TITLE

Systematic genetic characterization of the human PKR kinase domain highlights its functional malleability to escape a viral substrate mimic

AUTHORS

Michael J. Chambers (1,2), Sophia Scobell (1), Meru J. Sadhu (1)

AFFILIATIONS

1. Center for Genomics and Data Science Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA
2. Department of Microbiology & Immunology, Georgetown University, Washington DC, USA

ABSTRACT

Evolutionary arms races can arise at the contact surfaces between host and viral proteins, producing dynamic spaces in which genetic variants are continually pursued. However, the sampling of genetic variation must be balanced with the need to maintain protein function. A striking case is given by protein kinase R (PKR), a member of the mammalian innate immune system. PKR detects viral replication within the host cell and halts protein synthesis to prevent viral replication by phosphorylating eIF2α, a component of the translation initiation machinery. PKR is targeted by many viral antagonists, including poxvirus pseudosubstrate antagonists that inhibit PKR by interacting with the same binding surface as eIF2α. Remarkably, PKR has several rapidly evolving residues at this interface, suggesting it is engaging in an evolutionary arms race, despite the surface's critical role in phosphorylating eIF2α. To systematically explore the evolutionary opportunities available at this dynamic interface, we generated and characterized a library of 426 SNP-accessible nonsynonymous variants of human PKR for their ability to escape inhibition by the model pseudosubstrate inhibitor K3 from vaccinia virus. We identified key sites in the PKR kinase domain that harbor K3-resistant variants, as well as critical sites where variation leads to loss of function. We find K3-resistant variants are readily available throughout the interface and are enriched at sites under positive selection. Moreover, variants beneficial against K3 were also beneficial against an enhanced variant of K3, indicating resilience to viral adaptation. Overall, we find that the eIF2α-binding surface of PKR is highly malleable, potentiating its evolutionary ability to combat viral inhibition.

INTRODUCTION

Molecular interfaces between interacting host and pathogen proteins are atomic battlegrounds that give rise to evolutionary arms races (1). Successful infection typically involves critical interactions between host and pathogen proteins, such as binding to cell-surface receptors or inhibition of cellular defense mechanisms. In such scenarios, the host benefits from genetic mutations that disrupt the interaction while the pathogen benefits from mutations that maintain the interaction. In an evolutionary arms race the accumulation of mutations e can lead to abnormally high sequence diversity, which is used to identify sites experiencing positive selection.

A key member of the antiviral innate immune repertoire in vertebrates is protein kinase R (PKR) (2–4). PKR has two domains, a double-stranded RNA (dsRNA) binding domain and a kinase domain, separated by a long linker (5). The presence of dsRNA in the cytoplasm is a hallmark of viral infection (6). PKR binds dsRNA, initiating PKR homodimerization and autophosphorylation which activates the kinase (7–10). Active PKR binds and phosphorylates eIF2α on Ser51 (Figure 1A), a member of the translation initiation machinery. Phosphorylation of eIF2α leads to a halt in protein synthesis, thereby preventing viral replication (8,11,12).

Many viruses encode PKR inhibitors to evade its antiviral activity (13–15). One example is vaccinia virus, which encodes the well-characterized pseudosubstrate inhibitor protein K3 (Figure 1B) (16–18). K3 competitively inhibits eIF2α from binding PKR through its sequence and structural similarity to eIF2α (17,19–21). Vertebrate PKR homologs have many sites under positive selection, which suggests an evolutionary arms race between PKR and its inhibitors (22–24). Surprisingly, this includes PKR residues at its binding surface with eIF2α (Figure 1C) even though this interface is essential to PKR function (22,23), which could reflect evolutionary pressure from viral pseudosubstrate inhibitors. Evolutionary pressures on PKR would require that it recognizes eIF2α while evading pseudosubstrate antagonists at the same interface.

Here we use a prospective, systematic, and unbiased approach to map SNP-accessible PKR variants in the kinase domain that are resistant or susceptible to the K3 pseudosubstrate antagonist. We found many K3-resistant variants, especially at sites under positive selection. Essential PKR residues that did not tolerate variation were infrequent despite the critical nature of this interface; they were primarily found surrounding the ATP-binding site. These results paint a portrait of PKR balancing the tension between conserved function and evasion of viral inhibitors through mutational resilience in its substrate interface.

RESULTS

Generation of nonsynonymous PKR variant library

We systematically made genetic variants in four windows of interest in the kinase domain of PKR (Figure 1D). The four windows were selected based on having sites within five angstroms of eIF2α or K3 and sites under positive selection across vertebrate PKR homologs (22). Our windows of interest include helix αD and αG, as well as the glycine-rich loop and activation loop surrounding the ATP-binding site. In these windows, we focused on testing 426 variants that could be acquired with a single nucleotide change, as those are the most likely to be evolutionarily sampled by PKR.

Overexpression of PKR in the budding yeast *Saccharomyces cerevisiae* is toxic, as PKR auto-activates and shuts down protein synthesis by phosphorylating yeast eIF2α, which is highly conserved between yeast and humans (25,26). This toxicity can be partially counteracted by co-expression of vaccinia K3 (21). Thus, yeast provides a useful experimental system to test the functionality of PKR variants: variants that maintain the ability to phosphorylate eIF2α while escaping K3 inhibition will further arrest growth, while variants that lose eIF2α kinase activity or become more susceptible to K3 will improve growth.

We sought to leverage this assay in a high-throughput manner to determine the functionality of PKR variants by tracking their abundances in a pool over time. We generated the

426 single-residue variants of PKR on a single-copy plasmid with a galactose-inducible promoter in a plasmid pool (Figure 1E). Each unique plasmid in the pool carried a DNA barcode randomly generated during cloning. We used long-read circular consensus sequencing to pair DNA barcodes with PKR variants, identifying an average of 43 barcodes per variant (Supplemental Figure 1). We then created three libraries by combining the PKR variant library with wild-type K3, nonfunctional K3, or an enhanced mutant allele of K3 (21,27). The K3 alleles were expressed from a constitutive *TDH3* promoter. We transformed the three PKR-K3 paired libraries in duplicate into yeast grown in glucose media, such that PKR expression was repressed. We then transferred to liquid galactose media to induce PKR expression, and sampled cells at 0, 12, 16, and 20 hours post-inoculation. We quantified the abundance of the PKR variants at each timepoint using high-throughput short-read sequencing of the DNA barcodes. For each PKR variant we calculated a functional score against each K3: high scores corresponded to PKR variants that evaded K3 (identified by decreased abundance in the pool over time) and low scores corresponded to PKR variants that were either non-functional or well-inhibited by K3 (increased abundance in the pool over time; Supplemental Figure 2). The correlation of functional scores between replicates 1 and 2 was very strong (Pearson correlation coefficient > 0.98 for each K3 allele, Supplemental Figure 3), so we used PKR functional scores generated by combining the reads from replicates 1 and 2 for subsequent analyses.

Many PKR variants evade K3 while maintaining kinase function

We first examined how each PKR variant performed against wild-type K3. We were encouraged to find that variants previously characterized as resistant to K3-His47Arg (hereafter referred to as K3-H47R) (28) were resistant to wild-type K3 in our screen (Figure 2A). We found many additional K3-resistant variants across all four tested windows. Notably, every variant made at Glu375 increased evasion of K3, with the exception of Glu375Asp. Glu375 forms an attractive electrostatic interaction with Lys45 of K3 in the AlphaFold2-predicted complex (Figure

2B). This interaction would be disrupted by all tested variants at residue 375 aside from the negatively charged aspartic acid, consistent with our results. Interestingly, Glu375 is the only site under positive selection in vertebrates that was proximal to K3 but not eIF2α in the AlphaFold2 models, making it a prime candidate for disrupting K3 inhibition without affecting eIF2α phosphorylation.

We found a large number of K3-resistant variants in the vicinity of PKR's helix αG (Figure 2C,D). Compared to other protein kinase structures, eIF2α kinases have a noncanonical helix αG that is extended by one full turn and rotated 40° counterclockwise with a 5 Å translation relative to its C-terminus, owing to a reduced αF-αG linker (5). The noncanonical orientation of helix αG contributes to both eIF2α binding specificity and Ser51 presentation to PKR (5). It has been proposed that mutations causing movement in helix αG might be especially capable of disrupting K3 binding due to the rigid structure of K3 between its contact points on helix αG and the PKR active site. In contrast, these mutations would be tolerated in the context of eIF2α binding because eIF2α can accommodate helix αG movement due to its flexible loop that contacts PKR's active site (22). We observe a gradient of PKR functional scores along the helix αG region, starting with highly resistant or susceptible variants (positions 480-492) and shifting to neutral wild-type like scores (493-500) (Figure 2A). We identified four sites within or near the N-terminus of helix αG for which many variants evade K3: Glu480, Val484, Glu490, and Thr491 (Figure 2C).

We looked at what features of PKR were associated with sites of K3-resistant variants and were excited to find that sites under positive selection in vertebrates were strongly enriched for K3-resistant variants relative to sites conserved across vertebrate PKRs (Figure 2E). Comparing K3 contact and non-contact sites, we find significantly lower functional scores at K3 contact sites (Figure 2F). However, functional scores at K3 contact sites form a bimodal distribution (Hartigan's dip test statistic = 0.105, $p < 0.0001$). Considering only functional variants (PKR functional score > -2) we find that variants at K3 contact sites have significantly

higher functional scores than those at non-contact sites (Supplemental Figure 4). Thus, variants at K3 contact sites tend towards large effects either improving or diminishing PKR function versus K3.

Select sites are essential for PKR function

To understand the constraint placed on the PKR kinase domain, we examined variants that failed to maintain PKR function. We screened our PKR variant library against a nonfunctional K3 (hereafter referred to as K3Δ58) in which the first 173 bases were deleted, removing the first 58 residues, including the initial methionine. When paired with K3Δ58, most PKR variants were comparable to wild-type PKR (Figure 3A), suggesting the tested regions were not highly functionally constrained for kinase activity despite their proximity to eIF2α during its phosphorylation. The few highly constrained sites include Gly276 and Gly277, which compose the glycine-rich loop above the ATP-binding site; Gly450 and Thr451 in the activation loop; and Thr487 in helix αG, which stabilizes helix αG (5) and induces the conformational change in eIF2α required for its phosphorylation (Figure 3B) (29). We also found proline variants to generally disrupt PKR function (Figure 3C). We found nonfunctional variants of varying degrees in all four windows of interest, but with very few in helix αD and the C-terminus of helix αG (Figure 3C), consistent with these regions having pliable secondary structures in which variants can evade K3 while not disrupting eIF2α recognition.

More PKR variants had low functional scores in the presence of functional K3 (Figure 2A) than in its absence (Figure 3C). We plotted the functional scores against each other (Figure 4A), and found that variants with low PKR functional scores versus K3 separated into two classes: some switched to high functional scores in the absence of K3, indicating they caused enhanced susceptibility to K3 but otherwise retained PKR function, while others had low functional scores in both conditions, indicating they had diminished eIF2α kinase activity irrespective of K3 antagonism. We quantified this effect by calculating residuals from the line

connecting PKR nonsense variants to PKR-WT, which highlights variants whose functional scores diverged between the K3 conditions. We find that the activation loop and the start of helix αG, which were less tolerant of mutation when paired with wild-type K3 (Figure 2D), contain a mix of variants that lose function outright or that increase susceptibility to K3 (Figures 3B-C, 4B-C). Variants at five sites were especially susceptible to K3 inhibition: Phe278, Thr373, Arg453, Phe489, and Ser492 (Figure 4B). None appear under positive selection across vertebrates, and Phe278 is conserved across vertebrate PKR homologs. Interestingly, Phe278 is predicted to form a pi interaction with His47 of K3, which aligns to Arg52 of eIF2α (Supplemental Figure 5) (21). Mutation of K3-H47 to arginine, matching eIF2α, is known to improve K3's antagonism of PKR (21,27), suggesting PKR can discriminate against K3 based on its difference from eIF2α at this position. This is consistent with our observation of the importance of Phe278 specifically in the presence of K3.

K3-resistant variants were largely resistant to enhanced K3-H47R

We also characterized PKR variants in the presence of an enhanced mutant allele of K3, K3-H47R (Figure 5). This mutant allele was first identified in a genetic screen of K3 mutants as an enhanced antagonist of PKR (21) and independently identified in a directed evolution screen (27). We sought to understand the resilience of improved PKR variants to alternative K3 alleles by examining the concordance between beneficial PKR variants against wild-type K3 and K3-H47R. We found strong correlation between PKR functional scores versus K3 and K3-H47R (Spearman correlation = 0.832, Figure 5A). Indeed, many of the K3-resistant variants are also resistant to K3-H47R (Figure 5B-D), suggesting resilience of PKR to allelic variation in K3. Still, there were some differences in the patterns of resistance to the two K3 alleles. Surprisingly, positively selected PKR sites did not have significantly higher PKR functional scores versus K3-H47R than conserved sites (Figure 5E), a sharp contrast to the PKR functional scores versus wild-type K3 (Figure 2E). We also noted that variants at sites proximal to K3 had slightly

higher functional scores than those at non-contact sites (Figure 5F), the opposite of what we observed against wild-type K3 (Figure 2F). This could result from wild-type-like PKR variants having little more function than nonsense-like variants in the context of enhanced inhibition by K3-H47R, as wild-type-like variants are enriched at non-contact sites.

The correlation of PKR functional scores between wild-type K3 and K3-H47R is non-linear due to compression between wild-type and nonsense PKR variants in the K3-H47R condition. We fitted a nonlinear exponential function and calculated residuals (Figure 6A), yielding three sites whose variants deviated most strongly between the wild-type K3 and K3-H47R conditions: Arg382, Met455, and Glu480 (Figure 6B,C). We note that Glu480 harbors K3-resistant variants, yet variants at this site appear to be even more impactful at evading the enhanced K3-H47R mutant. All three sites harbored variants that are susceptible to wild-type K3 but resistant to K3-H47R (Figures 2C, 5C). Interestingly, these three sites were not proximal to His47 of K3 in the predicted binding surface, suggesting they acted allosterically. Overall, the strong correlation of functional scores versus K3 and K3-H47R suggests K3-resistant variants would often be resilient against alternate K3 alleles.

Previous work identified twelve PKR variants with enhanced resistance to K3-H47R (28), nine of which were included in our experiment (Supplemental Figure 6). We examined the other nonsynonymous variants introduced at these sites and found that most sites contained additional novel resistant variants, indicating the previously uncovered individual mutants reflect sites of general opportunity for pseudosubstrate evasion.

DISCUSSION

Here we characterize the local evolutionary space available to the human innate immunity protein PKR to maintain functionality while evading the well-studied antagonist, K3. One strategy viruses leverage against PKR is pseudosubstrate inhibition, in which proteins like K3 from vaccinia virus bind PKR by structurally mimicking its natural substrate eIF2α, preventing

PKR kinase activity. Interestingly, this binding interface contains many residues under positive selection. However, positively selected genetic variants across the binding interface would need to maintain binding with PKR's natural substrate while evading viral pseudosubstrate antagonists. In this study, we quantified the abundance of such "tightrope-walking" genetic variants relative to the abundance of functionally deleterious variants, which highlighted particular sites of opportunity and constraint within the kinase domain.

We identified many sites containing variants that evade K3 without losing eIF2α targeting. Resistant variants clustered around helices αD and αG, secondary structures with sites under positive selection and where resistant variants have been previously characterized (22,28). Overall, we found few genetic variants render the PKR kinase domain nonfunctional (Figure 3), underscoring the genetic resilience of PKR, with functionally constrained sites mostly localized to the vicinity of the ATP-binding site. Finally, we found a strong correlation between PKR variants that are resistant against both K3 and the enhanced antagonist K3-H47R, suggesting the K3-resistant variants are resilient to K3 allelic variation.

We identified sites across the PKR kinase domain that were K3-resistant, K3-susceptible, or essential to eIF2α phosphorylation (Figure 7). Sites enriched with genetic variants that improve evasion of K3 were classified as K3-resistant sites, such as Glu375 and Glu480. We find clusters of these sites located across helices αD and αG, at the binding surface of K3 and eIF2α. Positive selection sites are enriched for K3-resistant variants. There were sites under positive selection where we did not observe much K3 resistance, such as sites 269-272, which could reflect that a variety of viral inhibitors with different binding surfaces may have driven PKR evolution. Still, we found these sites to be generally able to adopt variants that maintain PKR function (Figure 3), making them prime "tightrope-walking" sites for thwarting any alternate viral inhibitors that interact at those sites. We also identified some K3-resistant sites that are not under positive selection, such as Glu480, which is in fact conserved across vertebrate PKR homologs and the four human eIF2α kinases (9,22). Perhaps it has an

important role in a function not tested in our assay, such as avoiding off-target phosphorylation of other human proteins.

We found PKR variants that were not tolerated, though perhaps fewer than we would expect given the kinase domain is responsible for carrying out its critical function. These came in two flavors: variants that increased susceptibility to K3, and variants that were deleterious for PKR function. We were surprised to find sites enriched for K3-susceptible variants, such as Phe278, Arg453, and Ser492. These could be sites where the wild-type residues help discriminate between eIF2α and K3. These sites do not cluster together; however, when projected onto the surface of PKR they appear to be intermingled with essential residues (Figure 7). This could suggest that while they are not needed to recognize eIF2α in normal conditions, they are well-positioned to discriminate between eIF2α and a pseudosubstrate inhibitor. Finally, the correlation of PKR functional scores between K3 and K3-H47R suggests resistant variants are resilient. This supports prior work screening a few PKR variants against vaccinia K3-H47R and the variola virus homolog C3 (22). We would be curious to know if this pattern holds for more highly diverged K3 orthologs (30) and other pseudosubstrate inhibitors, such as ranavirus vIF2α (31).

The genetic pliability observed in the PKR kinase domain is reminiscent of that seen with another innate immunity factor, TRIM5α, which is locked in an evolutionary arms race with retroviral capsids (32,33). Both PKR and TRIM5α have access to “rolling hills" of resistant variants (Supplemental Figure 7), in opposition to sharp cliffs that are highly optimized (33). Notably, unlike PKR, TRIM5α does not have to balance pathogen conflict with enzymatic function. We also considered contrasts between PKR and ACE2, a host enzyme with sites under positive selection likely due to being targeted by coronaviruses for cell entry (34,35). However, the ACE2 sites under positive selection are not located near the catalytic domain of the enzyme, in contrast to PKR (36–38). Ultimately, it appears that the resolution of this conundrum is that PKR maintains pliability in this region despite its criticality. This pliability is

likely made possible by the flexibility of the eIF2α's Ser51 loop in which the phosphorylation site resides; indeed, the loop is unstructured in the co-crystal of PKR and eIF2α (5). As such, while we still don't have a clear picture of what phosphorylation of eIF2α looks like, we can appreciate the flexible nature of eIF2α during its interaction with PKR (5,23,28,29).

We found an enrichment for K3-resistant variants at sites previously found to be under positive selection (22,23). In effect, these are two independent approaches that both pinpoint these sites as genetically pliable: positive selection analysis retrospectively considers existing variation in nature across multiple homologs, whereas deep mutational scanning prospectively tests novel variants in a single homolog. Thus, we view deep mutational scanning experiments as a complementary approach to positive selection analysis. One strength of deep mutational scanning is its potential to illuminate sites of opportunity against pathogens that have not exerted historical evolutionary pressure. It can also allow for the testing of insertions or deletions (39) which are generally discarded in positive selection analysis. On the other hand, a limitation of the deep mutational scanning approach is that it only tests variant function in a limited set of conditions, when the evolutionary selection could be a symphonic signal made up of a myriad of unknown pressures. Deep mutational scanning also requires the function of the protein to be testable in a selectable assay. We leveraged an existing yeast growth assay to test PKR function, and protein-protein interactions are generally amenable to high-throughput characterization (40–42). While proteins with cell-extrinsic or whole-organism phenotypes would be harder to test, recent advances in high-throughput genetic experimentation may allow deep mutational scans of those proteins as well (43,44).

IDEAS AND SPECULATION

It is curious that poxviral pseudosubstrate antagonists such as K3 are not closer mimics of the natural PKR substrate, eIF2α. The most extreme form would be a pseudosubstrate inhibitor identical to eIF2α aside from a mutation rendering it non-phosphorylatable, which would

be extremely difficult for PKR to distinguish from eIF2α. Perhaps pseudosubstrate divergence from eIF2α has been selected for, to allow the inhibitor to bind PKR with even higher affinity than eIF2α does. But mutations that improve binding to one species' PKR homolog might decrease binding to the divergent PKR homologs from other species (45–47). Thus, pseudosubstrate inhibitors may need to evolutionarily balance improving PKR binding for one host against maintaining PKR binding of multiple hosts. Interestingly, the structural component most diverged between K3 and eIF2α, a rigid helix in K3 in place of eIF2α's flexible phospho-acceptor site, primarily interacts with the residues around the activation loop of PKR, which we found to be highly functionally constrained. In contrast, amino acids 72-83 of K3 are highly identical to eIF2α and interact primarily with PKR's divergent helix αG (Supplemental Figure 5). Thus, vaccinia K3 may well have been evolutionarily constrained to avoid limiting host range.

Despite its pliability, PKR helix αG is believed to play a critical role in eIF2α binding and phosphorylation, as it induces a conformational shift in eIF2α that moves the Ser51 phospho-acceptor site into PKR's ATP-binding site. Based on our findings and positive selection analyses, this conformational shift is apparently amenable to variation in the PKR kinase domain. Perhaps the resilience of the binding surface between PKR and eIF2α could be informative for the budding field of de novo protein design, as resilient design is a critical consideration for vaccines and therapeutics faced with rapidly evolving infections agents.

MATERIALS AND METHODS

PKR and K3 plasmid construction

We generated the plasmid MSp508_MCS as the base plasmid for all experiments by inserting a multiple-cloning site (MCS) into the single-copy plasmid YCp50, a gift from Mark Rose (48). The MCS contained seven unique restriction enzyme cut sites: AgeI, BsiWI, BstEII, NotI, MluI, PvuII, and SacI. YCp50 was digested using restriction enzymes EcoRI-HF (NEB Cat#R3101S) and BspDI (NEB Cat#R0557S) with rCutSmart buffer, followed by purification and size selection on a 1% agarose gel with 0.6 µg/mL ethidium bromide (BioRad Cat#1610433), selecting for the 7,961 bp band. The extracted band was purified using a QIAquick Gel Extraction kit (Qiagen Cat#28706) and eluted with 30 µL of water. For Gibson Assembly the purified band was combined with a single-stranded DNA oligonucleotide containing the MSC and homology arms spanning the digest site, named Oligo 1 (Supplemental File), in a 1:5 molar ratio in a total volume of 5 µL, along with an additional 5 µL 2x Gibson Assembly Master Mix (NEB Cat#E2611L). The reaction mixture was incubated at 50°C for 1 hour. One microliter of the Gibson Assembly reaction was transformed into 10 µL 5-alpha competent E. coli (NEB Cat#C2987I) and plated onto 10-cm LB-AMP plates (IPM Scientific Cat#11006-016) and incubated overnight at 37°C. Colonies were selected for overnight outgrowth in 3 mL LB-AMP (IPM Scientific Cat#11006-004), followed by a QIAprep spin miniprep (Qiagen Cat#27106) and Sanger sequencing to validate the plasmid sequence using the primer Oligo 2. This produced the plasmid MSp508_MCS.

We cloned wild-type PKR, wild-type K3, and K3-H47R into the MCS of MSp508_MCS to produce the MSp509_PKR-WT, MSp510_K3-WT, and MSp511_K3-H47R plasmids, respectively. The plasmid p1419 (21) contains the *PKR* gene (standard name *EIF2AK2)* under the control of the galactose-inducible pGAL10/CYC1 promoter for yeast expression. We PCR-amplified the PKR expression construct using primers (Oligos 3 and 4) to place pGAL10/CYC1-*PKR* on MSp508_MCS between AgeI and BstEII. Oligo 4 incorporated a

synthetic terminator for *PKR*, T_{synth1} (49), a 28-bp spacer sequence, and a 26-bp barcode (CGCTTAATATGCAATGAAATTGCTTA); the primers added 20-bp homology arms for cloning onto the MSp508_MCS plasmid. PCR was performed using the polymerase PfuUltraII (Agilent Technologies Cat#600674) for 30 cycles with an annealing temperature of 55°C, followed by gel purification. MSp508_MCS was digested with restriction enzymes AgeI-HF (NEB Cat#R3552S) and BstEII-HF (NEB Cat#R3162S), followed by gel purification. The digested MSp508_MCS and pGAL10/CYC1-PKR fragments were joined via Gibson Assembly in a 1:5 molar ratio, followed by transformation into 5-alpha competent *E. coli* and plating onto LB-AMP. A resultant clone, validated by Sanger sequencing, was designated MSp509_PKR-WT.

To produce MSp510_K3-WT, the *K3L* coding sequence was amplified from the plasmid pC140 (21) using primers Oligo 5 and Oligo 6. Similarly, to produce MSp511_K3-H47R, the *K3L-H47R* coding sequence was amplified from the plasmid pC407 using primers Oligo 6 and Oligo 7. pC140 and pC407 were gifts from Thomas Dever. We note that the *K3L* alleles on these plasmids carry a Val2Leu mutation relative to the VACV-WR *K3L* (Genbank AAO89313.1), which was introduced during the cloning of *K3L* on the original plasmid pTM1 (17). Oligos 5 and 7 incorporated a synthetic terminator (T_{synth8}) for the K3 allele (49), a 28-bp nucleotide spacer sequence, a 26-bp barcode (ATCGTAATAGGTTTCCGGCTTGTTCG and ACAGGAAATGCTTTCGGGGTTGTATT, respectively), and a 20-bp homology arm to MSp508_MCS. The pTDH3 promoter, also known as pGPD, was amplified from genomic DNA of the yeast strain BY4742 (50) using primers Oligo 8 and Oligo 9). K3-WT, K3-H47R, and pTDH3 were amplified using PfuUltraII, followed by gel purification. MSp508_MCS was digested with restriction enzymes NotI-HF (NEB Cat#R3189S) and PvuII-HF (NEB Cat#R3151S). Each K3 allele (K3-WT and K3-H47R) was joined with a pTDH3 promoter onto the digested MSp508_MCS via Gibson Assembly in a 1:5:5 molar ratio, followed by transformation into 5-alpha competent *E. coli* and plating onto LB-AMP. Resultant clones, validated by Sanger sequencing, were designated as MSp510_K3-WT and MSp511_K3-H47R.

MSp512_K3Δ58 was made by removing nucleotides 1-173 of the *K3L* sequence of YCp50_K3, leaving codons 59-88 to generate a nonfunctional *K3L* allele (denoted K3Δ58) under the pTDH3 promoter. The K3Δ58 sequence was amplified from the MSp510_K3-WT plasmid using primers Oligo 10 and Oligo 11. MSp510_K3-WT was digested with restriction enzymes NotI-HF and MluI-HF (NEB Cat#R3198S), followed by gel purification and size selection for the 8,684 bp band. The K3Δ58 allele was joined with the MSp508_MCS digested fragment via Gibson Assembly in a 1:5 molar ratio, followed by transformation into 5-alpha competent E. coli and plating onto LB-AMP. A resultant clone was validated by Sanger sequencing and designated as MSp512_K3Δ58.

Generation of the *PKR* variant library

We made the *PKR* variant library by 15 Gibson Assembly reactions using pairs of PCR amplicons targeting 4 windows of PKR protein sequence: 255-278 (5 pairs of reactions), 371-385 (3 pairs of reactions), 448-455 (2 pairs of reactions), and 480-506 (5 pairs of reactions). In each pair, one of the PCR reactions, PCR-1, utilized a pool of forward primers, termed a variant primer tile set, containing single-base deviations from the *PKR* sequence to generate all SNP-accessible variants in a stretch of X to Y consecutive codons (Supplemental Figure 8). Each forward primer in a tile set was composed of a 20-bp homology arm, the variant region, and a priming region with a melting temperature of 50°C. The reverse primer of all PCR-1 reactions had a constant 20-bp homology arm, a barcode region with 20 fully mixed bases, and a constant PCR-priming region homologous to the end of *PKR* with a melting temperature of 55°C (Supplemental Figure 9). Mixed bases in the barcodes were limited to 5-nucleotide stretches to prevent the unintended generation of restriction sites, spacing these stretches with “AA” or “TT” sequences (41). The forward PCR-1 primer tile sets are listed in the Supplemental File as Oligos 12-26, and the reverse PCR-1 barcode primer is listed as Oligo 27. The paired PCR-2 reaction was an inverse PCR that generated the remainder of the *PKR* gene and

plasmid backbone, with overlaps to the homology arms of PCR-1; the forward primer for PCR-2 reactions was Oligo 28, and the reverse primers were Oligos 29-43. Thus, Gibson Assembly of these PCR products would generate sets of *PKR* variants in defined segments of *PKR* along with random barcodes at the end of the gene (Supplemental Figure 10).

We used custom Python (v3.8.18) scripts to design each variant-introducing primer, utilizing IUPAC degenerate base symbols (e.g. “D” represents a mix of “A”, “G”, and “T”) (51,52). The introduced variant amino acids were all within a single base change from the canonical human *PKR* coding sequence (Ensemble ENSG00000055332; CDS EIF2AK2-001); we note the *PKR* sequence encoded on MSp509_PKR-WT differs somewhat from the canonical sequence through synonymous substitutions. Primers were synthesized by Integrated DNA Technologies, then were pooled manually into variant primer tiles sets.

PCR-1 reactions were amplified from the MSp509_PKR-WT plasmid using PfuUltraII polymerase with an annealing temperature of 45°C. Each reaction generated a pool of *PKR* gene fragments containing a single nonsynonymous variant paired with a unique barcode sequence. In parallel, PCR-2 reactions amplified vector fragments from MSp509_PKR-WT plasmid using Herculase II polymerase (Agilent Technologies Cat#600677) using cycling conditions for vector targets >10 kb. PCR-1 and PCR-2 amplicons were digested with DpnI (NEB Cat#R0176S) for 2 hours at 37°C to remove the MSp509_PKR-WT template DNA, followed by gel purification. Each PCR-1 amplicon was paired with its corresponding PCR-2 amplicon for Gibson Assembly in a 1:5 molar ratio.

Two µL of each Gibson Assembly was transformed into 50 µL of 5-alpha competent *E. coli* and all cells were plated on LB-AMP and incubated overnight at 37°C. Colonies were counted on each plate and plates were bottlenecked at approximately 30 colonies per selected nonsynonymous variant (i.e. each nonsynonymous variant would be linked to approximately 30 unique barcode sequences). Plates were washed with 15 mL LB-AMP, cell densities were measured as the optical density at 600 nm (OD_{600}), and an equal number of cells (approximately

5×10^9 cells) from each reaction were pooled to form a single *PKR* variant library. A 200 mL LB-AMP outgrowth was grown to an OD_{600} measurement of approximately 3, followed by a Qiagen MAXI plasmid prep (Qiagen Cat#12963).

PacBio sequencing of barcoded variant libraries

The *PKR* variant library was sequenced using the PacBio Sequel II instrument to identify the barcodes that were linked to each nonsynonymous variant. Sixteen micrograms of the *PKR* variant library were digested with restriction enzymes AgeI-HF and NotI-HF to create linear fragments of approximately 2,100 bp for sequencing. The digest was incubated for 2 hours at 37°C then purified using a QiaQuick PCR purification kit (Qiagen Cat#28106). Fragments of 2,100 bp were size-selected using a SageELF instrument (Sage Science) before PacBio circular consensus sequencing (CCS) HiFi sequencing on a Sequel II instrument (Pacific Biosciences). From the CCS reads we generated a table of *PKR* barcodes paired with genetic variants using alignparse v0.2.6 and custom Python scripts (53).

Combining *PKR* variant library with K3 alleles

We next cloned the three K3 alleles (K3, K3-H47R, and K3Δ58) onto the PKR variant library plasmids. The PKR variant library was digested with NotI-HF and PvuII-HF to generate a 10,270 bp receiver DNA fragment, while the K3 alleles were separately digested with BstEII-HF and SacI-HF (NEB Cat#R3156S) to generate a 1,130 bp insert DNA fragment with 20-bp homology arms with homology to the receiver fragment for Gibson Assembly. For each of the reactions, 10 µg of plasmid DNA was digested with 5 µL of each restriction enzyme in a 50 µL reaction to generate a linear fragment, followed by gel purification. The PKR and K3 fragments were combined in three separate Gibson Assembly reactions in a 1:5 molar ratio with 100 ng of the PKR vector fragment in a reaction volume of 20 µL and incubated at 50°C for 1 hour. A

standard ethanol precipitation was performed on each reaction, adding 100 µL of 100% ETOH and incubating overnight at -20°C before resuspending the pellets in 2 µL of water (54).

We transformed each of the three concentrated reactions into *E. cloni* 10G SUPREME electrocompetent cells (Lucigen Cat#60080-2) using a MicroPulser Electroporator (BioRad), with 2 µL of the concentrated reaction combined with 25 µL of competent cells. Following the electroporation and recovery, cells were plated onto two 15-cm LB-AMP plates. A 1:1000 dilution was plated with 100 µL of water on a 10-cm plate to estimate the number of transformed colonies across the two 15-cm plates, which produced an estimate of approximately 2 million colonies per electroporation. Plates were washed with 25 mL LB-AMP, combining the two plates for each of the reactions. Each library was expanded in 200 mL LB-AMP to an OD_{600} measurement of approximately 3, followed by a Qiagen MAXI plasmid prep.

Screening the PKR-K3 variant pairs in a yeast growth assay

All yeast growth was at 30°C in CSM-Ura (Sunrise Science Products Cat#1004-010), and liquid cultures were shaken at 200 revolutions per minute. We transformed each of the three paired libraries (PKR + K3, PKR + K3Δ58, and PKR + K3-H47R) into yeast in duplicate. The paired plasmid libraries were transformed into the yeast strain BY4742 (*MAT*α *ura3*Δ0 *leu2*Δ0 *his3*Δ1 *lys2*Δ0; strain name MSY2) (50) using a standard large-scale high efficiency lithium acetate transformation protocol (55) and plated onto 15-cm plates with 2% dextrose, using 1:1,000 and 1:10,000 dilutions on 10-cm plates to estimate colony counts, which produced approximately 100,000 colonies. Colonies were washed off plates with 25 mL media with 2% dextrose.

To start the yeast growth assay, six cultures were seeded at low density (OD_{600} measurement of 0.01) in 40 mL media with 2% dextrose and grown overnight for 16 hours. The following morning cultures were moved from 30°C to 4°C to pause growth, then restarted at an OD_{600} of 0.25 in 40 mL media with 2% dextrose 3 hours prior to GAL induction. After 3 hours,

once all cultures were in a log growth phase, approximately 1×10^8 cells were taken from each culture as the starting timepoint sample (designated as “0 hours”). All timepoint samples were spun down in a 1.5 mL Eppendorf tube at 5,000 rpm for 1 minute, after which supernatant was removed and the pellet was frozen at -80°C. Approximately 2×10^8 cells from the remaining cultures were pelleted in 50 mL conical tubes at 3,000 rpm, supernatant was removed, and cells were resuspended in 80 mL CSM-Ura media with 2% galactose (OD_{600} measurement of 0.125 OD_{600}) to induce PKR expression. Cultures were incubated overnight at 30°C, 200 rpm. Additional samples were harvested at 12, 16, and 20 hours post-PKR induction, with approximately 1×10^8 cells taken at each timepoint. The culture density was monitored and back diluted to maintain a log growth phase (OD_{600} measurement less than 1). With three K3 allele conditions (K3, K3Δ58, and K3-H47R), four timepoints (0, 12, 16, and 20 hours) and two replicates, a total of 24 samples were taken across the yeast growth assay.

Plasmid extraction and barcode amplification

To quantify changes in barcode abundance between timepoints in the yeast growth assay, we harvested plasmids from the four timepoints for each K3 allele, then amplified and sequenced the adjacent PKR and K3 barcodes in each plasmid. Plasmids were harvested from the yeast samples using a modified QIAprep spin miniprep protocol (56). Sampled cell pellets were first thawed at room temperature, followed by the addition of 250 µL QIAprep P1 buffer and 2 µL zymolyase (2.5 units per µL), and incubation at 37°C for 30 minutes, followed by adding 250 µL P2 buffer and following the manufacturer instructions for the remainder of the protocol.

To identify and quantify the abundance of each PKR variant paired with each of the three K3 alleles we performed Illumina sequencing of the PKR barcodes from the plasmids. We used a two-step PCR protocol, the first to amplify the PKR barcode from the plasmids and the second to attach Illumina adapters and Nextera indices for downstream sequencing and sample

demultiplexing. For the first PCR reaction, we designed forward and reverse primers as a pool of five primers each, that had 0 to 4 “N” bases between the Nextera transposase adapter sequence and the priming sequence to stagger the base signal per cycle and maintain sequence diversity across the flow cell (57). The forward primers (Oligos 44-48) with a melting temperature of 60°C were pooled together and used for all PCR-1 reactions, annealing immediately upstream of the PKR barcode locus. Five reverse primers were designed and pooled for amplification from each of the three K3 alleles that annealed within the three distinct K3 barcode sequences (Oligos 49-53 for K3, Oligos 54-58 for K3Δ58, and Oligos 59-63 for K3-H47R). One microliter of each primer pool was combined with 50 ng of sample plasmid DNA and 25 µL Kapa Hifi Hotstart ReadyMix (Kapa Biosystems Cat# KK2602), and topped off to a total volume of 50 µL with water. PCR cycling was performed as follows: (1) 95°C for 3 minutes, (2) 98°C for 20 seconds, (3) 65°C for 15 seconds, (4) 72°C for 1 minute, repeat steps 2-4 for a total of 18 cycles, (5) 72°C for 1 minute, (6) 12°C hold. PCR products were purified using a MinElute PCR purification kit (Qiagen Cat#28006) using 11 µL water for the final elution, then quantified using a Qubit dsDNA High Sensitivity (HS) kit (ThermoFisher Scientific Cat#Q32851).

For the second PCR reaction, 10 ng of amplicon DNA from the first PCR reaction was combined with 10 µL of Nextera adapter index primers (Illumina Cat#20027213) and 25 µL Kapa HiFi Hotstart ReadyMix, and topped off to a total volume of 50 µL with water. PCR cycling was performed as follows: (1) 100°C for 45 seconds, (2) 100°C for 15 seconds, (3) 60°C for 30 seconds, (4) 72°C for 30 seconds, repeat steps 2-4 for a total of 8 cycles, (5) 72°C for 1 minute, (5) 12°C hold. PCR products were purified using a MinElute PCR purification kit using 11 µL of water for the final elution, then quantified using a Qubit dsDNA Broad Range (BR) kit (ThermoFisher Scientific Cat#Q32850). 500 ng of the second PCR reaction amplicons were gel purified on a Size Select II E-Gel (ThermoFisher Scientific Cat#G661012) for approximately 13 minutes to extract the approximately 250-bp amplicon band, followed by quantification with a Qubit dsDNA HS kit. Amplicon samples were diluted to 4 nM before being pooled together,

followed by manufacture denature and dilution protocols (Illumina Document#15039740v10) before sequencing on an Illumina NextSeq 2000 instrument.

PKR functional scores and screening analysis

Next, we extracted PKR barcode sequences from the Illumina reads and mapped the barcodes back to their corresponding select nonsynonymous variants using the table of PKR barcodes paired with genetic variants generated from the PacBio CCS HiFi reads. Paired reads were assembled into contiguous sequences using PEAR v0.9.11 (58), followed by Bartender v1.1 (59) to extract and cluster PKR barcodes using the barcode search pattern "CAAGG[25-27]GGTGA".

We wrote Python v3.8.18 scripts to tally the PKR barcodes and map them back to genetic variants using the table generated from the PacBio CCS HiFi reads. Barcode counts were normalized to the total read count for each of the four timepoints across each of the K3 alleles, followed by a fold change calculation for each PKR barcode across timepoints 0, 12, 16, and 20 hours:

$$FC_{TP} = -\ log_2\left(\frac{Normalized\ Reads_{TP}}{Normalized\ Reads_0}\right)$$

where FC_{TP} is the fold change in normalized reads for a PKR barcode from 0 hours to a subsequent timepoint (TP), $Normalized\ Reads_{TP}$ is the read count for a given barcode divided by the total number of reads at the given timepoint, and $Normalized\ Reads_0$ is the read count for a given barcode divided by the total number of reads at the 0-hour timepoint. Of note, the $\log_2$ of the fold change is inverted such that functional PKR variants that inhibit yeast growth and decreased in abundance were assigned positive fold change value. We then calculated PKR functional scores for each PKR barcode in each of the K3 alleles by calculating the area under the curve using the composite trapezoidal rule. We calculated the PKR functional score for each variant for each K3 allele by averaging the PKR functional score across all representative

barcodes for a given variant. As we found a strong correlation between replicates (Pearson correlation coefficient > 0.98 for each K3 allele, Supplemental Figure 3), we proceeded by combining all reads from the two replicate experiments for each K3 allele and recalculating the fold changes and PKR functional scores for each barcode as described above.

Predicted PKR complexes and substrate contacts

All molecular graphics and analyses were performed using USCF ChimeraX v1.5 and PyMol v2.5.4. To define PKR residues contacting K3 and eIF2α we used the AlphaFold2 structure prediction tool (ColabFold v1.5.5) in UCSF ChimeraX v1.5 (60,61) (Supplemental Figure 11). We compared the predicted multimers to crystal structures of K3 (PDB 1LUZ) (20) and PKR in complex with eIFα (2A1A) (5), noting the eIFα flexible loop containing Ser51 adopts a compact helical conformation in the predicted model, matching the conformation of this loop in the crystal structure of unbound yeast eIF2α (PDB 1Q46). As Ser51 of eIF2α would need to travel approximately 17 Å to reach PKR's active site (5), this predicted structure is likely reflective of a pre-phosphorylation complex, and thus would not capture contacts between PKR and eIF2α made during phosphorylation. The corresponding loop is unresolved in the crystal structure of human PKR in complex with eIF2α (PDB 2A1A). Note that we opted not to use the existing structure for defining contact residues, both for consistency with the K3 contact definitions, as there is no PKR-K3 co-crystal, as well as for the potential to capture interactions mediated by residues unresolved in the crystal structure.

To identify potential PKR contact sites from the predicted complex structures, we used PyMol v2.5.4 to select all PKR residue branches within 5 Å of K3 or eIFα using the command: "sele contacts, br. /best_model//A within 5 of /best_model//B" with chain "A" being PKR and chain "B" being either K3 or eIFα. Predicted PKR sites that contact K3 were: 275, 276, 278, 304, 339, 343, 345, 375, 379, 382, 414, 416, 435, 448, 449, 450, 451, 452, 453, 455, 460, 485, 486, 487, 488, 489, 490, 492, 493, and 496. Predicted PKR sites that contact eIFα were: 274, 275,

276, 279, 335, 337, 338, 339, 340, 341, 342, 379, 382, 451, 452, 453, 483, 486, 487, 488, 489, 490, 491, 492, and 493.

PKR sites in the kinase domain under positive selection across vertebrate species were identified by Rothenburg et al. (22): 261, 269-272, 307, 314, 322, 360, 368, 375, 378, 379, 382, 385, 389, 394, 405, 428, 448, 449, 462, 471, 483, 486, 488, 491, 493, 500, 502, 504, 505, 514, 520, and 524. Sites conserved across vertebrate PKR homologs were also identified by Rothenburg et al. (22) which were used for supplementary analysis: 263, 267, 276-279, 281, 283, 296, 298, 308, 309, 312, 315, 317, 319, 320, 323, 327, 362, 364-367, 369, 374, 377, 397, 401, 404, 406, 407, 410-417, 419, 420, 429-433, 437, 444-446, 450, 451, 454, 455, 457-459, 465, 469, 470, 474-477, 480, 490, 511, 519, 523, and 526.

Data access

Supplemental data and code are available under the DOI: 10.5281/zenodo.11095101

ACKNOWLEDGEMENTS

We thank Thomas Dever, Stefan Rothenburg, Adam Phillippy, Bernard Moss, Stephanie Jaquet, Darach Miller, and members of the Sadhu lab for helpful discussions. We thank Thomas Dever and Mark Rose for strains and plasmids. Next-generation sequencing was performed by both the NIH Intramural Sequencing Center (NISC) and the Microarrays and Single-Cell Genomics Core of the National Human Genome Research Institute. This work utilized the computational resources provided by the NIH HPC Biowulf Cluster (http://hpc.nih.gov). This work was supported by the Intramural Research Program of the National Human Genome Research Institute, NIH (1ZIAHG200401).

AUTHOR CONTRIBUTIONS

Designed research/conceptualization: M.J.C. and M.J.S. designed research;

Performed research: M.J.C. and S.B. performed research;

Analyzed data: M.J.C., S.B., and M.J.S. analyzed data;

Wrote paper: M.J.C. and M.J.S. wrote the paper.

All authors read and revised the paper.

AUTHOR DETAILS

Michael J Chambers

- Center for Genomics and Data Science Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA
- Department of Microbiology & Immunology, Georgetown University, Washington DC, USA

Contribution: Conceptualization, Investigation, Visualization, Writing - original draft

Competing interests: No competing interests declared

ORCID: 0000-0002-0658-6984

Sophia Scobell

Center for Genomics and Data Science Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA

Contribution: Investigation, Analysis, Writing - review and editing

Competing interests: No competing interests declared

ORCID: 0000-0003-4630-0167

Meru J Sadhu

Center for Genomics and Data Science Research, National Human Genome Research Institute, National Institutes of Health, Bethesda, Maryland, USA

Contribution: Conceptualization, Supervision, Analysis, Writing - original draft

For correspondence: meru.sadhu@nih.gov

Competing interests: No competing interests declared

ORCID: 0000-0002-8636-9102

FIGURES

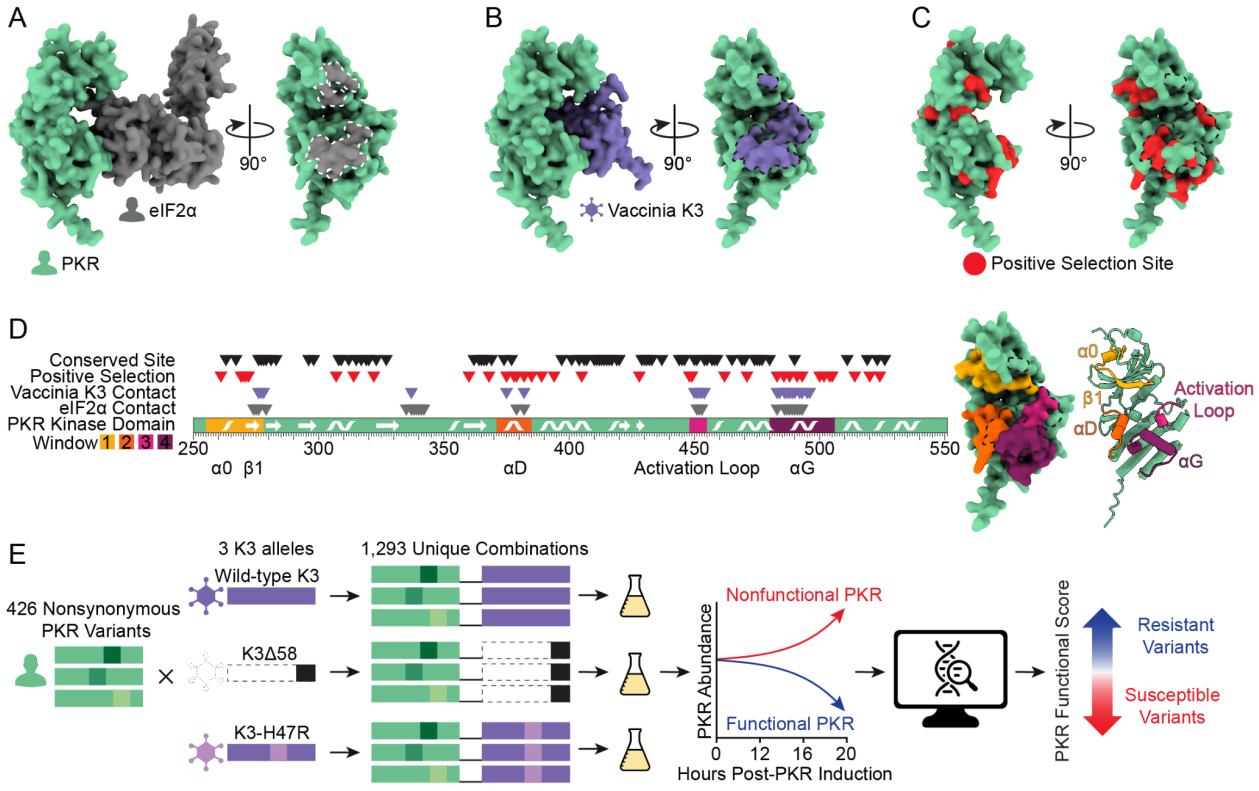


Figure 1. Exploring genetic variants of human PKR against the pseudosubstrate antagonist vaccinia K3. (A) (Left) AlphaFold2 model of the PKR kinase domain (green) bound to its target, eIF2α (gray). (Right) Rotated 90°, the eIF2α binding surface of PKR is shaded gray and outlined in white dashed line. (B) AlphaFold2 model of the PKR kinase domain bound to vaccinia K3 (purple). Rotated 90°, the K3 binding surface of PKR is shaded purple and outlined in black dashed line. (C) PKR kinase domain with vertebrate positive selection sites highlighted in red (22), with the K3 binding surface delineated with dashed line. (D) (Left) The PKR kinase domain with four windows highlighted in which we designed 426 nonsynonymous variants. Secondary structures occurring in these windows (alpha helices α0, αD, and αG, beta strand β1, and the kinase activation loop) are marked (5). The positions of conserved sites and sites under positive selection (22) are denoted above with black and red triangles, respectively, and the positions of sites within 5 angstroms of either K3 or eIF2α in an Alphafold2 binding model are denoted with purple and gray triangles, respectively. (Right) Mutated windows are highlighted on the surface of the PKR kinase domain, with the K3 binding surface delineated with dashed line, and in cartoon form with secondary structures marked. (E) Methodological approach to explore the effects of variants of human PKR in the presence of different K3 alleles. 426 nonsynonymous variants of PKR were generated and paired with wild-type K3, K3Δ58, and K3-H47R. Variant effects were characterized using a high-throughput yeast growth assay and massively parallel sequencing.

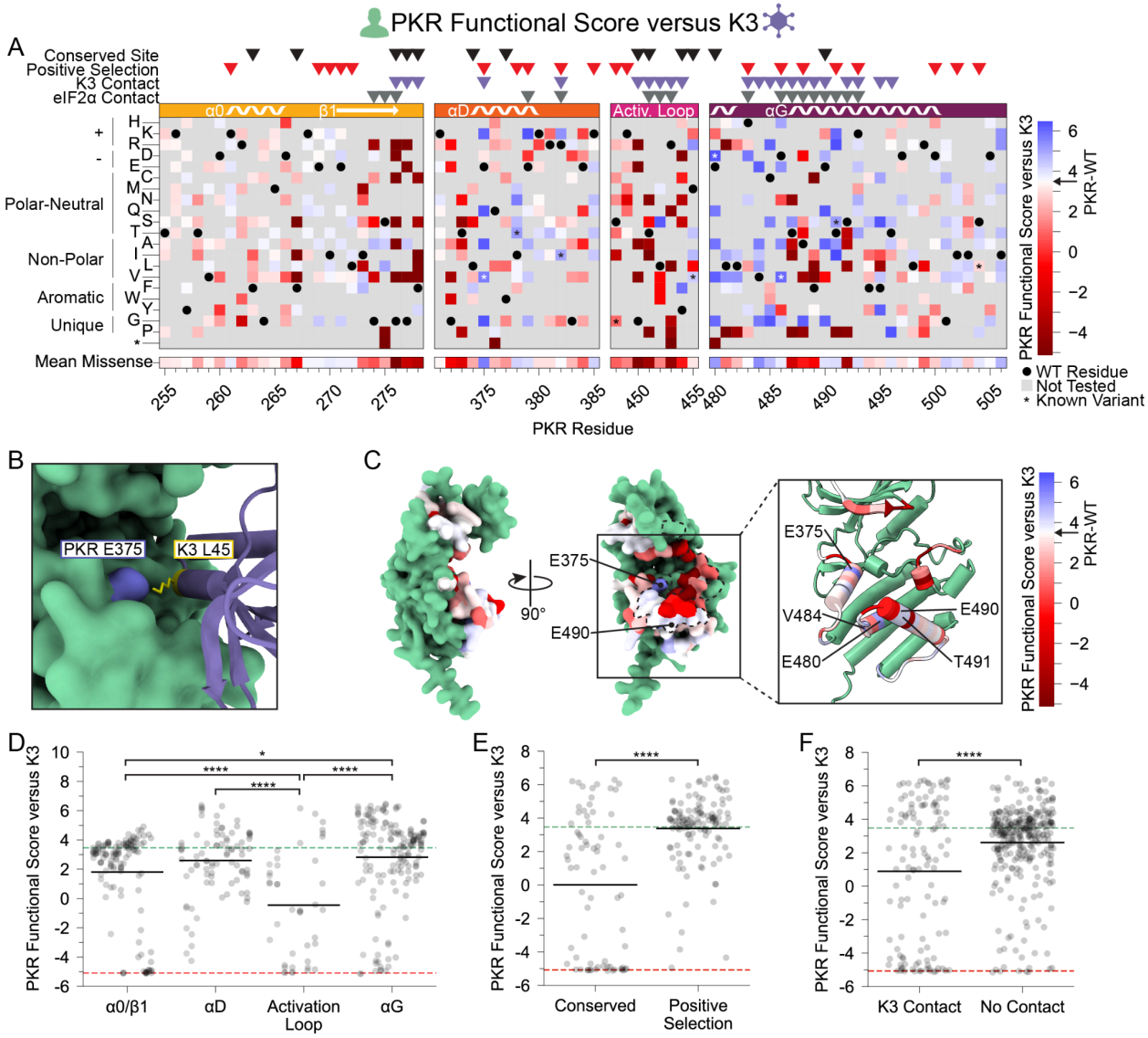


Figure 2. PKR variants that evade K3 and maintain kinase function are enriched at positive selection sites and helices αD and αG. (A-B) PKR functional scores versus K3 are colored ranging from susceptible (red) to WT-like (white) to resistant (blue). (A) Heatmap of PKR variants with cells colored by the PKR functional score versus K3 for SNP-accessible variants. Wild-type PKR residues and untested variants are denoted with black circles and gray squares, respectively. Previously characterized K3-H47R-resistant variants are noted with an asterisk (28). (B) (C) Attractive electrostatic interaction between PKR-E375 (blue) and K3-L45 (yellow) in the AlphaFold2 model of PKR bound to K3. (C) Surface structure of the PKR kinase domain with sites colored by the mean PKR functional score versus K3 for missense variants. The K3 binding surface is delineated with black dashed line. (Inset) Location of K3-resistant sites cited in the text. (D-F) Stripplots of PKR functional score versus K3 for variants, with green and red dashed lines representing mean scores for WT PKR and nonsense variants, respectively. Variants are partitioned by nearest secondary structural element (D), level of conservation in vertebrates (E), or predicted contact with K3 (F). Points in D are ordered left to right by their

position along the kinase domain, whereas they are randomly jittered along the x-axis in E and F. (* $p < .05$, **** $p < .0001$, Tukey's HSD (D) and two-sample t-test (E, F).)

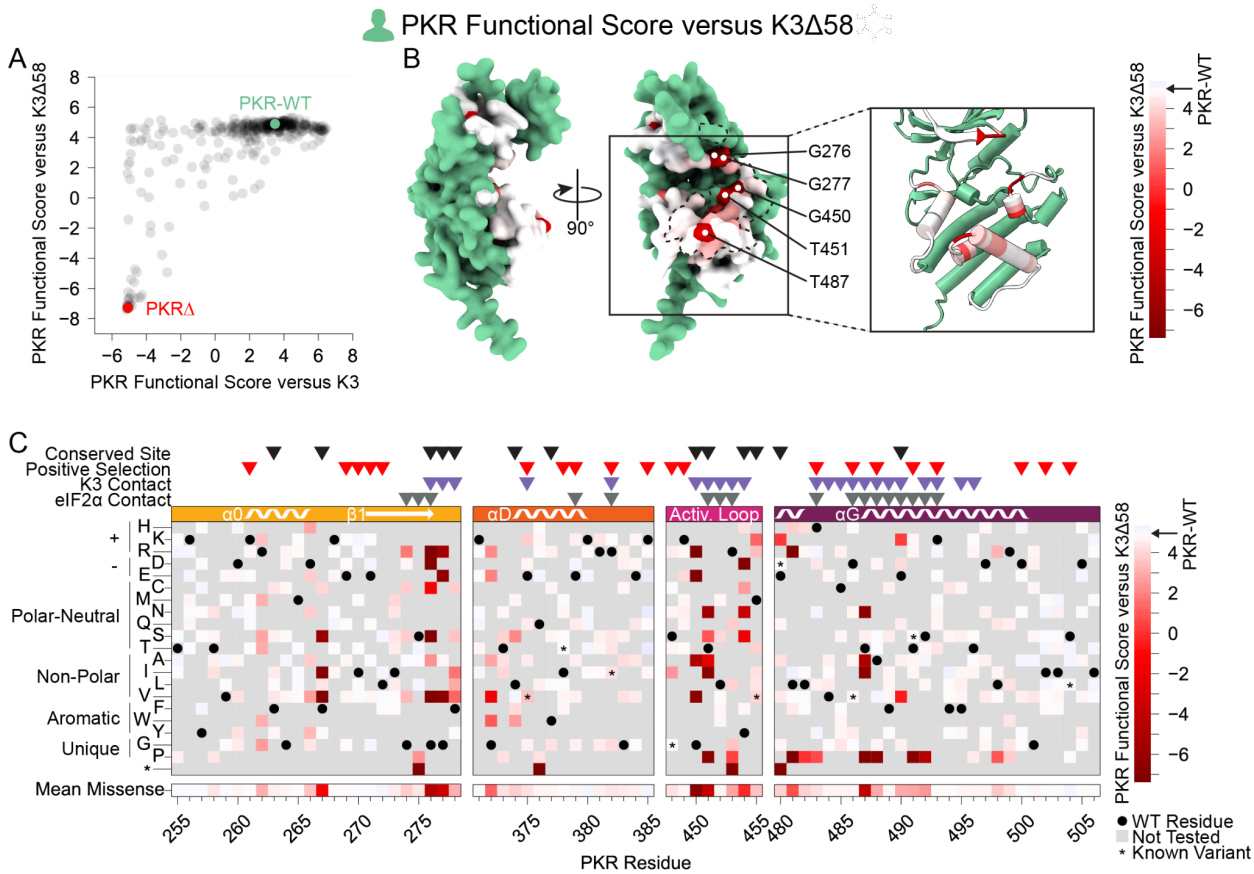


Figure 3. Few nonfunctional PKR variants identified in the absence of K3 inhibition. (A) Scatterplot showing each variant's PKR functional score versus wild-type K3 plotted against its PKR functional score versus K3Δ58. The data point for PKR-WT is colored green and a data point representing the average of the four nonsense variants (PKRΔ) is colored red. (B-C) PKR functional scores versus K3Δ58 are colored ranging from nonfunctional (red) to WT-like (white). (B) Surface structure of the PKR kinase domain with sites colored by the mean PKR functional score versus K3Δ58 for missense variants. The K3 binding surface is delineated with a black dashed line. (Inset) Location of highly constrained sites cited in the text. (C) Heatmap of PKR variants with cells colored by the PKR functional score versus K3Δ58 for each variant.

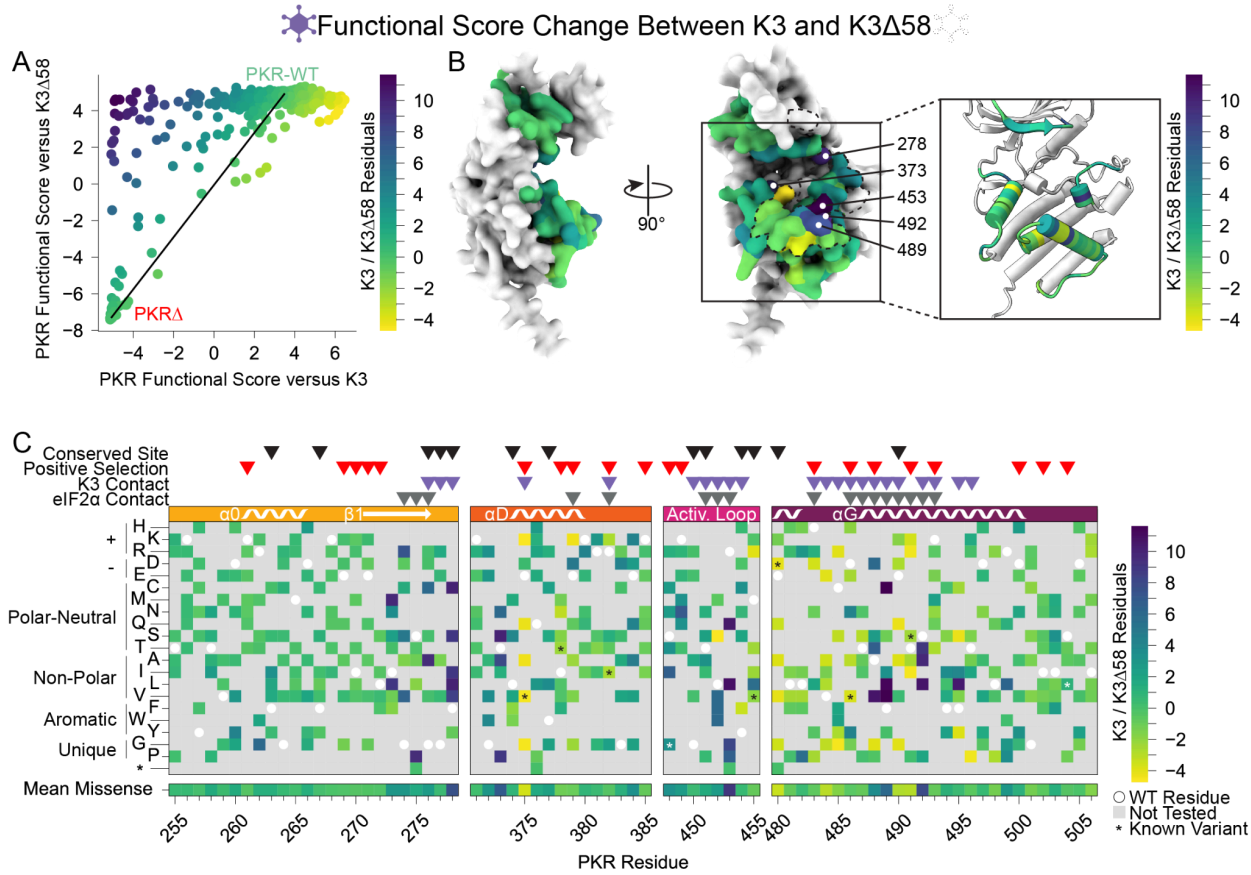


Figure 4. Identification of PKR sites highly susceptible to K3 inhibition. (A) A line was drawn connecting the data point for PKR-WT to the position of the average data point for the four nonsense variants (PKRΔ) from the data in the scatterplot from Figure 3A, with all other data points colored by their residual from that line (K3 / K3Δ58 Residuals), ranging from K3-susceptible (purple) to K3-indifferent (green) to K3-resistant (yellow). This color scheme is used in panels B and C. (B) Surface structure of the PKR kinase domain with sites colored by the mean K3 / K3Δ58 residuals for missense variants. The K3 contact site is delineated with a black dashed line. (Inset) Location of K3-susceptible sites cited in the text. (C) Heatmap of PKR variants with cells colored by the K3 / K3Δ58 residuals for each variant.

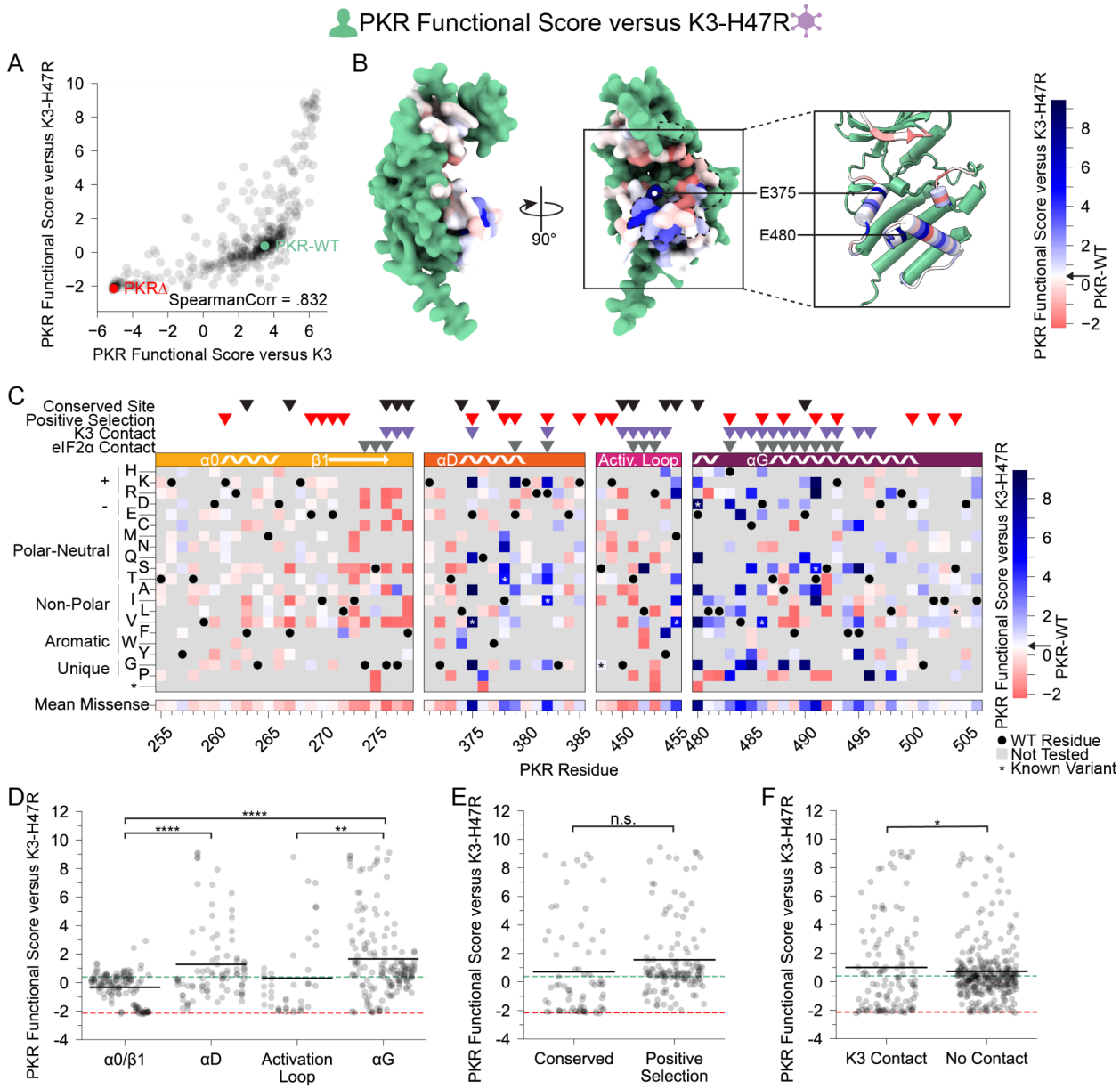


Figure 5. PKR variants that are K3-resistant are also largely K3-H47R-resistant. (A) Scatterplot of PKR functional scores versus wild-type K3 plotted against PKR functional scores versus K3-H47R. The data point for PKR-WT is colored green and a datapoint representing the average of the four nonsense variants (PKRΔ) is colored red. (B-C) PKR functional score versus K3-H47R ranging from susceptible (red) to WT-like (white) to resistant (blue). (B) Surface structure of the PKR kinase domain with sites colored by the mean PKR functional score versus K3-H47R for missense variants. The K3 binding surface is delineated with a black dashed line. (Inset) Location of K3-H47R resistant sites cited in the text. (C) Heatmap of PKR variants with cells colored by the PKR functional score versus K3-H47R for each variant. (D-F) Stripplots of PKR functional scores versus K3-H47R for variants, as plotted in Figure 2D-F. Variants are partitioned by nearest secondary structural element (D), level of conservation in vertebrates (E), or predicted contact with K3 (F). Points in D are ordered left to right by their position along the

kinase domain, whereas they are randomly jittered along the x-axis in E and F. (* $p < .05$, ** $p < .01$, **** $p < .0001$, Tukey's HSD (D) and two-sample t-test (E, F).)

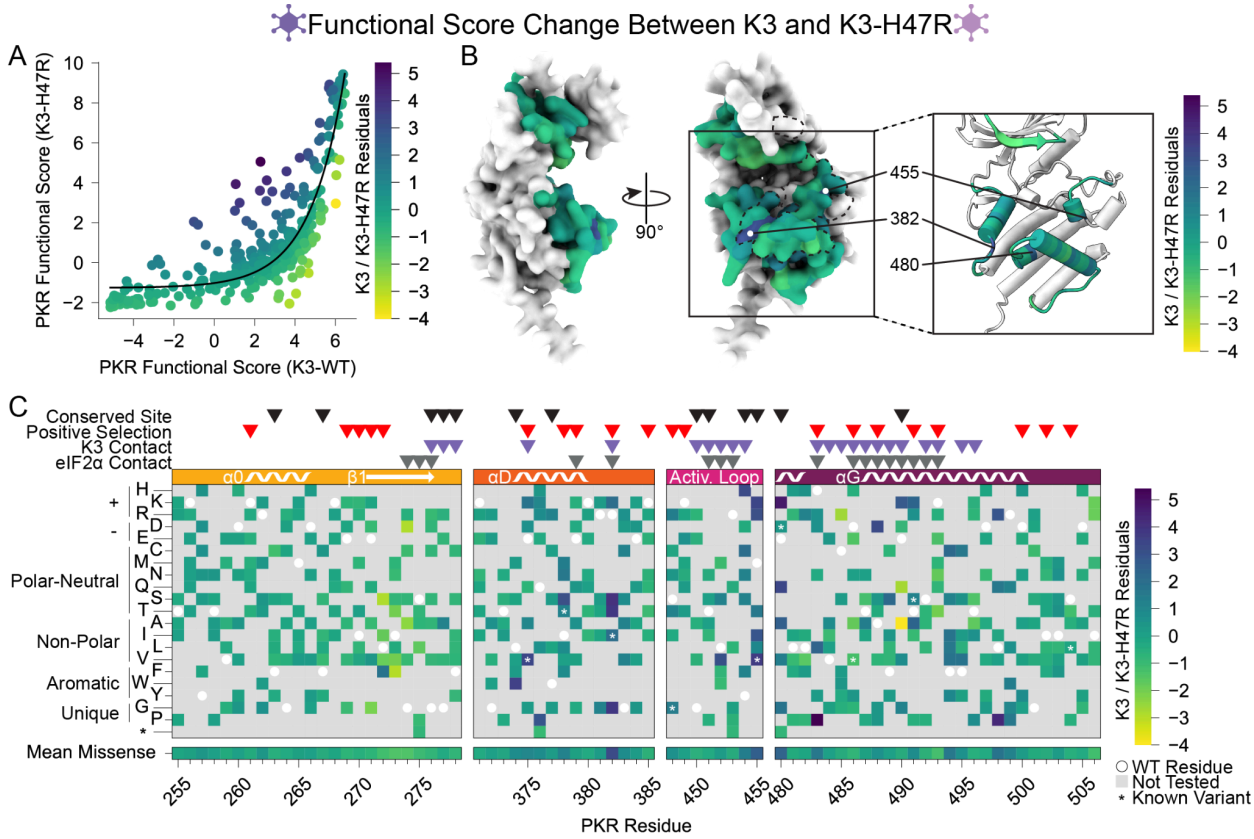


Figure 6. Identification of PKR variants with differing resistance to wild-type K3 and K3-H47R. (A) A nonlinear exponential curve (black line) was fitted to the data in the scatterplot from Figure 5A. Points are colored by their residuals from that curve (K3 / K3-H47R Residuals), ranging from enhanced (purple) to decreased (yellow) resistance to K3-H47R relative to the expectation from K3 resistance. This color scheme is used in panels B and C. (B) Surface structure of PKR kinase domain with sites colored by the mean K3 / K3-H47R residuals for missense variants. The K3 contact site is delineated with a black dashed line. (Inset) Location of K3-H47R-resistant sites cited in the text. (C) Heatmap of PKR variants with cells colored by the K3 / K3-H47R residuals for each variant.

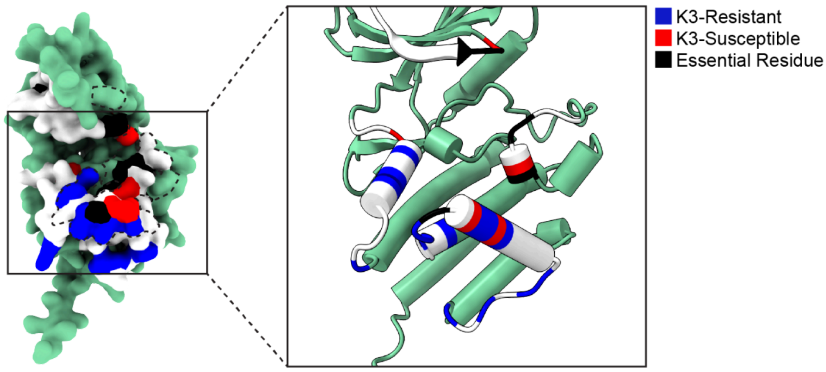
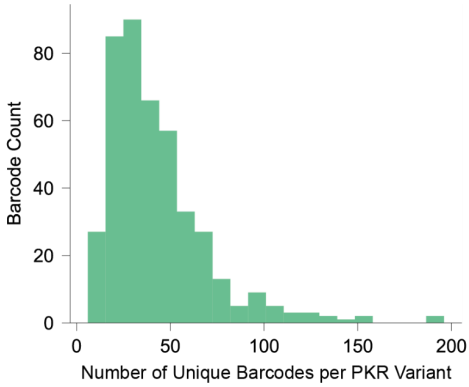
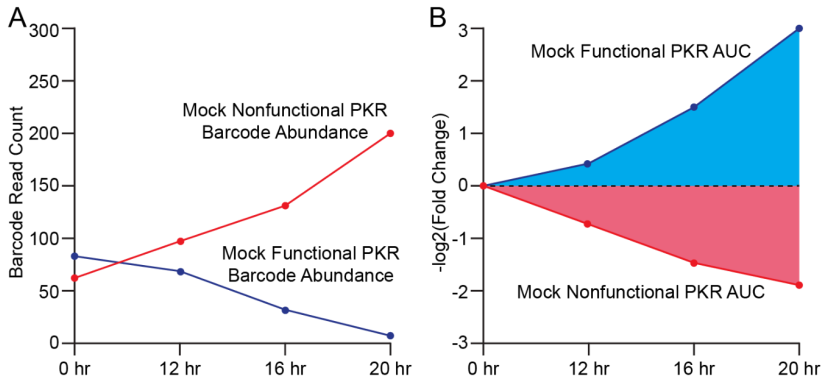


Figure 7. Unified spatial view of highlighted PKR sites across the K3-binding interface. Positions of sites classified as K3-resistant (blue), K3-susceptible (red), and essential residues (black) based upon the above analyses.

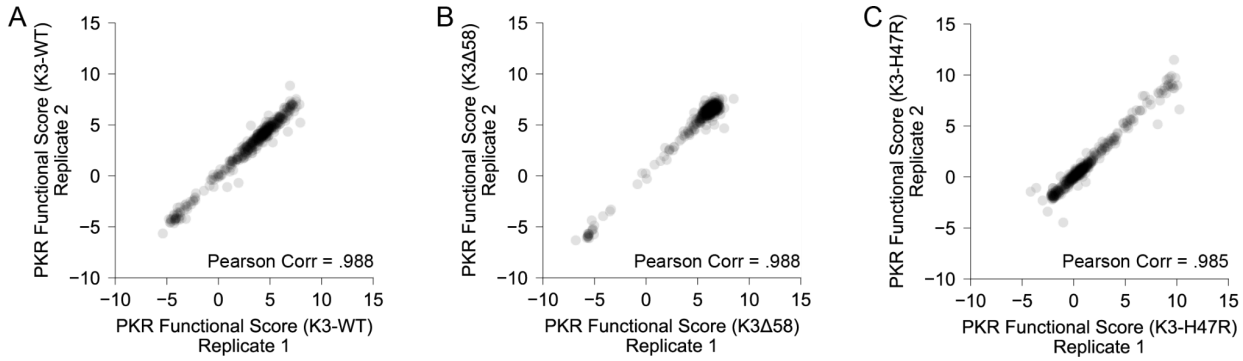
SUPPLEMENTAL FIGURES



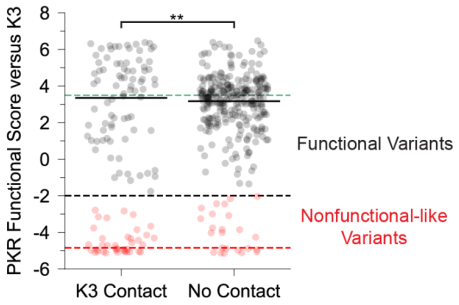
Supplemental Figure 1. Count of unique barcode sequences per PKR variant. Histogram depicts the number of unique barcodes linked to each PKR variant, with a mean of 42.83.



Supplemental Figure 2. Calculation of PKR functional scores from yeast growth assay. (A) Line plot of mock barcode read count data over time for cells expressing nonfunctional (red) or functional (blue) PKR across four sampled timepoints. As PKR activity is toxic to yeast, the number of cells in the pool expressing the functional PKR will decrease over time, and thus the associated barcode read count will also decrease, while the read count for the nonfunctional PKR will increase. (B) Line plot depicting the fold changes from timepoint 0 of barcode abundance derived from Panel A, with a -log2 transformation to assign positive values to functional PKR and negative values to nonfunctional PKR, from which the area under the curve (AUC) is calculated to produce a PKR functional score.



Supplemental Figure 3. Replication of PKR variants paired with K3 alleles. (A) Scatterplot of PKR functional scores for variants paired with K3-WT for two biological replicate experiments. Functional scores were calculated across four timepoints as the area under the curve (see Materials and Methods). Thus, variants with increased evasion of K3 would have higher PKR functional scores, while those with increased susceptibility to K3 or loss of eIF2α kinase activity would have lower scores.

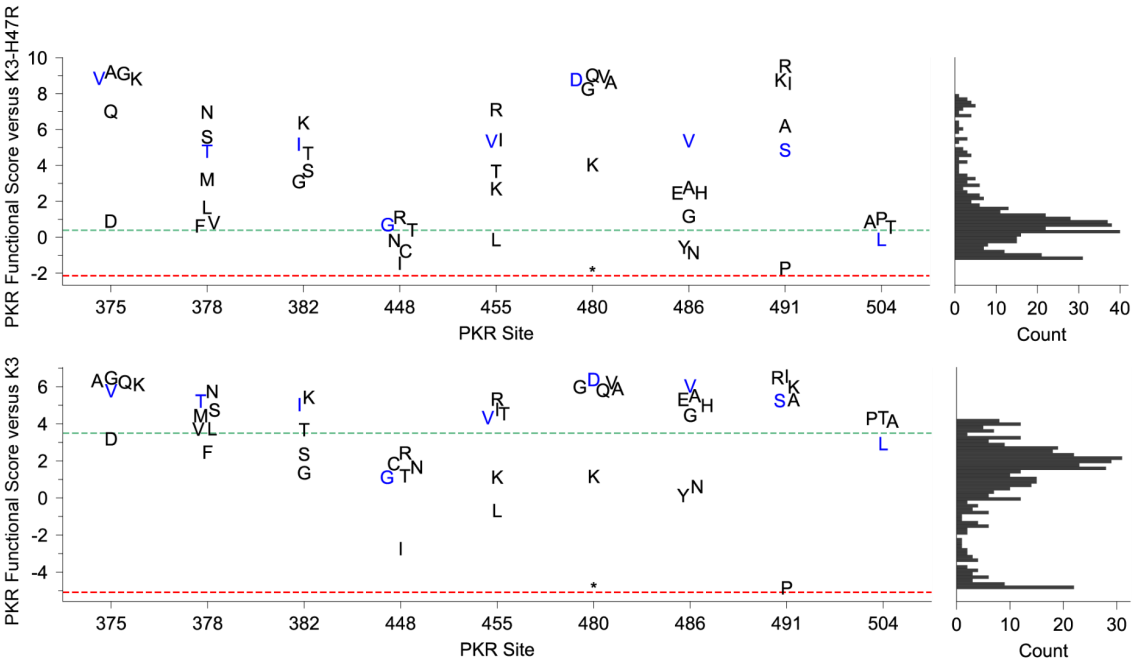


Supplemental Figure 4. Bimodal distribution of PKR functional scores at K3 contact sites. Stripplot of PKR functional scores versus K3 partitioned by predicted contact with K3. Black dashed line denotes the threshold at which variants were separated as nonfunctional-like or functional. Green and red dashed lines represent mean scores for WT PKR and nonsense variants. (** $p < 0.01$, two-sample t-test, comparing just the variants classified as functional in the two classes.)

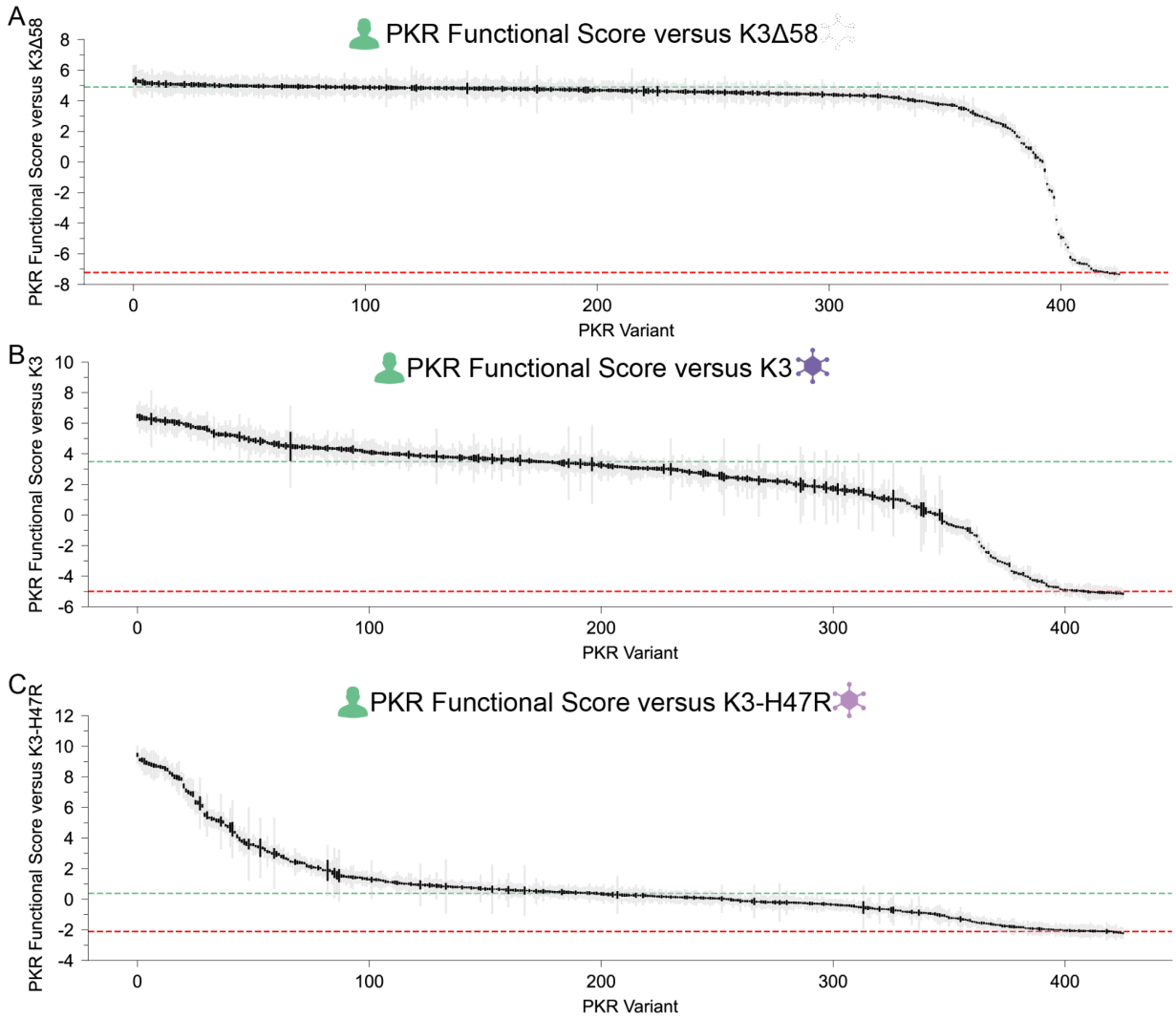
```
                                                                  #
                0         10        20        30        40        50
Human eIF2α     MPGLSCRFYQHKFPEVEDVVMVNVRSIAEMGAYVSLLEYNNIEGMILLSELSR--RRIRS
                     1        10        20         30        40        50
Vaccinia K3     -----MLAFCYSLPNAGDVIKGRVYE-KDYALYIYLFDYPHFEAILAESVKMHMDRYVEY
                         : :.:*:. **:  .* .  : . *: *::* ::*.::  *   .  * :

                  60        70        80        90        100       110
Human eIF2α     INKLIRIGRNECVVVIRVDKEKGYIDLSKRRVSPEEAIKCEDKFTKSKTVYSILRHVAEV
                55     60        70        80
Vaccinia K3     RDKL--VGKTVKVKVIRVDYTKGYIDVNYKRMCRHQ------------------------
                 :**  :*..  * *****  *****:. .*:.  :
```

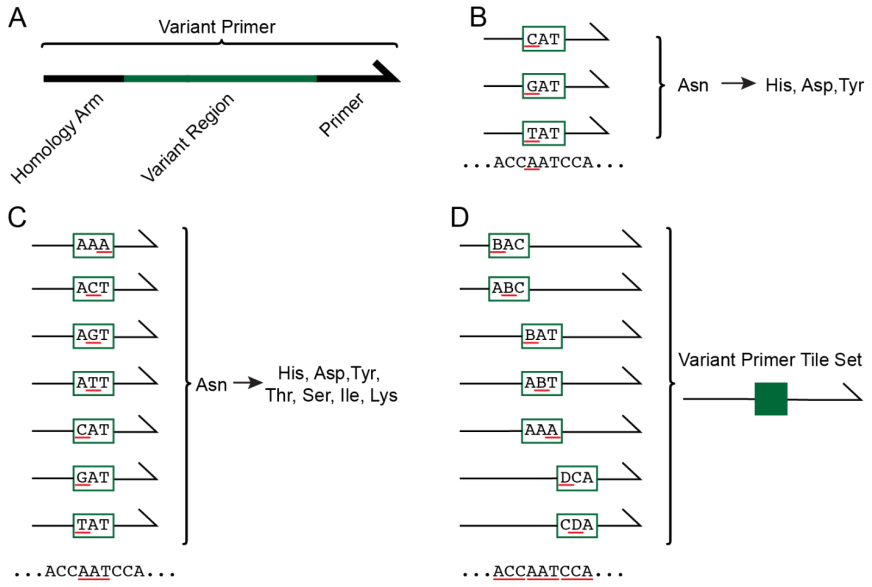
Supplemental Figure 5. Sequence similarity between human eIF2α and vaccinia K3. Alignment of human eIF2α (RefSeq Accession NP_004085.1, residues 0-117) to vaccinia K3 (RefSeq Accession YP_232916.1). The Ser51 site of phosphorylation in eIF2α is indicated by #. Sequences were aligned using Muscle 3.8 and displayed in Clustal format, * = fully conserved, : = strong group conservation, . = weak group conservation. Sites 41-58 of K3 correspond to the rigid helix insert, which is proximal to PKR's ATP-binding site in the Alphafold2-predicted complex, while sites 72-83 are proximal to PKR's helix αG.



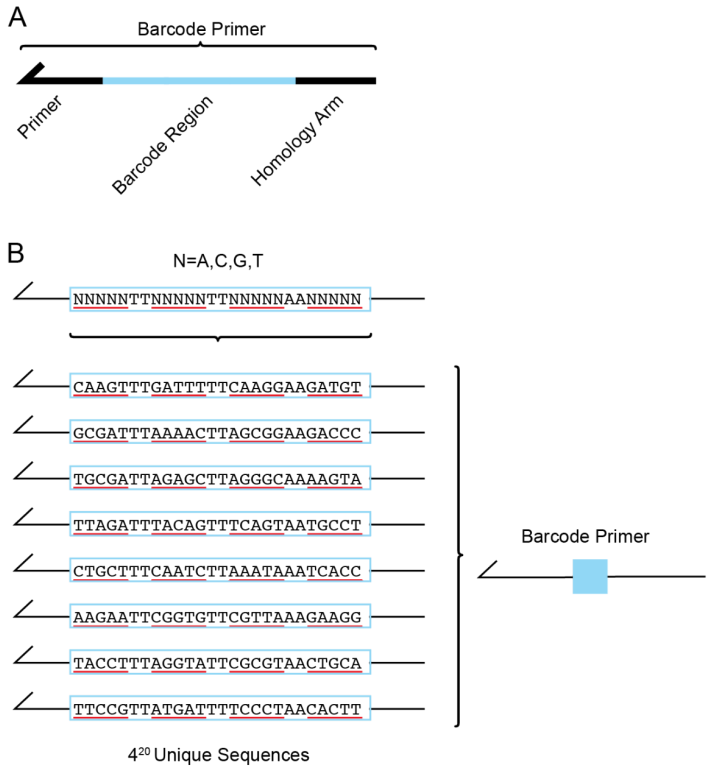
Supplemental Figure 6. Additional variants at the sites of previously identified K3-H47R-resistant variants often also conferred resistance to K3. Stripplots of PKR functional scores of variants paired with K3-H47R (top) and K3-WT (bottom). Each plot highlights PKR variants made at sites where an improved PKR variant was previously identified (28). Blue markers denote previously identified variants; black markers are additional variants made at the same site. Green and red dashed lines represent mean scores for WT PKR and nonsense variants, respectively. Histograms (right) show PKR functional scores versus K3-H47R (top) and K3 (bottom) for all variants. * = stop codon.



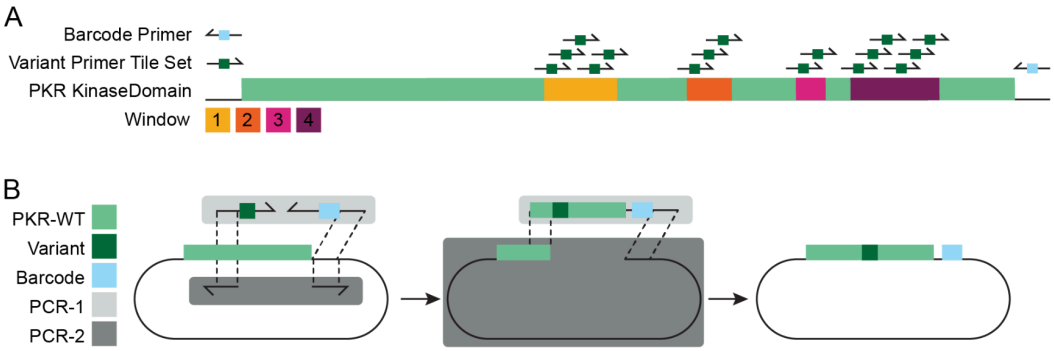
Supplemental Figure 7. Genetic functional resilience of PKR. PKR functional scores versus K3Δ58 (A), K3 (B), and K3-H47R (C), with each vertical line representing a PKR variant. Variants are sorted by functional score from high (left) to low (right), gray and black lines indicate the standard deviation and standard error, respectively in PKR functional scores across barcodes associated with the given variant.



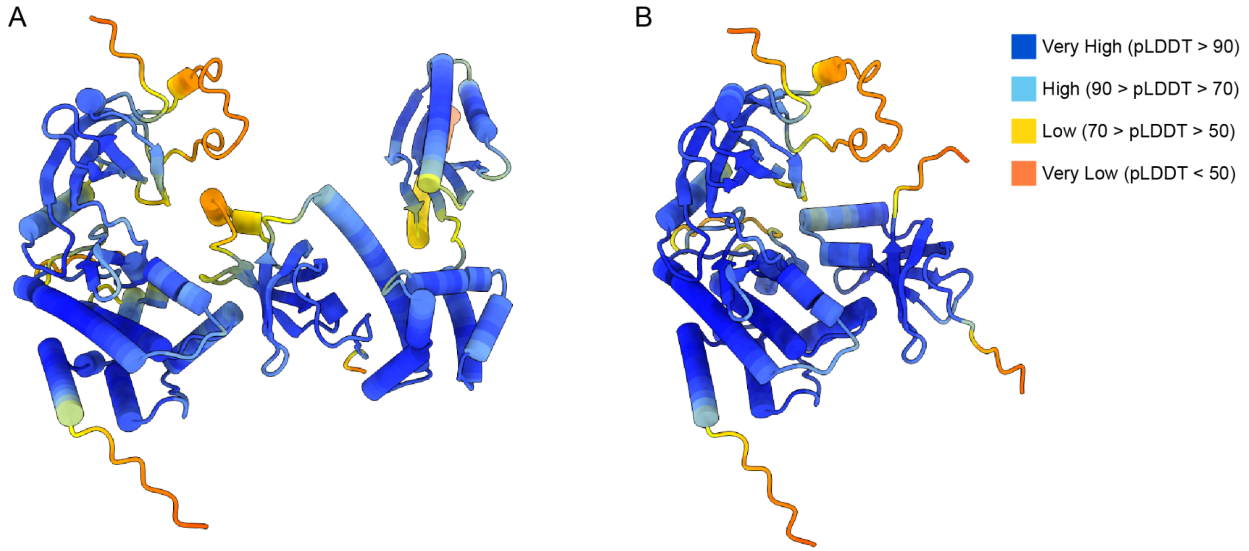
Supplemental Figure 8. Systematic generation of PKR variants using mixed-base primer tile sets. (A) Each variant primer is composed of a homology arm, variant region, and priming region. (B) Nonsynonymous SNP-accessible variants are generated by altering the codon in the variant region of the primer. This example depicts the codon “AAT” encoding Asn. The first nucleotide in the codon, “A”, is underlined in red, with three codons above having changes to “C”, “G”, and “T” underlined in red, which generate the nonsynonymous variants His, Asp, and Tyr. (C) Variant primers were designed across all three nucleotides in each codon, as underlined in red. (D) Variant primer tile sets, represented in dark green, were made by pooling variant primers that modify adjacent codons. Primers included in a given variant primer tile sets have differing variant regions but share homology arms and priming region sequences.



Supplemental Figure 9. Multiple unique barcode sequences were attached to each PKR variant. (A) The barcode primer is composed of a homology arm, barcode region, and priming region. The barcode primer is used as the reverse primer in the variant-generating PCR reactions to attach a unique nucleotide sequence after the PKR variant sequence. (B) The barcode region of the primer is composed of 20 “N”, representing an equal mix of the nucleotides “A”, “C”, “G”, and “T”. A barcode primer with 20 random nucleotides can take on 4^{20} unique nucleotide sequences. Dinucleotide sequences “TT” and “AA” are interspersed throughout the barcode region to avoid making unintended restriction enzyme cut sites.



Supplemental Figure 10. Assembly of PKR variant library using variant tile sets and barcode primers. (A) 15 variant primer tile sets were designed to generate variants across four windows of interest in PKR. The full-length PKR sequence is denoted in green, with Windows 1-4 overlaid in yellow, orange, magenta, and burgundy, respectively. Variant primer tile sets were used to generate 426 PKR missense variants, for a total of 15 PCR-1 reactions. (B) Two separate PCR reactions were used to generate complementary insert and vector fragments. PCR-1 primers (light gray box) include a single variant primer tile set (dark green, see Supplemental Figure 8) and a single doped barcode primer (light blue, see Supplemental Figure 9) that amplified from WT PKR (green) and made the PCR-1 insert fragment containing a select nonsynonymous variants (dark green) and a unique barcodes (blue). PCR-2 primers included a single forward and reverse primer that amplified from WT PKR and made a larger vector fragment with 20-bp homology arms that complement the homology arms of the PCR-1 insert fragment. The two fragments were combined via Gibson Assembly to form a pool of complete vectors, each vector containing a single, nonsynonymous variant with a unique barcode.



Supplemental Figure 11. Alphafold2 Multimer predictions used to identify PKR sites proximal to eIF2α and K3. Cartoon representations of AlphaFold2 Multimer predictions of PKR in complex with eIF2α (A) and K3 (B). Residues are colored AlphaFold2 pLDDT confidence scores per residue. In both panels, PKR is on the left and its binding partner is on the right, as in Figure 1A,B.

SUPPLEMENTAL DATA

Supplementary File 1 - 1_Oligo-Table.xlsx
Supplementary File 2 - 2_Variant-Primers.xlsx
Supplementary File 3 - 3_Plasmids.xlsx
Supplementary File 4 - 4_PKR-Functional-Scores.csv

REFERENCES

1. Daugherty MD, Malik HS. Rules of Engagement: Molecular Insights from Host-Virus Arms Races. Annual Review of Genetics. 2012;46(1):677–700.
2. Metz DH, Esteban M. Interferon inhibits Viral Protein Synthesis in L Cells infected with Vaccinia Virus. Nature. 1972 Aug;238(5364):385–8.
3. Clemens MJ. PKR—A protein kinase regulated by double-stranded RNA. The International Journal of Biochemistry & Cell Biology. 1997 Jul 1;29(7):945–9.
4. Balachandran S, Roberts PC, Brown LE, Truong H, Pattnaik AK, Archer DR, et al. Essential Role for the dsRNA-Dependent Protein Kinase PKR in Innate Immunity to Viral Infection. Immunity. 2000 Jul 1;13(1):129–41.
5. Dar AC, Dever TE, Sicheri F. Higher-Order Substrate Recognition of eIF2α by the RNA-Dependent Protein Kinase PKR. Cell. 2005 Sep 23;122(6):887–900.
6. Weber F, Wagner V, Rasmussen SB, Hartmann R, Paludan SR. Double-Stranded RNA Is Produced by Positive-Strand RNA Viruses and DNA Viruses but Not in Detectable Amounts by Negative-Strand RNA Viruses. Journal of Virology. 2006 May 15;80(10):5059–64.
7. Meurs E, Chong K, Galabru J, Thomas NSB, Kerr IM, Williams BRG, et al. Molecular cloning and characterization of the human double-stranded RNA-activated protein kinase induced by interferon. Cell. 1990 Jul 27;62(2):379–90.
8. Kaufman RJ. Double-stranded RNA-activated protein kinase PKR. Cold Spring Harbor Monograph Series. 2000;39:503–28.
9. Dey M, Cao C, Dar AC, Tamura T, Ozato K, Sicheri F, et al. Mechanistic Link between PKR Dimerization, Autophosphorylation, and eIF2α Substrate Recognition. Cell. 2005 Sep 23;122(6):901–13.
10. Lemaire PA, Anderson E, Lary J, Cole JL. Mechanism of PKR Activation by dsRNA. Journal of molecular biology. 2008;381(2):351–60.
11. Hershey JWB. Translational Control in Mammalian Cells. Annu Rev Biochem. 1991 Jun;60(1):717–55.
12. Sudhakar A, Ramachandran A, Ghosh S, Hasnain SE, Kaufman RJ, Ramaiah KVA. Phosphorylation of Serine 51 in Initiation Factor 2α (eIF2α) Promotes Complex Formation between eIF2α(P) and eIF2B and Causes Inhibition in the Guanine Nucleotide Exchange Activity of eIF2B. Biochemistry. 2000 Oct 1;39(42):12929–38.
13. Langland JO, Cameron JM, Heck MC, Jancovich JK, Jacobs BL. Inhibition of PKR by RNA and DNA viruses. Virus research. 2006;119(1):100–10.
14. Nallagatla SR, Toroney R, Bevilacqua PC. Regulation of innate immunity through RNA structure and the protein kinase PKR. Curr Opin Struct Biol. 2011 Feb;21(1):119–27.
15. Hull CM, Bevilacqua PC. Discriminating Self and Non-Self by RNA: Roles for RNA Structure, Misfolding, and Modification in Regulating the Innate Immune Sensor PKR. Acc Chem Res. 2016 Jun 21;49(6):1242–9.
16. Beattie E, Tartaglia J, Paoletti E. Vaccinia virus-encoded eIF-2α homolog abrogates the antiviral effect of interferon. Virology. 1991;183(1):419–22.
17. Carroll K, Elroy-Stein O, Moss B, Jagus R. Recombinant vaccinia virus K3L gene product prevents activation of double-stranded RNA-dependent, initiation factor 2 alpha-specific protein kinase. Journal of Biological Chemistry. 1993 Jun 15;268(17):12837–42.
18. Davies MV, Chang HW, Jacobs BL, Kaufman RJ. The E3L and K3L vaccinia virus gene products stimulate translation through inhibition of the double-stranded RNA-dependent protein kinase by different mechanisms. Journal of Virology. 1993 Mar;67(3):1688–92.
19. Sharp TV, Witzel JE, Jagus R. Homologous Regions of the α Subunit of Eukaryotic Translational Initiation Factor 2 (eIF2α) and the Vaccinia Virus K3L Gene Product Interact with the Same Domain within the dsRNA-Activated Protein Kinase (PKR). European Journal of Biochemistry. 1997;250(1):85–91.

20. Dar AC, Sicheri F. X-Ray Crystal Structure and Functional Analysis of Vaccinia Virus K3L Reveals Molecular Determinants for PKR Subversion and Substrate Recognition. Molecular Cell. 2002 Aug 1;10(2):295–305.
21. Kawagishi-Kobayashi M, Silverman JB, Ung TL, Dever TE. Regulation of the protein kinase PKR by the vaccinia virus pseudosubstrate inhibitor K3L is dependent on residues conserved between the K3L protein and the PKR substrate eIF2alpha. Molecular and Cellular Biology. 1997 Jul 1;17(7):4146–58.
22. Rothenburg S, Seo EJ, Gibbs JS, Dever TE, Dittmar K. Rapid evolution of protein kinase PKR alters sensitivity to viral inhibitors. Nature Structural & Molecular Biology. 2009 Jan;16(1):63–70.
23. Elde NC, Child SJ, Geballe AP, Malik HS. Protein kinase R reveals an evolutionary model for defeating viral mimicry. Nature. 2009 Jan;457(7228):485–9.
24. Jacquet S, Culbertson M, Zang C, Filali AE, Mory CDLM, Pons JB, et al. Adaptive duplication and functional diversification of Protein kinase R contribute to the uniqueness of bat-virus interactions [Internet]. bioRxiv; 2022 [cited 2022 Sep 30]. p. 2022.06.28.497829. Available from: https://www.biorxiv.org/content/10.1101/2022.06.28.497829v1
25. Chong K l., Feng L, Schappert K, Meurs E, Donahue T f., Friesen J d., et al. Human p68 kinase exhibits growth suppression in yeast and homology to the translational regulator GCN2. The EMBO Journal. 1992 Apr 1;11(4):1553–62.
26. Dever TE, Chen JJ, Barber GN, Cigan AM, Feng L, Donahue TF, et al. Mammalian eukaryotic initiation factor 2 alpha kinases functionally substitute for GCN2 protein kinase in the GCN4 translational control mechanism of yeast. PNAS. 1993 May 15;90(10):4616–20.
27. Elde NC, Child SJ, Eickbush MT, Kitzman JO, Rogers KS, Shendure J, et al. Poxviruses Deploy Genomic Accordions to Adapt Rapidly against Host Antiviral Defenses. Cell. 2012 Aug 17;150(4):831–41.
28. Seo EJ, Liu F, Kawagishi-Kobayashi M, Ung TL, Cao C, Dar AC, et al. Protein kinase PKR mutants resistant to the poxvirus pseudosubstrate K3L protein. Proceedings of the National Academy of Sciences. 2008 Nov 4;105(44):16894–9.
29. Dey M, Velyvis A, Li JJ, Chiu E, Chiovitti D, Kay LE, et al. Requirement for kinase-induced conformational change in eukaryotic initiation factor 2α (eIF2α) restricts phosphorylation of Ser51. Proceedings of the National Academy of Sciences. 2011 Mar 15;108(11):4316–21.
30. Haller SL, Peng C, McFadden G, Rothenburg S. Poxviruses and the evolution of host range and virulence. Infection, Genetics and Evolution. 2014 Jan 1;21:15–40.
31. Rothenburg S, Chinchar VG, Dever TE. Characterization of a ranavirus inhibitor of the antiviral protein kinase PKR. BMC Microbiology. 2011 Mar 18;11(1):56.
32. Sawyer SL, Wu LI, Emerman M, Malik HS. Positive selection of primate TRIM5α identifies a critical species-specific retroviral restriction domain. Proceedings of the National Academy of Sciences. 2005 Feb 22;102(8):2832–7.
33. Tenthorey JL, Young C, Sodeinde A, Emerman M, Malik HS. Mutational resilience of antiviral restriction favors primate TRIM5α in host-virus evolutionary arms races. Schoggins JW, Weigel D, Berthoux L, editors. eLife. 2020 Sep 15;9:e59988.
34. Demogines A, Farzan M, Sawyer SL. Evidence for ACE2-Utilizing Coronaviruses (CoVs) Related to Severe Acute Respiratory Syndrome CoV in Bats. Journal of Virology. 2012 Jun 1;86(11):6350–3.
35. Damas J, Hughes GM, Keough KC, Painter CA, Persky NS, Corbo M, et al. Broad host range of SARS-CoV-2 predicted by comparative and structural analysis of ACE2 in vertebrates. PNAS. 2020 Sep 8;117(36):22311–22.
36. Lan J, Ge J, Yu J, Shan S, Zhou H, Fan S, et al. Structure of the SARS-CoV-2 spike receptor-binding domain bound to the ACE2 receptor. Nature. 2020 May;581(7807):215–20.
37. Chan KK, Dorosky D, Sharma P, Abbasi SA, Dye JM, Kranz DM, et al. Engineering human ACE2 to optimize binding to the spike protein of SARS coronavirus 2. Science. 2020 Sep

4;369(6508):1261–5.
38. Heinzelman P, Romero PA. Discovery of human ACE2 variants with altered recognition by the SARS-CoV-2 spike protein. PLOS ONE. 2021 May 12;16(5):e0251585.
39. Ogden PJ, Kelsic ED, Sinai S, Church GM. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. Science. 2019 Nov 29;366(6469):1139–43.
40. Younger D, Berger S, Baker D, Klavins E. High-throughput characterization of protein–protein interactions by reprogramming yeast mating. PNAS. 2017 Nov 14;114(46):12166–71.
41. Liu X, Liu Z, Dziulko AK, Li F, Miller D, Morabito RD, et al. iSeq 2.0: A Modular and Interchangeable Toolkit for Interaction Screening in Yeast. Cell Systems. 2019 Apr 24;8(4):338-344.e8.
42. Dadonaite B, Crawford KHD, Radford CE, Farrell AG, Yu TC, Hannon WW, et al. A pseudovirus system enables deep mutational scanning of the full SARS-CoV-2 spike. Cell. 2023 Mar 16;186(6):1263-1278.e20.
43. Parvez S, Herdman C, Beerens M, Chakraborti K, Harmer ZP, Yeh JRJ, et al. MIC-Drop: A platform for large-scale in vivo CRISPR screens. Science. 2021 Sep 3;373(6559):1146–51.
44. Popp NA, Powell RL, Wheelock MK, Zapp BD, Holmes KJ, Sheldon KM, et al. Multiplex, multimodal mapping of variant effects in secreted proteins [Internet]. bioRxiv; 2024 [cited 2024 Apr 30]. p. 2024.04.01.587474. Available from: https://www.biorxiv.org/content/10.1101/2024.04.01.587474v1
45. Park C, Peng C, Brennan G, Rothenburg S. Species-specific inhibition of antiviral protein kinase R by capripoxviruses and vaccinia virus. Annals of the New York Academy of Sciences. 2019;1438(1):18–29.
46. Park C, Peng C, Rahman MJ, Haller SL, Tazi L, Brennan G, et al. Orthopoxvirus K3 orthologs show virus- and host-specific inhibition of the antiviral protein kinase PKR. PLOS Pathogens. 2021 Jan 14;17(1):e1009183.
47. Peng C, Haller SL, Rahman MM, McFadden G, Rothenburg S. Myxoma virus M156 is a specific inhibitor of rabbit PKR but contains a loss-of-function mutation in Australian virus isolates. PNAS. 2016 Apr 5;113(14):3855–60.
48. Rose MD, Novick P, Thomas JH, Botstein D, Fink GR. A *Saccharomyces cerevisiae* genomic plasmid bank based on a centromere-containing shuttle vector. Gene. 1987 Jan 1;60(2):237–43.
49. Curran KA, Morse NJ, Markham KA, Wagman AM, Gupta A, Alper HS. Short Synthetic Terminators for Improved Heterologous Gene Expression in Yeast. ACS Synth Biol. 2015 Jul 17;4(7):824–32.
50. Brachmann CB, Davies A, Cost GJ, Caputo E, Li J, Hieter P, et al. Designer deletion strains derived from Saccharomyces cerevisiae S288C: A useful set of strains and plasmids for PCR-mediated gene disruption and other applications. Yeast. 1998;14(2):115–32.
51. Cornish-Bowden A. Nomenclature for incompletely specified bases in nucleic acid sequences: recommendations 1984. Nucleic acids research. 1985;13(9):3021.
52. Guido VR, Drake Jr FL. Python 3 reference manual. Scotts Valley: CreateSpace. 2009;
53. Crawford KH d, Bloom JD. alignparse: A Python package for parsing complex features from high-throughput long-read sequencing. Journal of Open Source Software. 2019 Dec 11;4(44):1915.
54. Sambrook J, Russell DW. Standard Ethanol Precipitation of DNA in Microcentrifuge Tubes. Cold Spring Harb Protoc. 2006 Jun 1;2006(1):pdb.prot4456.
55. Gietz RD, Schiestl RH. Large-scale high-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method. Nat Protoc. 2007;2(1):38–41.
56. Singh MV, Anthony Weil P. A method for plasmid purification directly from yeast. Analytical Biochemistry. 2002 Aug 1;307(1):13–7.

57. Wu L, Wen C, Qin Y, Yin H, Tu Q, Van Nostrand JD, et al. Phasing amplicon sequencing on Illumina Miseq for robust environmental microbial community analysis. BMC Microbiol. 2015 Dec;15(1):1–12.
58. Zhang J, Kobert K, Flouri T, Stamatakis A. PEAR: a fast and accurate Illumina Paired-End reAd mergeR. Bioinformatics. 2014 Mar 1;30(5):614–20.
59. Zhao L, Liu Z, Levy SF, Wu S. Bartender: a fast and accurate clustering algorithm to count barcode reads. Bioinformatics. 2018 Mar 1;34(5):739–47.
60. Meng EC, Goddard TD, Pettersen EF, Couch GS, Pearson ZJ, Morris JH, et al. UCSF ChimeraX: Tools for structure building and analysis. Protein Science. 2023;32(11):e4792.
61. Mirdita M, Schütze K, Moriwaki Y, Heo L, Ovchinnikov S, Steinegger M. ColabFold: making protein folding accessible to all. Nat Methods. 2022 Jun;19(6):679–82.