restriction. *Sci. Adv.* (2017). doi:10.1126/sciadv.1701264

109. Eisfeld, A. J., Neumann, G. & Kawaoka, Y. At the centre: influenza A virus

ribonucleoproteins. *Nat. Rev. Microbiol.* **13,** 28–41 (2014).

110. Gotz, V. *et al.* Influenza A viruses escape from MxA restriction at the expense of efficient nuclear vRNP import. *Sci. Rep.* **6,** 23138 (2016).

111. Dornfeld, D. *et al.* SMARCA2-regulated host cell factors are required for MxA restriction of influenza A viruses. *Sci. Rep.* **8,** (2018).