Published in final edited form as:
*Virology*. 2014 November ; 0: 63–71. doi:10.1016/j.virol.2014.07.049.

# Escape variants of the XPR1 gammaretrovirus receptor are rare due to reliance on a splice donor site and a short hypervariable loop

**Xiaoyu Lu**[a,c], **Carrie Martin**[a,c,1], **Christelle Bouchard**[b,2], and **Christine A. Kozak**[a,*]

[a]Laboratory of Molecular Microbiology, National Institute of Allergy and Infectious Diseases, Bethesda, MD, 20892

[b]Whitney Laboratory for Marine Bioscience, Univ. of Florida, St. Augustine, FL 32136

## Abstract

Entry determinants in the XPR1 receptor for the xenotropic/polytropic mouse leukemia viruses (XP-MLVs) lie in its third and fourth putative extracellular loops (ECLs). The critical ECL3 receptor determinant overlies a splice donor and is evolutionarily conserved in vertebrate XPR1 genes; 2 of the 3 rare replacement mutations at this site destroy this receptor determinant. The 13 residue ECL4 is hypervariable, and replacement mutations carrying an intact ECL3 site alter but do not abolish receptor activity, including replacement of the entire loop with that of a jellyfish (Cnidaria) XPR1. Because ECL4 deletions are found in all X-MLV-infected *Mus* subspecies, we deleted each ECL4 residue to determine if deletion-associated restriction is residue-specific or is effected by loop size. All deletions influence receptor function, although different deletions affect different XP-MLVs. Thus, receptor usage of a constrained splice site and a loop that tolerates mutations severely limits the likelihood of host escape mutations.

### Keywords

XPR1 gammaretrovirus receptor; Xenotropic/polytropic gammaretroviruses; Retrovirus restriction genes; Retrovirus evolution; Retrovirus entry

## Introduction

The xenotropic/polytropic mouse leukemia viruses (XP-MLVs) are gammaretroviruses isolated from laboratory and wild mice that differ from one another in host range, in their pathogenic potential and their *M. musculus* subspecies of origin (Kozak, 2010). The multiple

*Corresponding author: NIH, NIAID, LMM, Bldg. 4, Room 329, 4 Center Drive MSC 0460, Bethesda, MD 20892-0460, Tel: 301-496-0972, FAX: 301-480-6477, ckozak@niaid.nih.gov.
[1]Present addresses:
Eshelman School of Pharmacy, University of North Carolina, Chapel Hill, NC 27599
[2]University of South Florida, Sarasota, FL 34243
[c]Contributed equally

XP-MLV host range variants all use the same XPR1 receptor but differ in their ability to infect cells of their natural *Mus* hosts and cells of other mammalian species (Cloyd et al., 1985; Fischinger et al., 1975; Hartley et al., 1977; Levy, 1973). X-MLVs infect most mammalian species, whereas P-MLVs infect fewer species, and there are wild mouse XP-MLV isolates with distinctive host ranges and interference patterns (Cloyd et al., 1985; Yan et al., 2009). These tropism differences are due to polymorphisms in the viral envelope (*env*) glycoprotein and in the XPR1 receptor.

In *Mus* species, there are 6 functionally distinct XPR1 receptor variants resulting from sequence variation in the putative third and fourth extracellular loops (ECL3 and ECL4) (Bamunusinghe et al., 2013; Marin et al., 1999; Yan et al., 2009; Yan et al., 2010). At least six XPR1 residues have roles in virus entry, 3 in ECL3 and 3 in ECL4 (Marin et al., 1999; Yan et al., 2007; Yan et al., 2010). Substitutions at these sites result in restriction of one or more of the viruses that use this receptor. Two of these sites are particularly important for X-MLV infection: K500 in ECL3 and T582 in ECL4 (Marin et al., 1999). Both of these sites are mutated in the X-MLV-restrictive $Xpr1^{n}$ variant, found in the majority of laboratory mouse strains. Repair of either of these sites is sufficient to generate a receptor for X-MLVs, demonstrating the presence of two independent X-MLV receptor determinants in this protein, although there is evidence that receptor function is modulated by cooperativity between these sites and by involvement of other residues (Yan et al., 2009).

With very few exceptions, mammalian XPR1 proteins have X-MLV receptor function (Xu and Eiden, 2011). This scarcity of escape mutations results in part from the fact that the critical ECL3 residue, K500, is evolutionarily conserved, and by the fact that functional X-MLV receptors can carry multiple substitutions in ECL4 (Yan et al., 2010). Among mammalian XPR1 receptors, ECL4 deletions are only found in *Mus musculus* subspecies. Three different in-frame deletions mark different lineages of house mice: $Xpr1^{c}$ (*M. m. castaneus*), $Xpr1^{m}$ (*M. m. molossinus* and *M. m. musculus*) and $Xpr1^{n}$ (laboratory mice and *M. m. molossinus*). These deletions remove residues in an 8 residue stretch in the 13 residue ECL4 (Marin et al., 1999; Yan et al., 2010) (Figure 1A). Each of the 3 deleted receptors inhibit two or more viruses in the XP-MLV family, and the appearance of these restrictive variants in *Mus* evolution coincides with the acquisition of X-MLV endogenous retroviruses (ERVs) in these 3 subspecies (Bamunusinghe et al., 2013; Kozak, 2013; Kozak and O'Neill, 1987). The fact that three different deletions lie in this receptor determining loop suggests either that the six deleted residues are important for receptor function, or alternatively, that altering the size or structure of this ECL may be a particularly effective way to disable or modify receptor function in mice that harbor mutagenic and pathogenic viruses.

To assess whether the 6 residues deleted in *Mus musculus* XPR1 are critical for entry, or whether it is the size of the ECL4 loop that is important for receptor function, we evaluated receptor function for XPR1 constructs in which each one of the 13 ECL4 residues was deleted. We also replaced the entire ECL4 with the corresponding segment of the very divergent jellyfish (*Cyanea capillata*) *XPR*1 ortholog to assess the tolerance of functional receptors for major sequence variation. Finally, we screened XPR1 orthologs in 60 sequenced vertebrate genomes for sequence variations, and we evaluated the replacement mutations found at receptor-critical sites for possible roles in virus entry.

## Results

### A fortuitous receptor choice: the K500 XPR1 receptor determinant overlaps a splice donor site

The site in the XPR1 ECL3 with the greatest influence on receptor function is K500. Most mammals are susceptible to XP-MLVs, and K500 is completely invariant in the mammals we examined previously with the single exception of the X-MLV-restrictive laboratory mouse which carries E500 (Figure 1A) (Yan et al., 2010). In the avian lineage, the homologous site, K496, is also critical for XPR1 receptor function, and avian XPR1 receptor function can be disabled by two naturally occurring mutations at this site, K496E and K496Q (Martin et al., 2013). The mouse mutation and both bird mutations all result from substitutions in the first position of this codon (CAA and GAA). The fourth possible substitution at this position generates a stop codon. The second and third positions in codon 500 overlap the consensus splice donor site at the end of *Xpr1* exon 11 (Figure 1B). Although the consensus sequence allows all 4 bases at these two positions (−3 and −2 relative to the splice site breakpoint), frequencies of some bases are greatly reduced, imposing some constraints on sequence variation (Mount, 1982).

While the ECL3 receptor determinants in bird and mouse XPR1are disabled by the equivalent mutations K496E and K500E, the functional consequences of K496/500Q have only been examined in some species of fowl. To determine if this mutation can affect virus entry in the context of other XPR1 sequence variants, we introduced K500Q into the fully functional $Xpr1^{sxv}$ receptor (Figure 1A). Receptors were expressed in E36 Chinese hamster cells and evaluated for function using lacZ pseudotypes carrying Env glycoproteins of 4 XPR1-dependent viruses: the CAST-X X-MLV, FrMCF P-MLV and two wild mouse viruses, CasE#1 and Cz524, with atypical host range (Figure 1A). The $Xpr1^{sxv}$–K500Q mutant was able to transduce all 4 XP-MLVs, although with reduced efficiency (Figure 2, top row). Because X-MLV entry is effected independently by K500 in ECL3 and T582 in ECL4 (Marin et al., 1999), the presence of T582 in this construct may fully account for the observed receptor function. Therefore, we also introduced Q500 into $Xpr1^{n}$ to determine if Q500 could, like K500, compensate for the $Xpr1^{n}$ ECL4 mutation T582Δ and generate an X-MLV receptor (Figure 1A). Of the 4 tested pseudoviruses, only FrMCF is transduced by wild type $Xpr1^{n}$ (Figure 2, bottom row), and this was also the only pseudovirus that utilized the mutated $Xpr1^{n}$-E500Q receptor, indicating that this substitution does not function as an X-MLV receptor determinant in the context of either the avian or the *Mus* XPR1 and that it has minimal influence on P-MLV entry.

We looked for additional naturally occurring substitutions in vertebrate XPR1 orthologs at the exon 11 splice donor. XPR1 orthologs have been identified in at least 66 eukaryotic species (http://useast.ensembl.org/Mus_musculus/Gene/Compara_Ortholog?db=core;g=ENSMUSG00000026469;r=1:155275701-155417415). Among the 60 vertebrate XPR1 orthologs, we found only one other replacement mutation at codon 500. This codon, AGA, encodes R500 and is found in the anole lizard and in pika (Figure 1B). Introduction of this mutation into the *Mus* $Xpr1^{sxv}$ construct produces a protein that is functional as a receptor for the 4 XP-MLVs (Figure 2, top row). In $Xpr1^{n}$, R500, like K500, supports entry

of 3 XP-MLVs, but not CasE#1 (Figure 2, bottom), so this rare mutation is thus functionally equivalent to the wild type K500.

### The tolerance of functional XPR1 receptors for sequence variation in ECL4

A comparative analysis of the 13 residue XPR1 ECL4 sequence in various mammals showed extensive sequence hypervariation and identified only 3 conserved residues (S578, T580, G589), the first 2 of which are also conserved in avian XPR1s (Martin et al., 2013; Yan et al., 2010). Sequence variation at the other 10 ECL4 residues is compatible with receptor function, as virtually all mammalian orthologs are functional as X-MLV receptors (Kozak, 2010). Comparison of protein sequences across a larger and more diverse set of 60 vertebrates shows that none of the 13 residues is conserved among these species (not shown). To define the limits of ECL4 sequence variation in functional receptors we evaluated 3 sets of replacement mutations in the *Mus Xpr1* ECL4 for their effects on virus entry.

First, because ECL4 variations in 5 of the 6 previously characterized *Mus* XPR1s restrict virus entry, we examined a seventh naturally occurring *Mus* XPR1 variant, carrying the ECL4 substitution K585R, found in the African pygmy mouse, *Mus tenellus* (Figure 3) (Yan et al., 2010). Introduction of K585R into $Xpr1^{sxv}$ did not appreciably alter receptor function (Figure 4, top row). To determine if the K585R-containing ECL4 has a functional receptor determinant in the absence of the ECL3 determinant K500, we also introduced K585R into $Xpr1^{n}$-Δ582T, a construct which carries E500 and a full length ECL4 (Figure 1A), but again, no change in receptor phenotype was noted (Figure 4, bottom row). These results indicate that the conservative substitution K585R does not detectably affect virus-receptor interactions. That *M. tenellus* is thus likely to have a fully functional XPR1 receptor is not surprising as this species, in the subgenus *Nannomys*, diverged from other *Mus* species well before virus exposure resulted in acquisition of X-MLV endogenous retroviruses in *M. musculus*, subgenus *Mus*, and the coincident appearance of restrictive receptors (Kozak, 2013; Kozak and O'Neill, 1987; Yan et al., 2010).

Second, our analysis of avian species had identified an ECL4 mutation that disables avian XPR1 receptor function, Q579E, a site which aligns with two adjacent Thr residues in *Mus Xpr1*, deletion of one of which, T582Δ, in the restrictive *Mus* $Xpr1^{n}$, is receptor critical (Marin et al., 1999; Martin et al., 2013). To determine whether this mutation alters receptor function in the XPR1 orthologs of other species, we introduced the replacement, T583E, into mouse $Xpr1^{sxv}$, which is 83% identical in protein sequence to the chicken XPR1, but only 38% identical in ECL4 (Figure 3). Transfected E36 cells were able to transduce all 5 XP-MLVs, but susceptibility to FrMCF, but not MoMCF, was reduced over 100-fold (Figure 4, top row). We also introduced T583 into $Xpr1^{n}$-Δ583T; infection was not detected for 4 of the 5 viruses (Figure 4, bottom row). This site is therefore a key receptor determinant, and the T583E mutation results in loss of the ECL4 receptor determinant in the very different bird and mammalian receptors.

Finally, we tested the functional limits of ECL4 sequence variation by replacing the ECL4 of $Xpr1^{sxv}$ with the equivalent region from the highly divergent ortholog of a primitive metazoan (Figure 3). The XPR1 ortholog of cnidarians (jellyfishes, corals, sea anemones,

Hydra) was isolated from *Cyanea capillata* (a jellyfish, class: Scyphozoa) using degenerate primers for GCPRs. A novel sequence encoding a protein of 675 amino acids (deduced MW 78 813) was obtained from a cDNA library enriched in neurons. Although 600 million years separate mammals from members of this phylum, the sequence shows significant homology with mammalian XPR1 and also with the XPR1-related PHO1 sequence of plants; the CcXPR1 sequence is 49% identical to *Mus Xpr1*. The sequence in the region corresponding to ECL4 is quite divergent, although there is significant sequence homology upstream and downstream of the ECL4 segment.

Expression of the jellyfish CcXPR1 in E36 hamster cells resulted in no detectable transduction of any XP-MLV (not shown). We generated a chimeric *Xpr1* by replacing ECL4 of $Xpr1^{sxv}$ with that deduced for CcXPR1 (Figure 3). E36 transfectants expressing the MuCcXPR1 chimera showed significant susceptibility to X-MLV and MoMCF, and 2–4 log (100–10,000 fold) reductions in susceptibility to FrMCF and the 2 wild mouse viruses (Figure 4, top row). Introduction of the K500E mutation at the critical ECL3 receptor site into this chimera eliminated most receptor function (Figure 4, bottom row), although trace levels of infection were detected for most viruses. This indicates that the substantially different ECL4 of jellyfish can contribute to an active XP-MLV receptor in the context of a functional *Mus* XPR1 carrying K500.

## The effect of ECL4 deletion mutations on receptor function

Among mammalian species, only the 3 XP-MLV infected *Mus musculus* lineages carry *Xpr1* receptors that restrict virus entry due to 3 different in-frame deletions in ECL4: deletion of T582 in $Xpr1^{n}$, deletion of I579 in $Xpr1^{m}$ and the 5-residue T582-P586 deletion in $Xpr1^{c}$ (Figure 1a). Unlike $Xpr1^{n}$, the $Xpr1^{m}$ and $Xpr1^{c}$ receptors, like $Xpr1^{sxv}$, carry K500 in ECL3. To determine if the effects of these various deletions on receptor function are due to loss of key residues or to a structural change resulting from reduction in loop size, we assessed receptor function after removal of each of 12 ECL4 residues from $Xpr1^{sxv}$.

All of the deleted receptors were able to transduce the two X-MLVs (CAST-X and XMRV) (Figure 5), although deletion of I579 substantially reduced infection by both X-MLVs as shown previously (Yan et al., 2010). Deletion of residue T583 did not affect X-MLV receptor function as demonstrated previously for deletion of T582 (Marin et al., 1999). Infection with the 2 wild mouse pseudoviruses (Cz524 and CasE#1) was sensitive to multiple residue deletions, especially residues in the more N-terminal end of this ECL, positions 578–585. This segment includes the 6 residues deleted in $Xpr1^{m}$, $Xpr1^{n}$, or $Xpr1^{c}$, and these 3 *Mus* orthologs all show restricted infection with these 2 viruses (Kozak, 2010) (Figure 1A).

Infection with two P-MLVs, FrMCF and MoMCF, was reduced for nearly all of the deletions, but different patterns were observed for these two P-MLVs. Deletions of S578, F584 and K585 largely abolish infection by FrMCF, but not by MoMCF (Figure 5). This receptor-mediated tropism difference for two viruses in the P-MLV host range group was also seen for MuCcXpr1 and T583E (Figure 4) and is consistent with their previously observed differential infectivity for mutated *Mus Xpr1* receptors (Yan et al., 2009). This confirms that there are different host range subgroups among polytropic MLVs.

Because this set of deletions was made in *Xpr1$^{sxv}$*, it is likely that the presence of K500 contributed to their ability to transduce X-MLVs. Therefore, we made a second set of 7 deletions spanning the ECL4 of *Xpr1$^{n}$*-Δ583T, which has E500 (Figure 1A). All 7 deletions restrict infectivity by CAST-X X-MLV by at least 1000-fold (Figure 6), indicating that the X-MLV receptor determinant in ECL4 is not strictly defined by T582 as previously thought, and suggesting either that these deletions may result in a structural change in this loop that impacts receptor function or that the receptor interface extends over this entire loop. These 7 deletions also nearly eliminate infection by the 2 wild mouse viruses (Figure 6) indicating that these viruses rely on K500 as well as determinants in ECL4 as shown in Figure 5.

These 7 deletions, however, minimally alter infection by the two P-MLVs (Figure 6). Also, these deletions did not show the same differential sensitivity of FrMCF and MoMCF for deletions of S578, F584 and K585 as seen in the *Xpr1$^{sxv}$* constructs (Figure 5, 6). The *Xpr1$^{n}$*-Δ582T constructs differ from *Xpr1$^{sxv}$* at another ECL4 site, D590N (Figure 1A). To assess whether residues at position 590 are responsible for this infectivity difference, we added the mutation N590D to *Xpr1$^{n}$*-Δ583T,K585 Δ (Figure 6, lower right). MoMCF replication was unchanged by this substitution, but FrMCF showed reduced infectivity in the presence of D590, explaining the observed differences between the two sets of constructs.

The data taken together indicate that the XP-MLV subtypes respond differently to ECL4 deletions: X-MLV entry is sensitive to deletions throughout ECL4 when the K500 receptor determinant is not present, the wild mouse viruses are also sensitive to N-terminal ECL4 deletions in the presence of K500, and the two P-MLV isolates differ in sensitivity to ECL4 sequence variants defining two distinct P-MLV host range subgroups.

## Discussion

XPR1 acts as a functional mouse gammaretrovirus receptor in virtually all mammalian species. The scarcity of XPR1 escape mutations lacking all receptor function is due in large part to the fact that virus entry relies on independent receptor determinants in XPR1, and because, as shown here, the two redundant sites for X-MLV entry include one which overlaps a constrained splice donor, while the other determinant, in the short ECL4 loop, is marked by significant sequence variation. Specific residue replacements or deletions in ECL4 can affect but not eliminate all receptor function, and replacement of the entire ECL4 with the corresponding sequence of the jellyfish does not, by itself, abolish receptor function.

It has long been recognized that virus-receptor interfaces are evolutionary battlefields in an “arms race” in which hosts evolve to avoid infection, and viruses evolve to bypass host restrictions. It is clearly disadvantageous to harbor an infectious agent that is pathogenic and mutagenic, and host species can inhibit virus entry through hypomorphic mutations that downregulate receptor expression, through factors that interfere with receptor binding, or through mutations that alter the virus-receptor interface. Virus survival depends on factors that minimize the consequences of the host escape mutations that inhibit replication. At the level of entry, multiple strategies can circumvent host restrictions. One possible survival mechanism is receptor switching, but the repertoire of alternative receptors for

gammaretroviruses is restricted to multipass transmembrane proteins. A second strategy is usage of functionally redundant binding sites, so that single receptor mutations cannot disable receptor function. Thus, for XPR1, separate mutations in the receptor-determining regions of ECL3 or ECL4 do not abolish receptor function, and most mutations with detectable effects restrict some but not all XP-MLVs. Thus, for example, P-MLVs and the wild mouse isolates are sensitive to specific ECL4 deletions, and 579Δ effectively restricts all but CAST-X.

Another virus survival mechanism involves use of receptor determinants that are also important for the protein's normal function. This limits mutations that could decrease the fitness of the host species, and usage of such sites provides viruses with a reliable port of entry. Our previous phylogenetic analyses demonstrated that XPR1 shows signatures of positive selection in rodent and avian lineages (Martin et al., 2013; Yan et al., 2010). While some of these mutational changes may be linked to the host cell function of XPR1, published evidence suggests this protein has roles in phosphate export or in signal transduction (Giovannini et al., 2013; Vaughan et al., 2012). On the other hand, restrictive XPR1 receptors are found in X-MLV-exposed rodent and avian species, and the sites under positive selection in both lineages govern antiviral activity. Although these mutational patterns have no known link to XPR1 function, there is, however, a clear functional constraint on mutational variation in *Xpr1* as the codon for the ECL3 receptor determinant K500 overlaps a splice donor site. The rarity of substitutions in this codon in vertebrate species indicates that this was a fortuitous receptor choice. Of the three naturally occurring replacement mutations at K500, two, K500E and K500Q, are found in mouse and/or avian species exposed to X-MLVs, and both mutations eliminate an XP-MLV receptor determinant. On the other hand, K500R is found in two vertebrate species not likely to have had virus exposure, and this mutation has no influence on receptor function.

In *Mus musculus* subspecies, three restrictive alleles are disabled by ECL4 deletions: $Xpr1^{n}$ in laboratory mice and some *M. m. molossinus*, $Xpr1^{c}$ in *M. m. castaneus* and $Xpr1^{m}$ in *M. m. musculus* and *molossinus* (Bamunusinghe et al., 2013; Kozak, 2010). These 3 deletions remove different residues in an 8 residue stretch of this 13-residue loop. Mutational analysis showed that in the presence of K500, all XP-MLVs except for X-MLVs are sensitive to deletions, especially within the N-terminal end of this loop. In the presence of the K500E mutation, however, deletions throughout ECL4 effectively reduce receptor function for X-MLV and the wild mouse viruses. These data suggest that the length or structure of this loop is important for X-MLV receptor function and that residues nearer the N-terminus are especially important for other XP-MLVs.

The XPR1 ECL4 sequence is highly variable among mammalian orthologs that encode functional receptors. Mutation of $Xpr1^{sxv}$ to evaluate two naturally occurring replacement mutations and the replacement of this loop with that from the nonfunctional Cnardian *XPR1* receptor produced some reductions in receptor function. The functional reduction for MuCcXpr1 could be due to the ECL4 size difference as well as the sequence variation. In the presence of the K500E mutation, however, receptor function was largely eliminated by T583E and the jellyfish replacement. These results establish the importance of specific residues for the ECL4 receptor determinant, and show that while the jellyfish XPR1 does not

contain an ECL4 virus attachment site, the presence of this divergent sequence does not disrupt construction of the K500 receptor site.

Finally, P-MLVs and X-MLVs use different XPR1 receptor determinants (Van Hoeven and Miller, 2005; Yan et al., 2009). Our data show that the Friend and Moloney MCF P-MLVs show distinctive receptor-defined tropisms although both are classed as polytropic with some differences in species host range (Cloyd et al., 1985; Yan et al., 2010). This variation in receptor usage by the polytropic viruses likely results from the fact that infectious P-MLVs are recombinants that contain ERV-derived *env* segments that differ in sequence and size (Chattopadhyay et al., 1982). Deletions of 3 particular ECL4 residues (S578, F584, K585) had profoundly different consequences for infection by these two P-MLVs, as did the jellyfish/mouse chimera, indicating that MoMCF is more accommodating of receptor variation. Further analysis of one of those deletions, K585Δ, showed that this differential infectivity is modulated by residues at ECL4 position 590, a site not previously shown to contribute to receptor function.

## Conclusions

The XPR1 protein functions as a gammaretrovirus receptor in nearly all mammals. This is because it carries 2 functionally independent receptor determinants, one of which overlies a splice donor site and is therefore evolutionarily constrained. The second site lies in a short extracellular loop that shows significant natural sequence variation due to replacement and deletion mutations. The great majority of mutations at one or the other of these receptor-determining regions have minimal effects on receptor function, including replacement of the entire ECL4 with that of the jellyfish XPR1; when both receptor regions are mutated, different mutations inhibit entry of different XP-MLV subtypes.

## Materials and Methods

### Viruses

CAST-X is an X-MLV isolated in our laboratory from the spleen of a CAST/EiJ mouse (Yan et al., 2007). Cz524 is a novel MLV isolated from the spleen of a CZECHII/EiJ mouse 2 months after inoculation with MoMLV (Yan et al., 2009). The human xenotropic-related virus, XMRV (Dong et al., 2007), was kindly provided by R. Silverman (Cleveland Clinic, Cleveland, OH). CasE#1 (Cloyd et al., 1985), the Friend mink cell focus inducing P-MLV (FrMCF), and Moloney MCF P-MLV (MoMCF) were originally obtained from J. Hartley (NIAID, Bethesda, MD).

### Pseudotype assay and Western analysis

Viral pseudotypes carrying the LacZ reporter were generated by infecting GP2-293 cells transfected with pCL-MFG-LacZ with the various XP-MLVs as described previously (Yan et al., 2009).

E36 Chinese hamster cells (Gillin et al., 1972) were transfected with variants of mouse *Xpr1*. Stable transfectants were generated using Fugene6 (Promega, Madison, WI) and selected with geneticin (830 μg/ml) (Corning Cellgro, Manassas, VA). Cells were infected

with dilutions of the pseudotype virus stocks in the presence of 4–8 μg/ml polybrene (Aldrich, Milwaukee, WI). One day after infection, cells were fixed with 0.4% glutaraldehyde and assayed for β-galactosidase activity using as substrate 5-bromo-4-chloro-3-indolyl-β-D-galactopyranoside (X-Gal, 2 mg/ml; ICN Biomedicals, Aurora, Ohio). Infectious titers were determined from the number of blue cells per 50 microliters of virus supernatant.

Proteins were extracted from transfected cells with M-PER Mammalian Protein Extraction Reagent (Pierce, Rockford, IL). Western blots were used to confirm expression of tagged XPR1 using anti-V5 antibody (Invitrogen, Carlsbad, CA) followed by goat anti-mouse IgG conjugated with HRP (Invitrogen). The membrane was then stripped and incubated with mouse anti-α-tubulin (Sigma, St. Louis, Mo) and the goat anti-mouse IgG.

## Cloning of the Cyanea capillata XPR1 sequence

The initial fragment of the jellyfish *(Cyanea capillata) XPR1* was obtained from a cDNA λgt22 phage library made from peri-rhopalial tissues (consisting of neurons, myoepithelial cells, endoderm, and mesoglea) using a degenerate PCR amplification strategy designed to amplify opioid-like and somatostatin receptors (Marchese et al., 1998) (Table 1). BLASTn searches revealed homology to XPR1. Five prime and 3’ ends of the *XPR1* cDNA were PCR amplified using oligonucleotides deduced from the *XPR1* initial fragment, and an oligonucleotide specific to the phage arms (Table 1). The cDNA was initially cloned in-frame into oocyte expression vector EGFP-PXOOM (Jespersen et al., 2002), a gift from D.Y. Boudko (Rosalind Franklin University, Chicago, IL), and was subsequently transferred into the pcDNA3.1/V5-His TOPO (Invitrogen, Carlsbad, CA). RT-PCR was used to confirm expression of the full-length *XPR1* cDNA in the peri-rhopalial tissue. The CcXPR1 cDNA sequence is publicly available in GenBank (accession number: KF638274).

## Generation of *Xpr1* mutants

Variants of the mouse *Xpr1* gene with replacement or deletion mutations were generated by mutagenesis PCR using QuikChange II XL Site-Directed Mutagenesis Kit (Stratagene, La Jolla, CA) using as templates previously described clones of $Xpr1^{n}$-Δ582T and $Xpr1^{sxv}$ (Yan et al., 2007) (Figure 1A). Mutations were produced using the primers and their reverse complements listed in Table 1. Twelve deletion mutations were produced for the 13 residue ECL4 because there are adjacent T residues at positions 582 and 583.

ECL4 (39 bp) of $Xpr1^{sxv}$ was replaced with the corresponding sequence of jellyfish XPR1 (33 bp) using overlap extension PCR. Forward primer A within the vector (Table 1) was used with antisense primer B at the 5’end of ECL4 to amplify upstream sequence from mouse $Xpr1^{sxv}$, and forward primer E at the 5’end of *Xpr1* and downstream vector primer F amplified the downstream sequence. Primers B and E were designed with overlapping CcXPR1 ECL4 overhangs. Primers C and D were then used to amplify CcXPR1 ECL4 with mouse overhangs. Fragments AB and CD were used to generate AD, and fragments AD and EF were used to make AF. Fragment AF and mouse $Xpr1^{sxv}$ were digested with PflMI and EcoRV, which cut at sites that flank ECL4, and the pieces were ligated to create $Xpr1^{sxv}$

carrying CcXPR1 ECL4, termed MuCcXPR1. K500E was introduced into this clone to create MuCcXPR1-K500E.

All mutants were confirmed by sequencing.

## Acknowledgments

This work was supported by the Intramural Research Program of the National Institute of Allergy and Infectious Diseases, NIH.

We thank Alicia Buckler-White and Ronald Plishka for sequencing.

## References

Bamunusinghe D, Liu Q, Lu X, Oler A, Kozak CA. Endogenous gammaretrovirus acquisition in *Mus musculus* subspecies carrying functional variants of the XPR1 virus receptor. J Virol. 2013; 87:9845–9855. [PubMed: 23824809]

Chattopadhyay SK, Cloyd MW, Linemeyer DL, Lander MR, Rands E, Lowy DR. Cellular origin and role of mink cell focus-forming viruses in murine thymic lymphomas. Nature. 1982; 295:25–31. [PubMed: 6276750]

Cloyd MW, Thompson MM, Hartley JW. Host range of mink cell focus-inducing viruses. Virology. 1985; 140:239–248. [PubMed: 2982233]

Dong B, Kim S, Hong S, Das Gupta J, Malathi K, Klein EA, Ganem D, DeRisi JL, Chow SA, Silverman RH. An infectious retrovirus susceptible to an IFN antiviral pathway from human prostate tumors. Proc. Natl. Acad. Sci. U. S. A. 2007; 104:1655–1660. [PubMed: 17234809]

Fischinger PJ, Nomura S, Bolognesi DP. A novel murine oncornavirus with dual ecoand xenotropic properties. Proc. Natl. Acad. Sci. U. S. A. 1975; 72:5150–5155. [PubMed: 1061100]

Gillin FD, Roufa DJ, Beaudet AL, Caskey CT. 8-Azaguanine resistance in mammalian cellsIHypoxanthine-guanine phosphoribosyltransferase. Genetics. 1972; 72:239–252. [PubMed: 4345996]

Giovannini D, Touhami J, Charnet P, Sitbon M, Battini JL. Inorganic phosphate export by the retrovirus receptor XPR1 in metazoans. Cell reports. 2013; 3:1866–1873. [PubMed: 23791524]

Jesperson T, Grunnet M, Angelo K, Klaerke DA, Olesen SP. Vector for protein expression in both mammalian cells and Xenopus laevis oocytes. BioTechniques. 2002; 32:536–540. [PubMed: 11911656]

Hartley JW, Wolford NK, Old LJ, Rowe WP. New class of murine leukemia-virus associated with development of spontaneous lymphomas. Proc. Natl. Acad. Sci. U. S. A. 1977; 74:789–792. [PubMed: 191826]

Kozak CA. The mouse "xenotropic" gammaretroviruses and their XPR1 receptor. Retrovirology. 2010; 7:101. [PubMed: 21118532]

Kozak CA. Evolution of different antiviral strategies in wild mouse populations exposed to different gammaretroviruses. Current opinion in virology. 2013; 3:657–663. [PubMed: 23992668]

Kozak CA, O'Neill RR. Diverse wild mouse origins of xenotropic, mink cell focus-forming, and two types of ecotropic proviral genes. J Virol. 1987; 61:3082–3088. [PubMed: 3041030]

Levy JA. Xenotropic viruses: murine leukemia viruses associated with NIH Swiss, NZB, other mouse strains. Science. 1973; 182:1151–1153. [PubMed: 4356281]

Marchese A, Nguyen T, Malik P, Xu S, Cheng R, Xie Z, Heng HH, George SR, Kolakowski LF Jr, O'Dowd BF. Cloning genes encoding receptors related to chemoattractant receptors. Genomics. 1998; 50:281–286. [PubMed: 9653656]

Marin M, Tailor CS, Nouri A, Kozak SL, Kabat D. Polymorphisms of the cell surface receptor control mouse susceptibilities to xenotropic and polytropic leukemia viruses. J Virol. 1999; 73:9362–9368. [PubMed: 10516044]

Martin C, Buckler-White A, Wollenberg K, Kozak CA. The avian XPR1 gammaretrovirus receptor is under positive selection and is disabled in bird species in contact with virus-infected wild mice. J Virol. 2013; 87:10094–10104. [PubMed: 23843647]

Mount SM. A catalogue of splice junction sequences. Nucleic Acids Res. 1982; 10:459–472. [PubMed: 7063411]

Van Hoeven NS, Miller AD. Use of different but overlapping determinants in a retrovirus receptor accounts for non-reciprocal interference between xenotropic and polytropic murine leukemia viruses. Retrovirology. 2005; 2:76. [PubMed: 16354307]

Vaughan AE, Mendoza R, Aranda R, Battini JL, Miller AD. Xpr1 is an atypical G-protein-coupled receptor that mediates xenotropic and polytropic murine retrovirus neurotoxicity. J Virol. 2012; 86:1661–1669. [PubMed: 22090134]

Xu WQ, Eiden MV. Primate Gammaretroviruses Require an Ancillary Factor Not Required for Murine Gammaretroviruses To Infect BHK Cells. Journal of Virology. 2011; 85:3498–3506. [PubMed: 21270153]

Yan Y, Knoper RC, Kozak CA. Wild mouse variants of envelope genes of xenotropic/polytropic mouse gammaretroviruses and their XPR1 receptors elucidate receptor determinants of virus entry. J Virol. 2007; 81:10550–10557. [PubMed: 17634227]

Yan Y, Liu Q, Kozak CA. Six host range variants of the xenotropic/polytropic gammaretroviruses define determinants for entry in the XPR1 cell surface receptor. Retrovirology. 2009; 6:87. [PubMed: 19811656]

Yan Y, Liu Q, Wollenberg K, Martin C, Buckler-White A, Kozak CA. Evolution of functional and sequence variants of the mammalian XPR1 receptor for mouse xenotropic gammaretroviruses and the human-derived XMRV. J Virol. 2010; 84:11970–11980. [PubMed: 20844050]

## Highlights

XPR1 functions as a gammaretrovirus receptor in most mammals.

One critical receptor determinant overlies a splice donor and is highly conserved.

The second receptor determining loop is largely unaffected by single deletions.

This loop is also unaffected by small or large replacements.

These determinants severely limit the likelihood of host escape mutations.

**A**

| | C-Term ECL3 | ECL4 | Virus Susceptibility Patterns X-MLV | P-MLV | CasE#1 | Cz524 |
|---|---|---|---|---|---|---|
| *Xpr1*$^{sxv}$ | THKEQNHSDTV | SITATTFKPHVGD | ++ | ++ | ++ | ++ |
| *Xpr1*$^{n}$ | ..E..... | ....-.....N | – | ++ | – | – |
| *Xpr1*$^{c}$ | ..... | ....-----.... | ++ | – | – | + |
| *Xpr1*$^{m}$ | ..... | .-...K..... | ++ | – | – | – |
| *Xpr1*$^{n}$-Δ582T | ..E..... | .....N | ++ | ++ | ++ | ++ |
| *Xpr1*$^{n}$-E500K | ..... | ....-.....N | ++ | ++ | – | ++ |
| | 500 | 578 590 | | | | |

**B**

| | | -4 | -3 | -2 | -1 | EXON +1 | +2 | +3 | +4 | +5 | +6 | INTRON +7 |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Consensus | | | A/C | A | G | G | T | A | A | G | T | |
| | %A | 30 | 40 | 64 | 9 | 0 | 0 | 62 | 68 | 9 | 17 | 39 |
| | %T | 20 | 7 | 13 | 12 | 0 | 100 | 6 | 12 | 5 | 63 | 22 |
| | %C | 30 | 43 | 12 | 6 | 0 | 0 | 2 | 9 | 2 | 12 | 21 |
| | %G | 19 | 9 | 12 | 73 | 100 | 0 | 29 | 12 | 84 | 9 | 18 |
| Mammals | C57BL | G | A | A | G | G | T | A | T | G | C | T |
| | | | Glu | | | | | | | | | |
| | *M. dunni* | A | A | A | G | | | | | | | |
| | | | Lys | | | | | | | | | |
| | Human | A | A | A | G | G | T | A | T | T | C | T |
| | | | Lys | | | | | | | | | |
| | Pika | A | G | A | G | G | T | A | T | T | T | T |
| | | | Arg | | | | | | | | | |
| Other vertebrates | Anole lizard | A | G | A | G | G | T | A | T | T | G | T |
| | | | Arg | | | | | | | | | |
| | Duck | A | A | A | G | G | T | A | C | G | G | G |
| | | | Lys | | | | | | | | | |
| | Chicken | C | A | A | G | G | T | A | C | T | G | G |
| | | | Gln | | | | | | | | | |
| | Gray junglefowl | G | A | A | G | G | T | A | C | T | G | G |
| | | | Glu | | | | | | | | | |

**Figure 1.**

Sequence variation at receptor critical sites in XPR1. A) Protein sequence in the C-terminal end of ECL3 and ECL4 in fully susceptible *Xpr1*$^{sxv}$ and in *Mus musculus* variants carrying ECL4 deletions. Virus infectivity patterns are shown for each naturally occurring receptor and for *Xpr1*$^{n}$ mutants with corrections at the two X-MLV receptor sites. $Log_{10}$ titer: +, 1–2; ++, >3. B) DNA sequence spanning the exon/intron junction at the end of XPR1 exon 11. Nucleotide distribution in the splice donor consensus site is from (Mount, 1982). Sequences are shown for variants in mammals and other vertebrates.

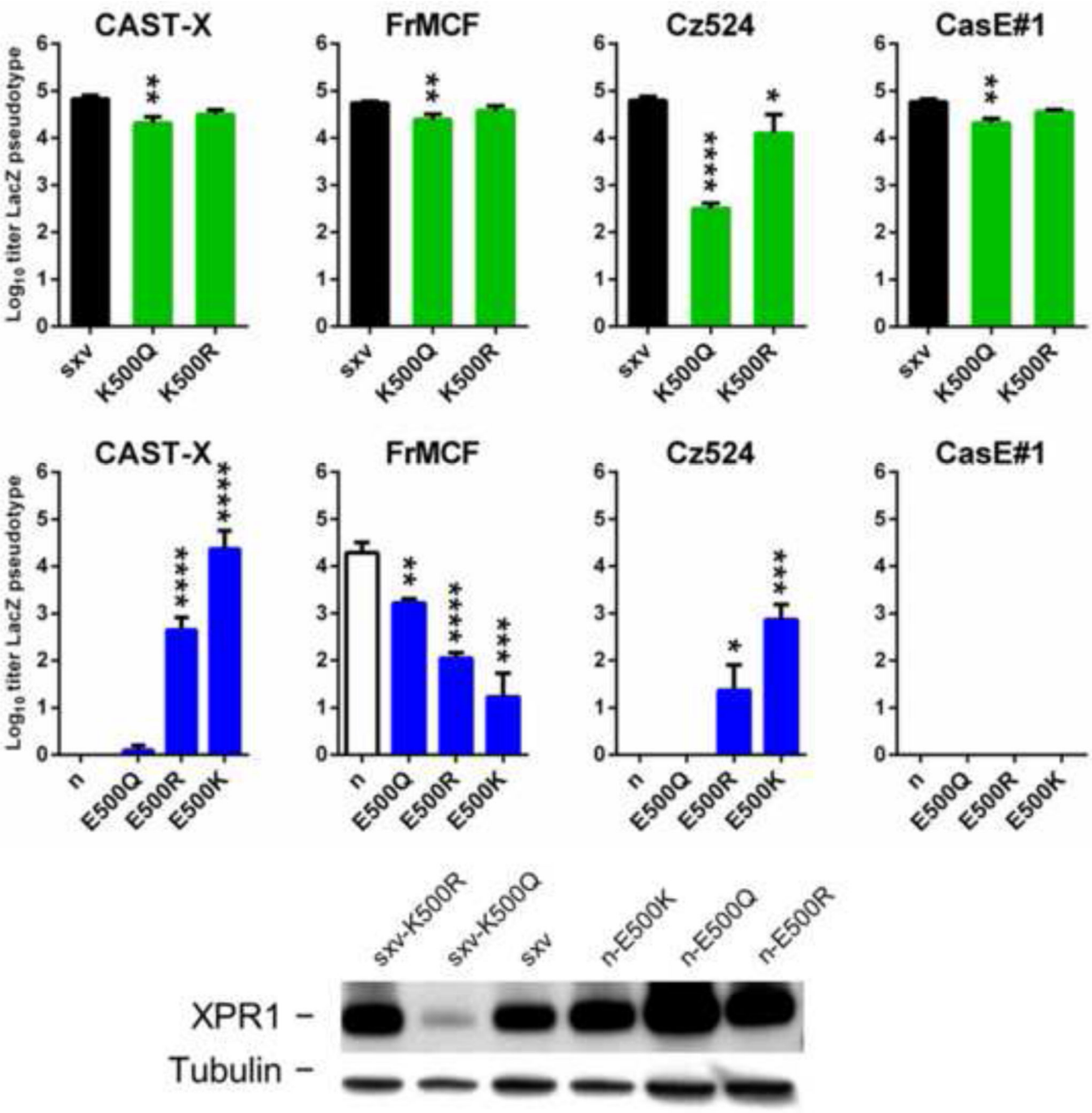


**Figure 2.**

Susceptibility of E36 hamster cells expressing *Xpr1* ECL3 mutants to LacZ XP-MLV pseudoviruses carrying the indicated Env glycoproteins. Titers represent blue cells in 50 μl of virus stock and are presented as the means of 3–5 tests plus SEM. The asterisks indicate significant P values using Student's t test (*, $p \leq 0.05$; **, $p \leq 0.01$; ***, $p \leq 0.001$; ****, $p \leq 0.0001$). Untransfected E36 cells show trace levels of susceptibility to CAST-X ($\log_{10}$ = 0.2 +/− 0.3). Graphs at the top show mutants introduced into $Xpr1^{sxv}$, with transfectants expressing the *sxv* wild type in black. Graphs at the bottom show mutants of $Xpr1^{n}$ mutants,

and the *n* wild type is in white. Westerns showing V5-tagged receptor expression in each transfectant are at the bottom.

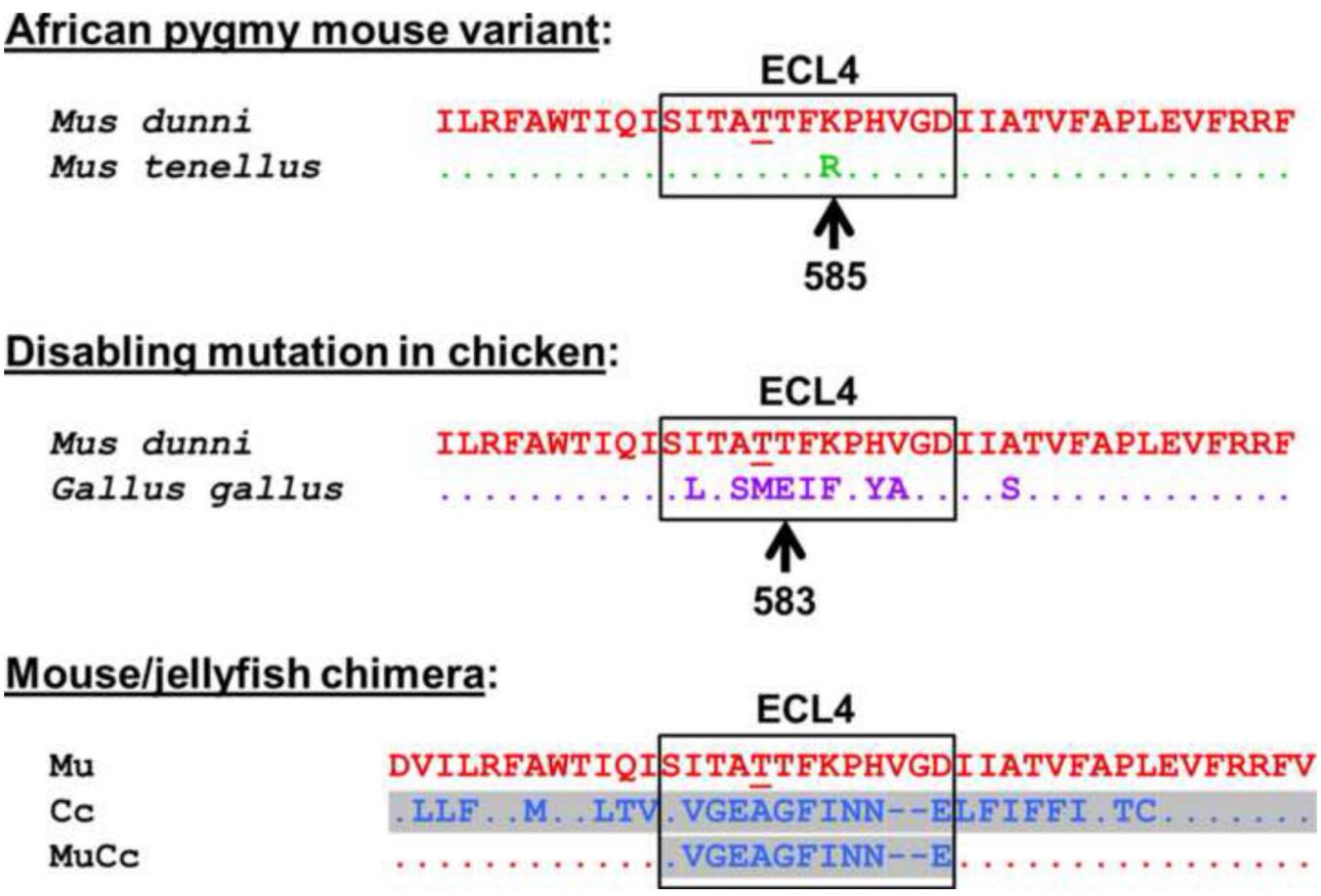


**Figure 3.**

Naturally occurring ECL4 mutations introduced into *Xpr1*$^{sxv}$. K585R was found in the African pygmy mouse, *M. tenellus*, and T583E is equivalent to the Q579E mutation in the restrictive chicken XPR1. At bottom is a chimera in which the ECL4 of *Xpr1*$^{sxv}$ is replaced by the homologous segment from the jellyfish, *C. capillata*.

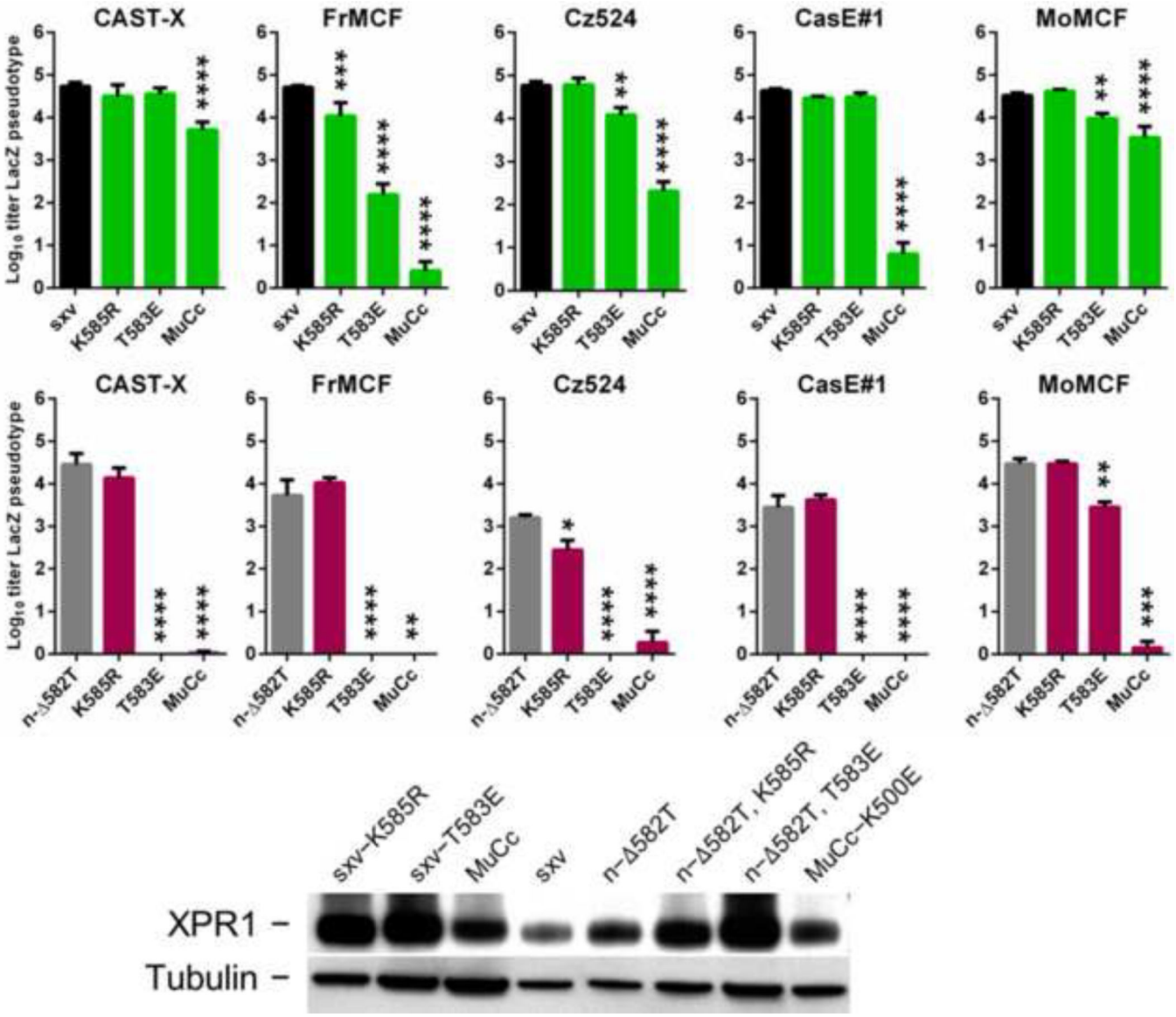


**Figure 4.**

Susceptibility of E36 hamster cells expressing *Xpr1* ECL4 replacement mutants to LacZ XP-MLV pseudoviruses. Titers represent blue cells in 50 μl of virus stock and graphs were produced as in Figure 2. Untransfected E36 cells are poorly infectious for CAST-X ($log_{10}$ = 0.4 +/− 0.4). Graphs at the top show mutations introduced into $Xpr1^{sxv}$. Graphs at the bottom show $Xpr1^{n}$-Δ582T mutants. At the bottom are westerns showing receptor expression.

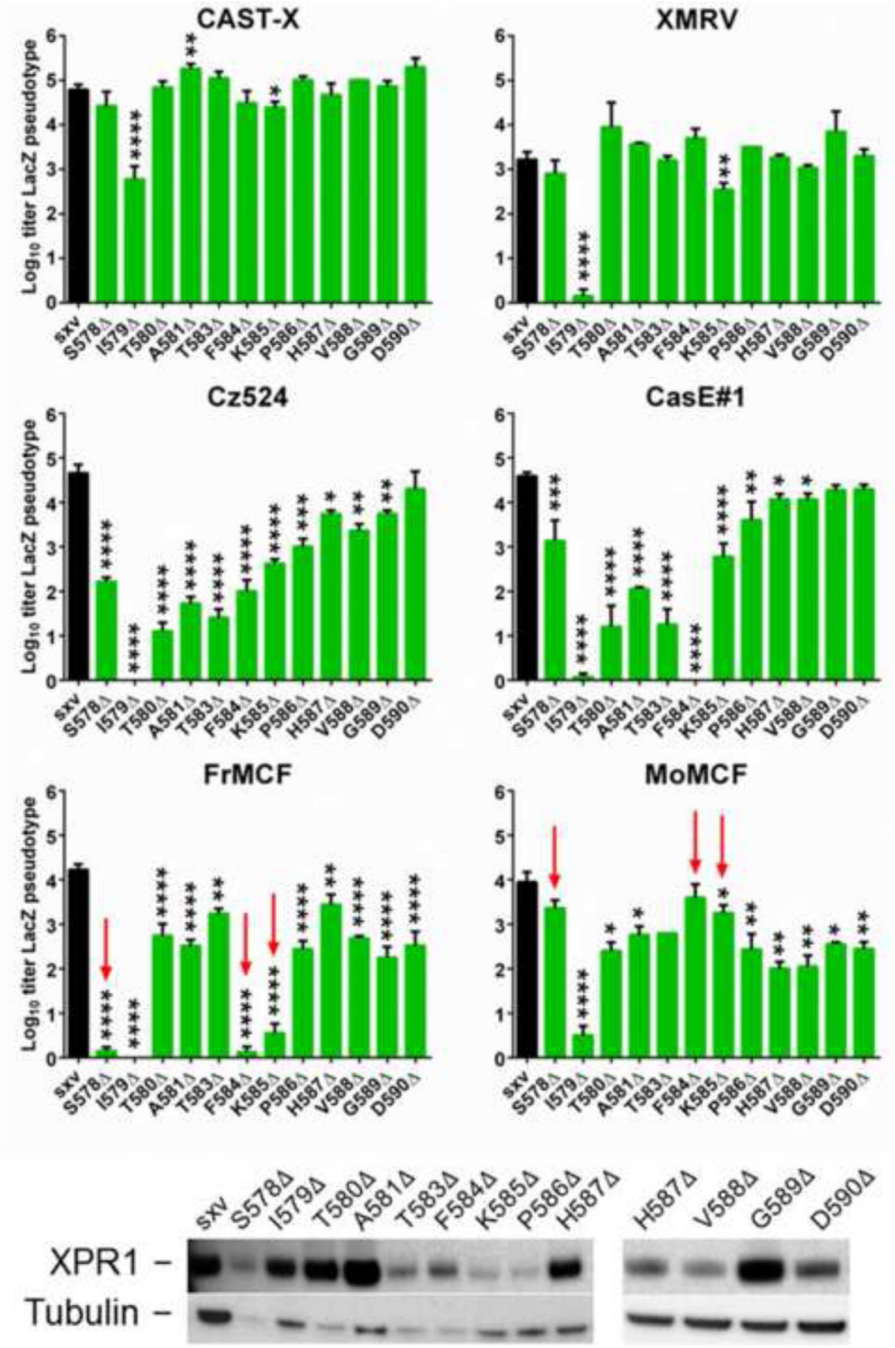


**Figure 5.**

Susceptibility of E36 hamster cells expressing *Xpr1*$^{sxv}$ ECL4 deletion mutants to LacZ XP-MLV pseudoviruses. Untransfected E36 cells are poorly infectious for CAST-X ($log_{10}$ = 1.2 +/− 0.6). Graphs were produced as for Figure 2. Red arrows mark titer differences between FrMCF and MoMCF. At the bottom are westerns showing XPR1 expression.

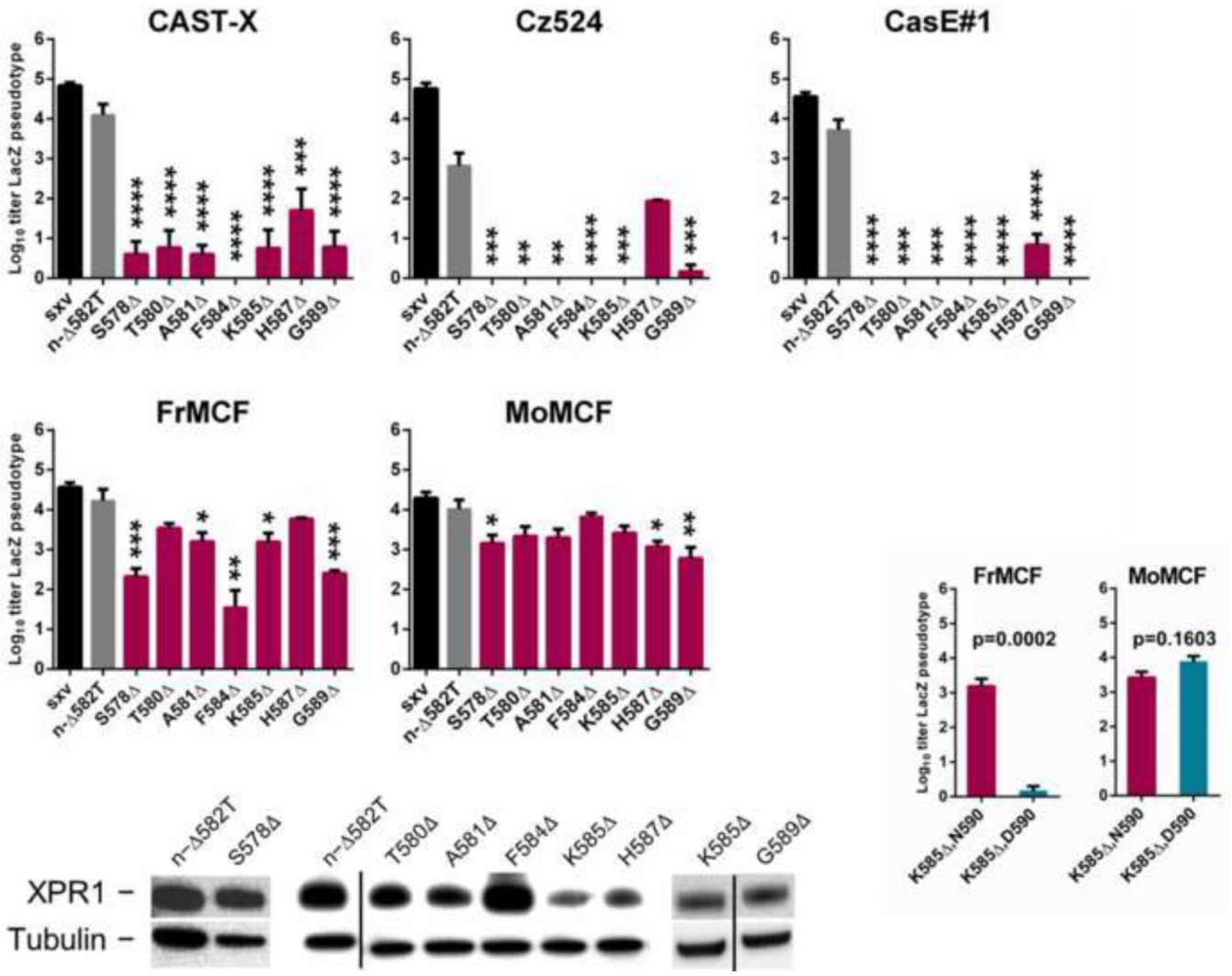


**Figure 6.**

Susceptibility of E36 hamster cells expressing *Xpr1* ECL4 mutants to LacZ XP-MLV pseudoviruses. Mutations were done in *Xpr1*$^{n}$-Δ582T which carries the ECL3 mutation K500E. Graphs were produced as for Figure 2. The extra panel on the lower right shows the effect of residues at position 590 on P-MLV infection. Westerns are presented on the lower left.

## Table 1

Primers for cloning and mutagenesis.

| | |
|---|---|
| **Deletion Mutations*** | |
| S578Δ | CGCTTGGACTATCCAAATCATTACTGCTACAACG |
| I579Δ | GACTATCCAAATCTCTACTGCTACAACGTTTAAGC |
| T580Δ | CGCTTCGCTTGGACTATCCAAATCTCTATTGCTACAACG |
| A581Δ | CGCTTGGACTATCCAAATCTCTATTACTACAACGTTTAAGCC |
| T583Δ | CAAATCTCTATTACTGCTACATTTAAGCCTCATGTTGGG |
| F584Δ | CCAAATCTCTATTACTGCTACAACGAAGCCTCATGTTGGGG |
| K585Δ | CTATTACTGCTACAACGTTTCCTCATGTTGGGGACATCATTGC |
| P586Δ | CTATTACTGCTACAACGTTTAAGCATGTTGGGGACATCATTGC |
| H587Δ | GCTACAACGTTTAAGCCTGTTGGGGACATCATTGC |
| V588Δ | GCTACAACGTTTAAGCCTCATGGGGACATCATTGCTACTG |
| G589Δ | CGTTTAAGCCTCATGTTGACATCATTGCTACTG |
| D590Δ | GCTACAACGTTTAAGCCTCATGTTGGGATCATTGCTACTG |
| K585Δ (in *Xpr1*$^{n}$) | CTATTACTGCTACAACGTTTCCTCATGTTGGGAACATCATTGC |
| H587Δ (in *Xpr1*$^{n}$) | GCTACAACGTTTAAGCCTGTTGGGAACATCATTGC |
| G589Δ (in *Xpr1*$^{n}$) | CGTTTAAGCCTCATGTTAACATCATTGCTACTGTCTTTGCC |
| **Replacement Mutations*** | |
| K500R | GCCCTTTACAGCACTCACAGAGAACAAAACCACTC |
| T583E | CTCTATTACTGCTACAGAGTTTAAGCCTCATGTTGGGGAC |
| K500Q | GCCCTTTACAGCACTCACCAAGAACAAAATCACTC |
| K500E | GCCCTTTACAGCACTCACGAGGAACAAAATCACTC |
| K585R | ACGTTTAGACCTCATGTTGGGG |
| **Cyanea XPR1 cloning** | |
| TM2 | ATCYTCAACCTKGCYMTSGCMGA |
| TM7 | CAGGAAGGCGTARAGRAMKGGRTT |
| CcXPR1-HindIII | GGGGAAGCTTAAGTTCACAGAACACTTAGGTGCACAC |
| CcXPR1-NotI | GGGGGCGGCCGCTCATCAAACGGTTCTAACAGCGCCATTTTGCTGCG |
| **Mouse/JellyfishChimera**** | |
| A | GCTGGAGTAAATCATGTCCTC |
| B | caccgaccgaGATTTGGATAGTCCAAGCGAA |
| C | ATCCAAATCtcggtcggtgaggctggc |
| D | GGGCAAAGACAGTAGCAATGATttcgttgttgatgaagcc |
| E | caacaacgaaATCATTGCTACTGTCTTTGCCCCCCTTGAGG |
| F | CAGAAGCCATAGAGCCCACCGCATC |

* Reverse primers were reverse complements.

** Jellyfish sequence shown in lowercase letters.