Evolution-guided design of super-restrictor antiviral proteins reveals a breadth-versus-specificity tradeoff

Rossana Colón-Thillet[1,4], Emily Hsieh[1,2,4,*], Laura Graf[6,*], Richard N McLaughlin Jr[1,5,§], Georg Kochs[6,7], Michael Emerman[1,2], Harmit S. Malik[1,3,8]

[1]Division of Basic Sciences, [2]Division of Human Biology, and [3]Howard Hughes Medical Institute, Fred Hutchinson Cancer Research Center, Seattle, WA, USA; [4]Molecular and Cellular Biology Graduate Program, University of Washington, Seattle, WA, USA; [5]Pacific Northwest Research Institute, Seattle WA, USA; [6]Institute of Virology, Medical Center-University of Freiburg; [7]Faculty of Medicine, University of Freiburg, Freiburg, Germany

*These authors contributed equally to this study

§Current address: Pacific Northwest Research Institute, Seattle, WA

[8]Correspondence should be addressed to:

Harmit S. Malik, 1100 Fairview Avenue N. A2-025, Seattle WA 98109; Tel: +001 (206) 667-5204; Fax: +001 (206) 667-6522; hsmalik@fhcrc.org

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.

Introduction

The innate arm of mammalian immunity includes dozens of antiviral proteins that act cell-autonomously to block viral replication [1, 2]. Over long evolutionary periods, innate immune proteins must adapt to combat rapidly-evolving pathogenic viruses. This adaptation often occurs at the interfaces between viral and host immune proteins, to either evade or increase binding interactions [3]. This back-and-forth evolution between host and viral proteins can result in signatures of positive selection (higher rates of non-synonymous substitutions, dN, compared to synonymous substitutions, dS) that can be identified by a comparison of orthologous genes over a phylogenetic tree. Positive selection is often concentrated in just a few residues that maximally impact binding affinity and thereby dictate the outcomes of viral infection in host cells [3]. Indeed, experimental tests show that variation in positively selected sites can explain the species-specific differences in antiviral specificity between orthologous antiviral proteins [3-5].

Such retrospective analyses are a powerful means of identifying amino acids and protein-protein interaction interfaces that have been shaped by past episodes of selection. Nonetheless, they are limited in terms of describing the available landscape of possible adaptation against viruses. For example, epistasis between different rapidly evolving sites could have shaped and constrained the available landscape of possible adaptation against viruses [6-8]. We sought to leverage the previous insights gained from our retrospective analyses to design a prospective

means to elicit higher-specificity versions of antiviral proteins and understand the available paths to adaptation of a host innate immune protein to potential viral pathogens.

We focused our studies on MxA, an interferon-induced large dynamin-like GTPase, which inhibits a diverse range of viruses by interacting with multiple distinct viral proteins [9, 10]. 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) [11, 12]. Previously, we had identified the L4 loop (L4), which protrudes from the MxA C-terminal stalk, as a hotspot for recurrent positive selection in primates [4]. We showed that five residues (amino acids 540, 561, 564, 566, 567 in human MxA) in L4 have evolved under positive selection during simian primate evolution. Moreover, single amino acid changes at amino acid 561 in the L4 region of MxA were necessary and sufficient to explain dramatic differences in species-specific antiviral activity of primate MxA proteins against Thogoto virus (THOV) and influenza A virus (IAV), two distantly related orthomyxoviruses [4].

We hypothesized that combinatorial mutagenesis of all five positively selected residues in L4 of human MxA might reveal the contributions of positions other than residue 561 in antiviral restriction and generate MxA variants with increased antiviral activity (here, called 'super-restrictors') relative to human MxA. Our combinatorial analyses revealed strict amino acid requirements at some L4 positions (*e.g.,* residue 561) but also significant contributions of other L4 positions (*e.g.,* residues 540, 564) to gain and loss of restriction against THOV. Our analyses also revealed a context-specific epistasis between L4 residues, thereby reiterating the merit of our combinatorial mutagenesis approach. Finally, consistent with our predictions, our analyses reveal 'super-restrictor' variants of human MxA that have 10-fold higher antiviral activity against THOV, in spite of the fact that human MxA is itself one of the most potent natural restrictors of THOV. We reveal that the basis of this super-restriction is increased binding to the THOV nucleoprotein (NP). Unexpectedly, we find an inverse correlation between the antiviral activity of MxA variants against the THOV and IAV orthomyxoviruses. Our analyses thus reveal not only a powerful means to elicit super-restrictor versions of antiviral proteins, but also reveal an unexpected breadth-versus-specificity tradeoff that shapes the adaptive landscape of antiviral proteins in nature.

Results

Combinatorial mutagenesis reveals L4 sequence requirements for THOV restriction

Retrospective analyses can not only identify which residues have recurrently evolved due to relentless adaptation (positive selection) [3], but also identify residues that have not undergone any changes despite millions of years of protein divergence (purifying selection). We reasoned that a prospective approach focusing on positively selected sites, while preserving sites that have evolved under purifying selection, might reveal insights into the selective pressures that mold antiviral defense repertoires and uncover unexplored potential for enhanced antiviral activity. Furthermore, we reasoned that focusing on only a subset of sites revealed to be important for past adaptation by retrospective studies would allow us to use combinatorial mutagenesis to more comprehensively survey potential paths of antiviral adaptation.

We therefore generated a MxA gene variant library (Figure 1A, Figure S1) encoding MxA proteins containing random combinations of all amino acids in the five positively selected sites of L4 in wild-type human MxA (hereafter referred to as wtMxA). Five residues in the L4 of wtMxA protein have evolved under recurrent positive selection in simian primates: 540 (G), 561 (F), 564 (F), 566 (S) and 567 (S). These residues were combinatorially mutagenized to all 20 amino acids using primers with NNS codons (Methods). We randomly selected ~600 clones and individually evaluated their antiviral activity against THOV using a minireplicon assay (Methods). In this minireplicon assay, the viral polymerase components (PB2, PB1, PA) as well as the genomic-RNA-binding nucleoprotein (NP) drive the expression of a firefly luciferase reporter. Overexpression of wtMxA reduces THOV-mediated luciferase expression in a dose-responsive manner (Figure S2), while a catalytic mutation in the MxA GTPase domain, T103A, renders human MxA non-restrictive against THOV regardless of MxA plasmid input concentration [4]. By measuring the restrictive activities of T103A and wtMxA multiple times, we determined the range of variation of wtMxA THOV restrictive activity (Figure S3) and then compared the restriction of all MxA variants (Figure 1B, Table S1).

Despite only altering the 5 most rapidly evolving L4 positions, we found that most MxA variants (95%) had worse THOV restriction than wtMxA, consistent with the evolutionary finding that these residues are crucial for MxA's antiviral function. We searched for patterns to explain the sequence requirements to maintain MxA antiviral activity against THOV. We found that all restrictive variants possessed a hydrophobic aromatic amino acid (Phe F, Tyr Y or Trp W) at residue 561 (Figure 1B, right). These results extend our previous findings that a hydrophobic

aromatic amino acid at residue 561 could confer antiviral restriction against THOV in the context of wtMxA [4]. While many non-restrictive variants had non-aromatic residues at position 561, we obtained at least 21 combinatorial MxA variants that had weaker anti-THOV activity than wtMxA despite possessing F, Y, or W at residue 561 (Figure 1C). Therefore, we conclude that an aromatic residue at position 561 is necessary but not sufficient to confer anti-THOV restriction. This implies that the other four rapidly evolving L4 residues also directly contribute to MxA antiviral activity against THOV.

'Super-restrictor' MxA variants and their molecular basis

In our initial screen, we also discovered four MxA variants (LFVKG, GYKDA, TFGNF and GYGQL in the randomized L4 positions) with 2- to 5-fold better restrictive ability against THOV than wtMxA (Figure 1B, 1C). We validated that these variants are indeed more potent than wtMxA throughout a dose-response curve. This increased potency cannot be explained by expression levels. These super-restrictor MxA variants validate our premise that combinatorial mutagenesis of positively selected residues could yield increase antiviral potency. However, because only 6% of the initial mutant pool possessed the required aromatic amino acid at position 561 (Figure 1B, 1C), our initial screen did not fully explore the super-restrictor sequence space. Therefore, we designed a second combinatorial library, in which we fixed site 561 as F, the amino acid present in wtMxA, while randomizing the other four amino acids under positive selection in L4.

For the four-site screen, we randomly selected 168 variants and assessed their antiviral activity in triplicate against THOV using the minireplicon assay (Figure 2A, Table S2). As predicted, a greater proportion (~65%) of MxA variants in the 4-site screen have anti-THOV activity comparable to wtMxA (Figure 2A) as opposed to only ~6% in the five-site screen (Figure 1B). Moreover, we found that ~13% of the 4-site variants are super-restrictors (Figure 2A) compared to only ~0.5% in the initial screen. To validate the top super-restrictors from both screens, we retested their activity in the THOV minireplicon assay (Figure 2B). Most of the variants identified as super-restrictors again showed enhanced restriction of THOV. In particular, the three most potent super-restrictors obtained with the 4-site screen (*i.e.,* QFAYS, VFRSV and TFAMC) are 7 to 10-fold more potent than wtMxA (Figure 2B). Although these super-restrictor variants are expressed at similar levels as wtMxA (Figure 2C, bottom), they have higher anti-THOV restriction at all levels of expression (Figure 2C). This significant improvement in THOV

restriction is remarkable because human wtMxA is the most potent naturally-occurring MxA ortholog for THOV restriction identified to date [4, 13].

Since we previously found that changes at amino acid 561 correlated with gain or loss of binding to the target nucleoprotein (NP) [4, 14], we next tested if increased MxA-NP interaction could explain the super-restriction phenotype [15-17]. MxA forms higher-order oligomers that are required for antiviral activity [11, 18, 19]. Thus, increased affinity of MxA monomers to NP might result in higher avidity in the multimeric MxA complex that interacts with THOV NP [4, 11, 14, 18, 20]. We assayed the association of NP with FLAG-tagged MxA super-restrictors in THOV-infected cells. We found that the most potent super-restrictors QFAYS, VFRSV and TFAMC pull down a larger fraction of cellular NP compared to wtMxA (GFFSS) (Figure 2D), whereas non-restrictor MxA versions (variant PFFSS or an MxA protein with a deletion of L4) have greatly reduced ability to pull down THOV NP. This suggests that the mechanism for super-restriction involves a more avid interaction of MxA with the THOV NP target protein.

We next investigated whether there was a common sequence pattern that could explain the enhanced anti-THOV restriction activity of the super-restrictor variants. We noticed that MxA variants with glutamine (Q) at position 540 were enriched among the 4-site super-restrictors (Figure 2B, Figure S4), including the most potent super-restrictor, QFAYS. Swapping the glutamine (Q) residue at position 540 in QFAYS, QFQSM and QFVVM to the glycine (G) present in wtMxA led to a significant drop in restriction activity – in one case to below wtMxA levels (Figure 3A). This finding suggested that position 540 could play an important role in determining super restriction. In order to better explore the role of position 540 in super-restriction, we mutated the glycine at residue 540 (G540) in the wtMxA backbone to every other amino acid (Figure 3B). We found that changing position 540 to Q in wtMxA leads to only a 2-fold improvement in antiviral activity (Figures 3A, B), significantly less than the super-restriction of the QFAYS, QFQSM and QFVVM variants. Moreover, although individual substitutions (*e.g.,* G540A or G540Q) could increase THOV restriction 2-3 fold, none of the single substitutions at residue 540 achieved the level of super-restriction in the most potent variants (compare Figures 3B with 2B). These results imply that Q540 is necessary but not sufficient to explain super-restriction. Consistent with these findings, although variants QFQSM and QFVSM are both super-restrictors, variant QFLSM is a non-restrictor despite differing only at position 564 (Figure 3C). Our findings imply that multiple positively selected L4 residues in addition to residue 540, can enhance or epistatically interfere (*e.g.,* QFLSM) with the degree of MxA restriction.

Our single residue swap studies revealed 540A to be the most potent restrictor (Figure 3B) in the context of the wtMxA backbone. However, we speculated that the enhancement of MxA restriction by 540A mutations might also be context-dependent. To test this possibility, we made 540A mutations in three super restrictors – HFSGR, TFAMC and VFRSV – that each encode for an 'unfavorable' amino acid at position 540 in the wtMxA context (Figure 3B). If the enhancement in MxA restriction by 540A were context-independent, we would expect to see further enhancement of anti-THOV activity in each of the HFSGR, TFAMC and VFRSV super-restrictors via 540A mutation. Instead, we found that HFSGR and AFSGR MxA variants were equivalent in their restriction, as were the TFAMC and AFAMC variants (Figure 3D). However, a V540A swap in VFSRV dramatically lowered restriction activity by 10-fold (Figure 3D), even though the same V540A swap in a wtMxA context (XFFSS) increased restriction by 5-fold (Figure 3B). Together, our results show that a simple model of context-independent contribution of all L4 positions cannot account for the broad repertoire of super-restrictors. If our analyses had relied on combining 'restriction favorable' single residues like A540 while avoid 'restriction-unfavorable' residues like V540, we would have never been able to discover super-restrictors like VFRSV. Instead, we find that epistasis shapes the restriction profile of many MxA super-restrictor variants such as VFRSV. This discovery of context-specificity demonstrates the power an unbiased combinatorial mutation approach to provide a more thorough exploration of the adaptive landscape for discovering novel super-restrictors.

A breadth-versus-specificity tradeoff underlies MxA antiviral specificity

MxA restricts a number of different viruses and its evolution is likely to have been shaped by a number of host-virus interactions and MxA from humans restricts both THOV and IAV more potently than other primate MxA variants [4]. Therefore, we investigated how restriction against one virus (*i.e.,* THOV) affects the antiviral activity against the other (*i.e.,* the H5N1 strain of IAV). This comparison is especially appropriate because restriction of both viruses depends on the L4, specifically on residue F561 in MxA [4]. Thus, one might expect that gaining super-restriction against one virus via changes in the MxA L4 might also confer super-restriction against other viruses that are restricted by a similar interface. We first investigated several 'non-restrictor' variants that still preserved an aromatic residue at residue but had lost anti-THOV activity (Figure 4A). We tested these MxA variants for their anti-IAV restriction activity using the IAV minireplicon assay described previously (Methods) as a proxy for IAV replication. We found

that these variants still retained anti-IAV activity (Figure 4B). Thus, loss of restriction against THOV is not necessarily linked to loss of IAV restriction. Next, we tested 'super-restrictor' MxA variants that had 10-fold higher activity against THOV (Figure 2). Surprisingly, we found that all anti-THOV super-restrictors tested have lower antiviral activity against IAV than wtMxA (Figure 4C). Thus, MxA optimization to restrict THOV also impairs its ability to restrict IAV. These results demonstrate that MxA has different optimal sequence states for different viruses. 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 renders the protein ineffective against other viruses.

Discussion

Taking advantage of an approach that combines combinatorial mutagenesis and evolution-guided insight into residues that have been subject to positive selection, we have been able to identify MxA variants that are 10-fold better anti-THOV restrictors than wtMxA, itself the most potent MxA ortholog known against THOV [4, 13]. Our analyses reveal that these 'super-restrictors' require contribution from several positions in L4 including the necessary requirement of an aromatic residue at position 561. Conversely, we find that 'unfavorable' residues can also render MxA variants to become 'non-restrictors' against THOV in spite of having an aromatic residue at position 561. Thus, multiple L4 residues that have evolved under recurrent positive selection over primate evolution shape the antiviral specificity of MxA.

Our analyses further reveal that context-specific epistasis helps determine the contributions of individual L4 residues in shaping MxA specificity. We note instances where the same amino acid mutation (A540V) can have opposite effects on restriction activity in MxA proteins that only differ at three other L4 residues. Our results suggest that the path to adaptation in MxA is indeed shaped by historical contingency (*i.e.,* what MxA sequence was present when adaptation occurred) as well as by epistasis. These findings that are reminiscent of previous work delineating the evolutionary path to adaptation is often shaped by contingency and epistasis [6-8].

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 [21] by relaxing the constraints imposed by epistasis and historical contingency. Using this method, we are thus able to obtain super-restrictor MxA versions that would not necessarily have been obtained by combining 'favorable' residues identified by a deep-mutational scan of all L4 positions.

It is not necessarily true that the positive selection of MxA L4 was shaped via interactions with NP proteins of orthomyxoviruses like THOV or IAV. However, we can view both of these extant viruses are proxies for other selective events, including some that occurred deep in simian primate evolution. Nonetheless, our analysis MxA variants derived via combinatorial mutagenesis reveals an unexpected breadth-versus-specificity tradeoff; MxA 'super-restrictor' variants with 10-fold higher THOV restriction appear to lose all IAV restriction. Our findings suggest that, under threat by multiple viruses, antiviral genes such as MxA appear to be under evolutionary pressure to harbor more broadly active alleles even at the expense of more potent, specific antiviral activity. This result is reminiscent of a recent study, which reported that generalist (promiscuous) MHC class II alleles are selected for in human populations in response to high pathogen diversity [22].

Similar breadth-specificity tradeoffs have also been invoked in the case of cytochrome p450 detoxification genes that protect herbivorous insects from plant counterdefenses [23] and in plant disease resistance (R) genes encoded by the Leucine Rich Repeat (LRR) gene family [24]. Furthermore, although it was not the intention of our combinatorial mutagenesis strategy to do so, the breadth-specificity tradeoff we observe is nevertheless similar to the outcome of an artificial selection experiment that made the GroEL chaperone highly specific for one substrate but at the expense of its substrate-binding breadth [25]. It is remarkable that critical components of the innate immune defense apparatus may evolve under the same constraints as an intracellular chaperone, sacrificing high avidity of binding to a given viral substrate in order to maintain a breadth of target binding and thereby antiviral response.

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 [3]. We propose that by leveraging the signature of positive selection, our evolution-guided approach can similarly increase the potency of other antiviral proteins. In many of these cases (*e.g.,* TRIM5, MxB), positive selection appears enriched in unstructured loops [4, 5, 26-

29] like MxA L4 [18]. We speculate that unstructured loops like L4 represent an ensemble of many conformations [18] 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.

1. Shaw AE, Hughes J, Gu Q, Behdenna A, Singer JB, Dennis T, et al. Fundamental properties of the mammalian innate immune system revealed by multispecies comparison of type I interferon responses. PLoS Biol. 2017;15(12):e2004086. doi: 10.1371/journal.pbio.2004086. PubMed PMID: 29253856; PubMed Central PMCID: PMCPMC5747502.
2. Yan N, Chen ZJ. Intrinsic antiviral immunity. Nat Immunol. 2012;13(3):214-22. doi: 10.1038/ni.2229. PubMed PMID: 22344284; PubMed Central PMCID: PMCPMC3549670.
3. Daugherty MD, Malik HS. Rules of engagement: molecular insights from host-virus arms races. Annu Rev Genet. 2012;46:677-700. doi: 10.1146/annurev-genet-110711-155522. PubMed PMID: 23145935.
4. Mitchell PS, Patzina C, Emerman M, Haller O, Malik HS, Kochs G. Evolution-guided identification of antiviral specificity determinants in the broadly acting interferon-induced innate immunity factor MxA. Cell Host Microbe. 2012;12(4):598-604. doi: 10.1016/j.chom.2012.09.005. PubMed PMID: 23084925; PubMed Central PMCID: PMCPMC3540999.
5. Sawyer SL, Wu LI, Emerman M, Malik HS. Positive selection of primate TRIM5alpha identifies a critical species-specific retroviral restriction domain. Proc Natl Acad Sci U S A. 2005;102(8):2832-7. doi: 10.1073/pnas.0409853102. PubMed PMID: 15689398; PubMed Central PMCID: PMCPMC549489.
6. Starr TN, Flynn JM, Mishra P, Bolon DNA, Thornton JW. Pervasive contingency and entrenchment in a billion years of Hsp90 evolution. Proc Natl Acad Sci U S A. 2018;115(17):4453-8. doi: 10.1073/pnas.1718133115. PubMed PMID: 29626131; PubMed Central PMCID: PMCPMC5924896.
7. Starr TN, Picton LK, Thornton JW. Alternative evolutionary histories in the sequence space of an ancient protein. Nature. 2017;549(7672):409-13. doi: 10.1038/nature23902. PubMed PMID: 28902834.
8. Shah P, McCandlish DM, Plotkin JB. Contingency and entrenchment in protein evolution under purifying selection. Proc Natl Acad Sci U S A. 2015;112(25):E3226-35. doi: 10.1073/pnas.1412933112. PubMed PMID: 26056312; PubMed Central PMCID: PMCPMC4485141.
9. Haller O, Staeheli P, Schwemmle M, Kochs G. Mx GTPases: dynamin-like antiviral machines of innate immunity. Trends Microbiol. 2015;23(3):154-63. doi: 10.1016/j.tim.2014.12.003. PubMed PMID: 25572883.
10. Mitchell PS, Emerman M, Malik HS. An evolutionary perspective on the broad antiviral specificity of MxA. Curr Opin Microbiol. 2013;16(4):493-9. doi: 10.1016/j.mib.2013.04.005. PubMed PMID: 23725670; PubMed Central PMCID: PMCPMC3763691.
11. Gao S, von der Malsburg A, Dick A, Faelber K, Schroder GF, Haller O, et al. Structure of myxovirus resistance protein a reveals intra- and intermolecular domain interactions required for the antiviral function. Immunity. 2011;35(4):514-25. doi: 10.1016/j.immuni.2011.07.012. PubMed PMID: 21962493.

12. Haller O, Gao S, von der Malsburg A, Daumke O, Kochs G. Dynamin-like MxA GTPase: structural insights into oligomerization and implications for antiviral activity. J Biol Chem. 2010;285(37):28419-24. doi: 10.1074/jbc.R110.145839. PubMed PMID: 20538602; PubMed Central PMCID: PMCPMC2937866.
13. Fuchs J, Holzer M, Schilling M, Patzina C, Schoen A, Hoenen T, et al. Evolution and Antiviral Specificities of Interferon-Induced Mx Proteins of Bats against Ebola, Influenza, and Other RNA Viruses. J Virol. 2017;91(15):e00361-17. doi: 10.1128/JVI.00361-17. PubMed PMID: 28490593; PubMed Central PMCID: PMCPMC5512242.
14. Patzina C, Haller O, Kochs G. Structural requirements for the antiviral activity of the human MxA protein against Thogoto and influenza A virus. J Biol Chem. 2014;289(9):6020-7. doi: 10.1074/jbc.M113.543892. PubMed PMID: 24448803; PubMed Central PMCID: PMC3937669.
15. Mänz B, Dornfeld D, Gotz V, Zell R, Zimmermann P, Haller O, et al. Pandemic influenza A viruses escape from restriction by human MxA through adaptive mutations in the nucleoprotein. PLoS Pathog. 2013;9(3):e1003279. doi: 10.1371/journal.ppat.1003279. PubMed PMID: 23555271; PubMed Central PMCID: PMCPMC3610643.
16. Riegger D, Hai R, Dornfeld D, Manz B, Leyva-Grado V, Sanchez-Aparicio MT, et al. The nucleoprotein of newly emerged H7N9 influenza A virus harbors a unique motif conferring resistance to antiviral human MxA. J Virol. 2015;89(4):2241-52. doi: 10.1128/JVI.02406-14. PubMed PMID: 25505067; PubMed Central PMCID: PMCPMC4338896.
17. Zimmermann P, Manz B, Haller O, Schwemmle M, Kochs G. The viral nucleoprotein determines Mx sensitivity of influenza A viruses. J Virol. 2011;85(16):8133-40. doi: 10.1128/JVI.00712-11. PubMed PMID: 21680506; PubMed Central PMCID: PMCPMC3147989.
18. Gao S, von der Malsburg A, Paeschke S, Behlke J, Haller O, Kochs G, et al. Structural basis of oligomerization in the stalk region of dynamin-like MxA. Nature. 2010;465(7297):502-6. doi: 10.1038/nature08972. PubMed PMID: 20428112.
19. Nigg PE, Pavlovic J. Oligomerization and GTP-binding Requirements of MxA for Viral Target Recognition and Antiviral Activity against Influenza A Virus. J Biol Chem. 2015;290(50):29893-906. doi: 10.1074/jbc.M115.681494. PubMed PMID: 26507657; PubMed Central PMCID: PMC4706002.
20. Kochs G, Haller O. GTP-bound human MxA protein interacts with the nucleocapsids of Thogoto virus (Orthomyxoviridae). J Biol Chem. 1999;274(7):4370-6. PubMed PMID: 9933640.
21. Fowler DM, Fields S. Deep mutational scanning: a new style of protein science. Nat Methods. 2014;11(8):801-7. doi: 10.1038/nmeth.3027. PubMed PMID: 25075907; PubMed Central PMCID: PMC4410700.
22. Manczinger M, Boross G, Kemeny L, Muller V, Lenz TL, Papp B, et al. Pathogen diversity drives the evolution of generalist MHC-II alleles in human populations. PLoS Biol. 2019;17(1):e3000131. doi: 10.1371/journal.pbio.3000131. PubMed PMID: 30703088.
23. Li X, Baudry J, Berenbaum MR, Schuler MA. Structural and functional divergence of insect CYP6B proteins: From specialist to generalist cytochrome P450. Proc Natl Acad Sci U S A. 2004;101(9):2939-44. doi: 10.1073/pnas.0308691101. PubMed PMID: 14981232; PubMed Central PMCID: PMC365724.
24. Farnham G, Baulcombe DC. Artificial evolution extends the spectrum of viruses that are targeted by a disease-resistance gene from potato. Proc Natl Acad Sci U S A. 2006;103(49):18828-33. doi: 10.1073/pnas.0605777103. PubMed PMID: 17021014; PubMed Central PMCID: PMC1693747.

25. Wang JD, Herman C, Tipton KA, Gross CA, Weissman JS. Directed evolution of substrate-optimized GroEL/S chaperonins. Cell. 2002;111(7):1027-39. PubMed PMID: 12507429.
26. Biris N, Yang Y, Taylor AB, Tomashevski A, Guo M, Hart PJ, et al. Structure of the rhesus monkey TRIM5alpha PRYSPRY domain, the HIV capsid recognition module. Proc Natl Acad Sci U S A. 2012;109(33):13278-83. doi: 10.1073/pnas.1203536109. PubMed PMID: 22847415; PubMed Central PMCID: PMCPMC3421187.
27. Yang H, Ji X, Zhao G, Ning J, Zhao Q, Aiken C, et al. Structural insight into HIV-1 capsid recognition by rhesus TRIM5alpha. Proc Natl Acad Sci U S A. 2012;109(45):18372-7. doi: 10.1073/pnas.1210903109. PubMed PMID: 23091002; PubMed Central PMCID: PMCPMC3494900.
28. Fribourgh JL, Nguyen HC, Matreyek KA, Alvarez FJD, Summers BJ, Dewdney TG, et al. Structural insight into HIV-1 restriction by MxB. Cell Host Microbe. 2014;16(5):627-38. doi: 10.1016/j.chom.2014.09.021. PubMed PMID: 25312384; PubMed Central PMCID: PMC4252739.
29. Mitchell PS, Young JM, Emerman M, Malik HS. Evolutionary Analyses Suggest a Function of MxB Immunity Proteins Beyond Lentivirus Restriction. PLoS Pathog. 2015;11(12):e1005304. doi: 10.1371/journal.ppat.1005304. PubMed PMID: 26658285; PubMed Central PMCID: PMC4687636.
30. Nettling M, Treutler H, Grau J, Keilwagen J, Posch S, Grosse I. DiffLogo: a comparative visualization of sequence motifs. BMC Bioinformatics. 2015;16:387. doi: 10.1186/s12859-015-0767-x. PubMed PMID: 26577052; PubMed Central PMCID: PMCPMC4650857.

We thank HHMI EXROP interns María Gutierrez (UT El Paso) and Isabel Delwel (Univ. North Texas) for assistance with THOV and IAV minireplicon assays, J. Young for help with DiffLogo plots, and P. Mitchell and members of the Malik and Emerman labs for advice and discussions. We also thank J. Bloom, N. Elde, P. Mitchell, S. Soh, J. Tenthorey and J. Young for their comments on the manuscript. This work was funded by the NIH CMB training grant T32GM007270 (to RCT), NSF predoctoral fellowship 2015198621 (to EH), NIH grant K99/R00 GM112941 (to RNM), DFG (German Research Foundation) grant KO 1579/8-2 (to GK), grant P50GM082250 (PI: Nevan Krogan, subaward to ME, HSM), R01 AI030927 (to ME), and a grant from the Mathers Foundation (to HSM). HSM is an Investigator of the Howard Hughes Medical Institute.

FIGURE LEGENDS

Figure 1. Combinatorial random mutagenesis of rapidly evolving sites in the MxA L4 loop alters antiviral potency against THOV. A. A positive selection-based mutagenesis scheme to obtain MxA combinatorial variants of the 5 rapidly-evolving sites (arrowheads) in L4 (N = any nucleotide, S = G or C). **B.** Random MxA variants are presented in order of increasing THOV fold restriction, with their consensus sequences indicated to the right. 'Non-restrictor'(NR)

variants with lower restriction than wtMxA are shown below the dashed green line (these contain very diverse residues in the randomized L4 positions). Approximately 6% of the variants have equivalent restriction to wtMxA (green restrictors or 'R', between the pink and green dashed lines) and these all possess a F, Y, or W residue at position 561. L4 sequences of four variants that restrict THOV better than wtMxA (pink 'super-restrictors' or 'SR') are shown. **C.** Fold restriction of MxA variants recovered with F, Y, or W residues at position 561. Data is presented as means of three experiments (error bars ±SEM). For super-restrictor variants (pink) values were compared to the wtMxA control in the experiment *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$ (t-test).

Figure 2. A modified mutagenesis strategy yields potent THOV super-restrictor variants. A. A modified combinatorial mutagenesis strategy with fixed F561 and four randomized L4 positions. Non-restrictor, restrictor, and super-restrictor variants are presented in order of increasing restriction activity (X-axis). **B.** Re-testing of anti-THOV restriction activity of super-restrictor MxA variants (relative to the empty vector) obtained from the original five-site screen (pink) and the modified 4-site screen (black). The three most potent super-restrictors (blue), have 7- to 10-fold higher anti-THOV restriction activity relative to wtMxA. **C.** Dose-responsiveness of super-restriction activity for MxA variants. **D.** Co-immunoprecipitation studies of viral NP target protein with different MxA variants using cell lysates of THOV-infected 293T cells transfected with the FLAG-tagged MxA expression constructs. Data is presented as mean of three experiments (error bars ±SEM). Western blots of the FLAG-tagged MxA variants and a loading control (β-actin or β-tubulin) are shown.

Figure 3. Molecular basis of super-restriction. A. Comparison of anti-THOV activity of Q540 (super restrictor) to Q540G MxA variants. A G540Q mutation in wtMxA resulted in a gain in restriction (white versus green bar), and Q-to-G changes reduced super-restriction capability (black versus gray bars). **B.** Anti-THOV activity of super-restrictor QFQSM MxA variant compared to QFLSM and QFVSM MxA variants. **C.** Effect of single residue G540X mutations in wtMxA (green) on anti-THOV restriction relative to the empty vector. **D.** Effect of single residue X540A changes in super-restrictor HFSGR, TFAMC and VFRSV MxA variants. All variants are compared to wtMxA *$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, ****$p \leq 0.0001$ (t-test).

Figure 4. A breadth-versus-specificity tradeoff in MxA restriction of IAV and THOV. of L4 loop variants. A, B. MxA 'non-restrictor' variants still encoding a necessary aromatic residue at

position 561 were re-tested in a THOV (**A**) and IAV (**B**) minireplicon assays. **C.** Restriction activity of MxA 5-site super-restrictor variants (pink) and 4-site super-restrictor variants (blue) against the H5N1 strain of IAV in a minireplicon assay. In all panels, data is presented as means of three experiments (error bars ±SEM). FLAG-tagged MxA and β-actin or β-tubulin protein levels were detected in THOV minireplicon lysates by Western blot. (*$p \leq 0.05$, **$p \leq 0.01$, ***$p \leq 0.001$, **** $p \leq 0.0001$, based on t-tests).

SUPPLEMENTARY FIGURE LEGENDS

Figure S1. Schematic of combinatorial mutagenesis screen. The mutant library was generated using oligonucleotide-directed mutagenesis of all five rapidly-evolving sites in wtMxA L4 (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 minireplicon assay (Methods) in triplicate. In a second iteration, we screened combinatorial mutants of L4 except for position 561 that was fixed as a Phenylalanine (F).

Figure S2. Calibrating the restriction sensitivity range of the minireplicon assay. Based on the overall activity of the variants analyzed for anti-THOV restriction activity, we had to adjust the dose-responsiveness of the minireplicon assay to maximize the dynamic range, by adjusting the amount of MxA plasmid that was transfected. To set appropriate dose-sensitivity, we transfected increasing amounts of wtMxA and the catalytically inactive mutant MxA-T103A expression plasmids with the THOV minireplicon into HEK-293T cells. The amounts used for the less-sensitive 5-site screen (50 ng) (Figure 1) and more sensitive 4-site screen (15 ng) (Figure 2) are indicated on the graph. Data is presented as a mean average of three experiments (error bars ±SEM).

Figure S3. Setting thresholds for wtMxA restriction levels for the 5-site MxA variant screen. We tested 91 distinct clones of wtMxA and 13 MxA-T103A mutants in triplicates in the minireplicon assay. The range of restriction potential represents the wtMxA-equivalent range of restriction (dashed lines, Figure 1B).

Figure S4. Identifying amino acid residues enriched in the 4-site super-restrictor pool. We used difference logo plots to compare super-restrictors to the full library. For these clones, we constructed an amino acid position frequency matrix, and generated difference logo plots using the DiffLogo R package. The stack height for each motif position is calculated using the Jensen-Shannon divergence (JS) (22). 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. For example, Q540 is enriched in the super-restrictor pool in our analysis of 168 4-site MxA variants.

Table S1. THOV minireplicon measurements of restriction activity for ~ 600 5-site MxA variants.

Table S2. THOV minireplicon measurements of restriction activity for all 168 4-site MxA variants.

MATERIALS AND METHODS

NNS library construction

The L4 loop five-point variant 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 bp in the region encoding for the rapidly evolving residues. For the targeted positions, the oligonucleotides contain NNS codons (N = A, T, C or G, and S = G or C). This biased randomization results in 32 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. The maximum complexity of this library is thus 20^5, or 3.2 million variants. One round of PCR was carried out with either the sense or antisense oligonucleotide and a flanking antisense or sense 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 µl dH_2O, which was used to transform ElectroMAX™ DH10B™ cells (ThermoFisher). 836 colonies were randomly selected and overnight in LB liquid medium containing 100 µg/ml of ampicillin. All variants containing stop codons introduced by NNS mutagenesis were removed and not analyzed further.

The L4 loop four-site variant library was similarly constructed, except that only four L4 positions were randomized, whereas position 561 was fixed as F. This library has a maximum complexity of 20^4, or 160,000 variants.

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. 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 (50 ng for 5-ite screen and 15 ng for 4-site screen) of 200 ng MxA plasmids were co-transfected into HEK 293T cells using the reagent TransIT-LT1 (Mirus Bio) in a 96-well plate format. After 24 h, firefly luciferase activity was measured using the Dual-Glo system (Promega) and normalized to the constitutively co-expressed *Renilla* luciferase (transfection efficiency control). Results are presented relative to the luciferase readings in the absence of MxA, the empty vector control, and as the mean average of fold restriction in three independent experiments. Expression of MxA proteins was detected by Western blot analysis using mouse ANTI-FLAG® monoclonal antibody F4042 (Sigma) and rabbit anti-beta-actin polyclonal antibody ab8227 (Abcam) or anti-beta-tubulin ab6046 (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.

For the IAV (H5N1) restriction analyses, we performed similar experiments but with the H5N1 minireplicon system of A/Vietnam/1203/04 as described (13) in a 12-well format, including a reporter construct encoding firefly luciferase under the control of the viral promoter, and 300 ng of the MxA expression plasmid. Just like with the THOV minireplicon experiments, restriction was assayed relative to empty vector control by measuring firefly luciferase expression relative to the *Renilla* luciferase transfection control.

Co-Immunoprecipitation

293T cells seeded in 6-well plates were transfected with 1 μg of expression plasmids coding for flag-tagged wtMxA and MxA variants using the JetPEI transfection reagent (Polyplus). 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 were subjected to standard Western blot analysis using antibodies against the MxA protein and THOV NP as described previously (12).

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 [30]. 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.

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.

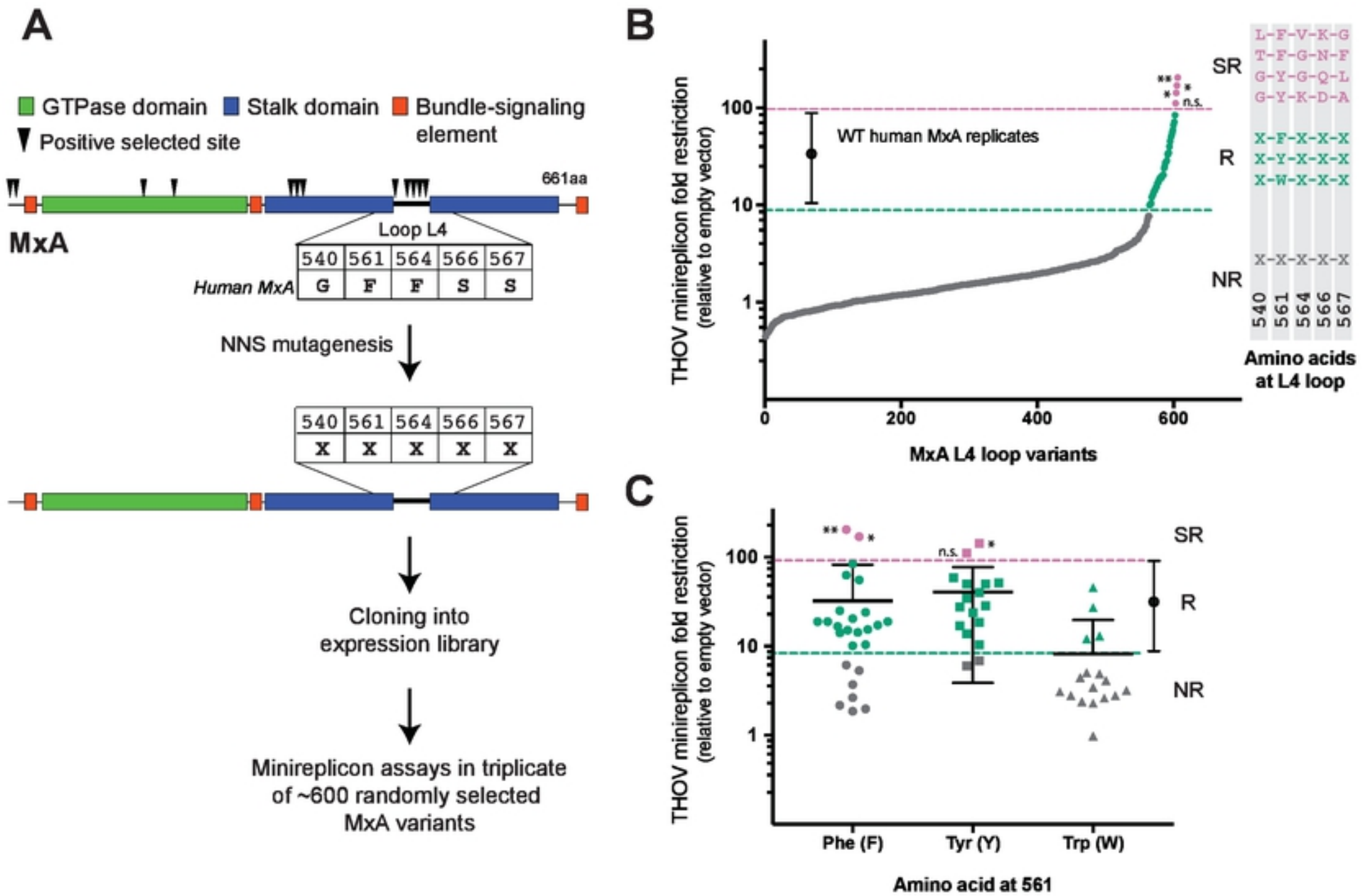

Figure 1

Figure 1

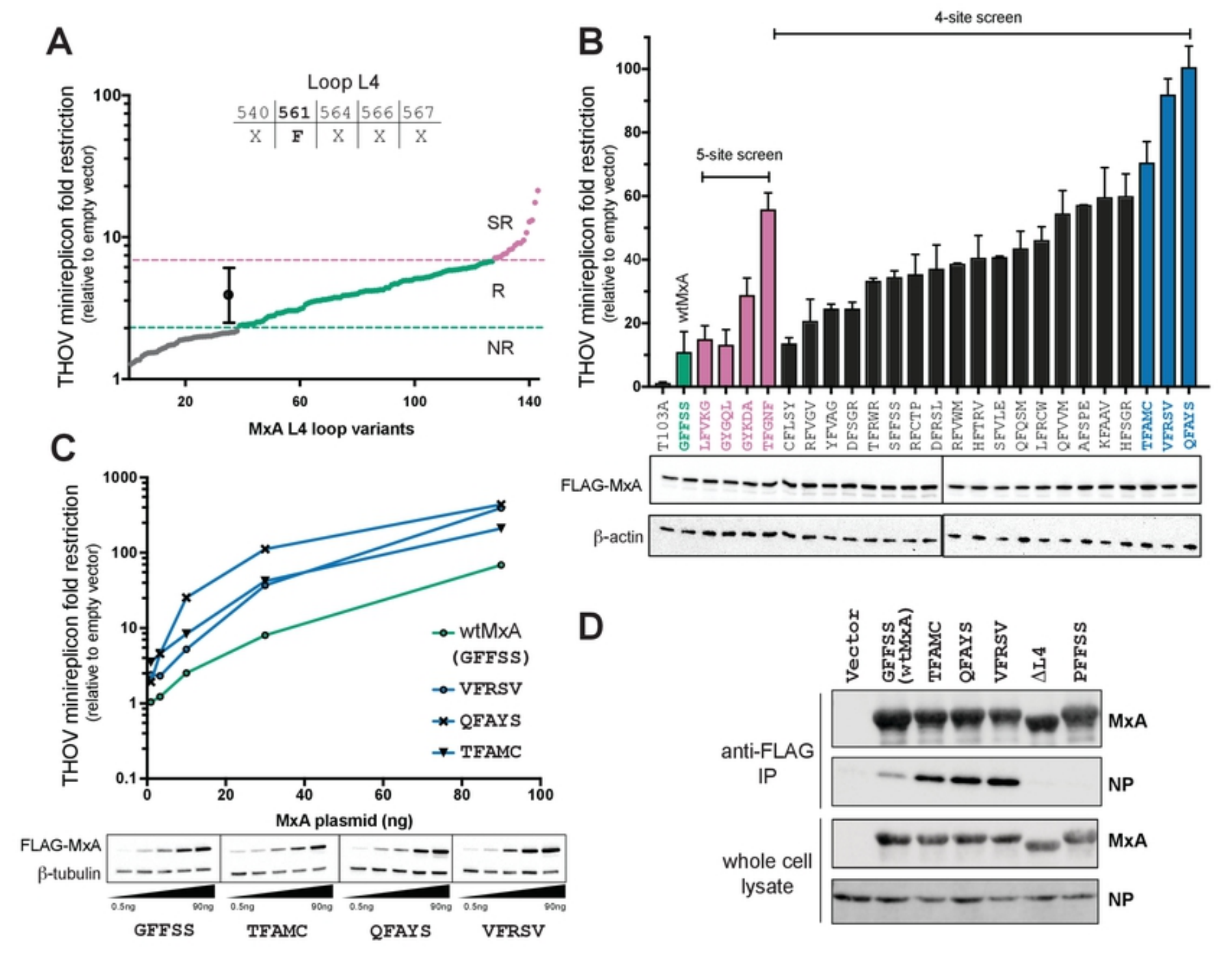

Figure 2

Figure 2

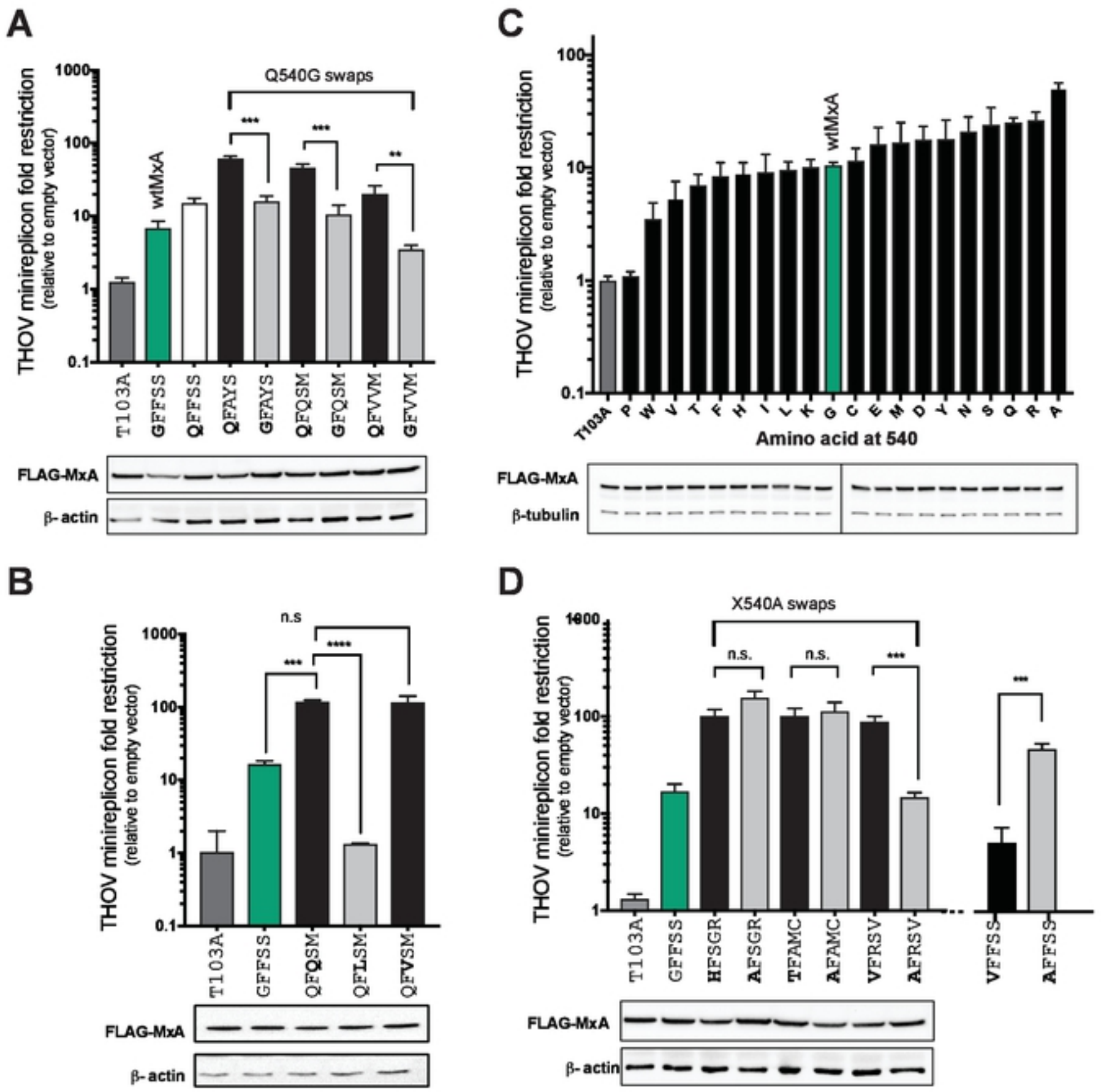

Figure 3

Figure 3

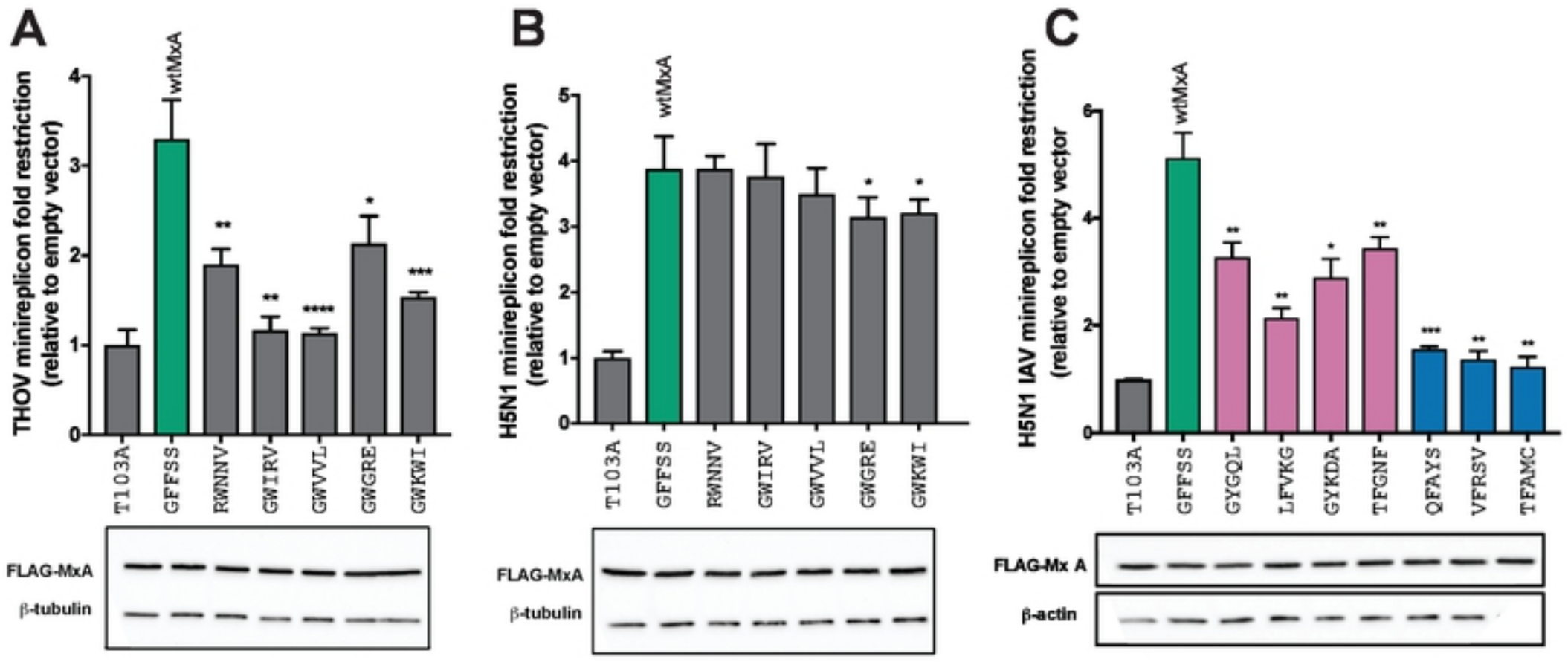

Figure 4

Figure 4