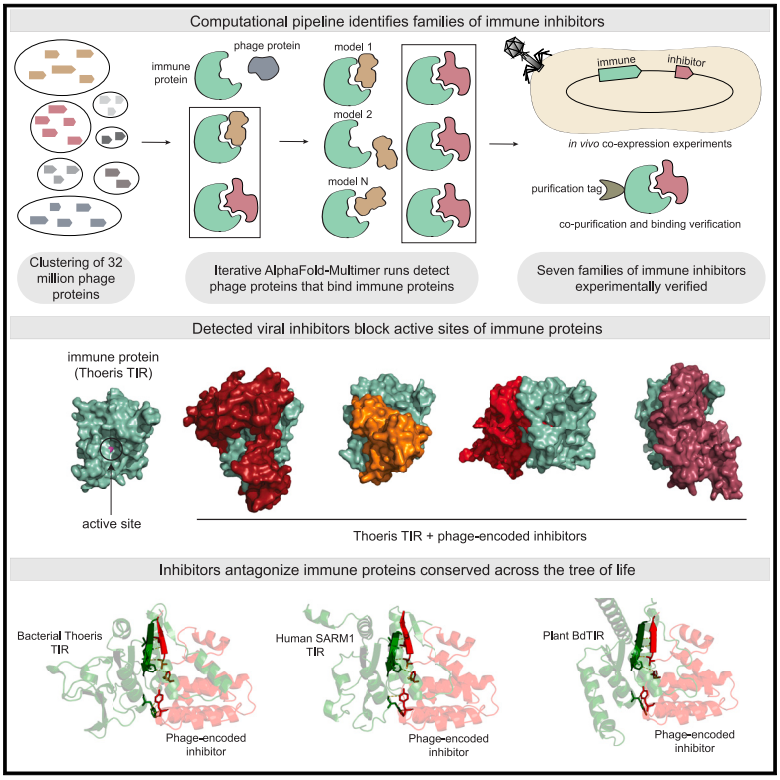
Cell

Article

Structure-guided discovery of viral proteins that inhibit host immunity

Graphical abstract



Highlights

- Structure-guided pipeline sifts through >30 M phage proteins to detect immune inhibitors
- Detected inhibitors bind TIR and cGAS-like proteins and block their active sites
- Phage-derived inhibitors also block human and plant TIRs as well as the human cGAS
- Phage genomes contain proteins that can inhibit bacterial, plant, and animal immunity

Authors

Erez Yirmiya, Samuel J. Hobbs, Azita Leavitt, ..., Gil Amitai, Philip J. Kranzusch, Rotem Sorek

Correspondence

rotem.sorek@weizmann.ac.il

In brief

Large-scale, structure-guided computational pipeline sifts through millions of phage proteins to identify those that bind and antagonize host immune proteins. Detected proteins are shown to inhibit not only bacterial but also plant and animal immune proteins.

Yirmiya et al., 2025, Cell *188*, 1–12
March 20, 2025
https://doi.org/10.1016/j.cell.2024.12.035

CellPress

Cell

CellPress

Article

Structure-guided discovery of viral proteins that inhibit host immunity

Erez Yirmiya,[1] Samuel J. Hobbs,[2,3] Azita Leavitt,[1] Ilya Osterman,[1] Carmel Avraham,[1] Dina Hochhauser,[1] Barak Madhala,[1] Marharyta Skovorodka,[1] Joel M.J. Tan,[2,3] Hunter C. Toyoda,[2,3] Igor Chebotar,[4] Maxim Itkin,[5] Sergey Malitsky,[5] Gil Amitai,[1] Philip J. Kranzusch,[2,3] and Rotem Sorek[1,6,*]

[1]Department of Molecular Genetics, Weizmann Institute of Science, Rehovot 7610001, Israel
[2]Department of Microbiology, Harvard Medical School, Boston, MA 02115, USA
[3]Department of Cancer Immunology and Virology, Dana-Farber Cancer Institute, Boston, MA 02115, USA
[4]High Performance Computing Section, Weizmann Institute of Science, Rehovot 7610001, Israel
[5]Life Sciences Core Facilities, Weizmann Institute of Science, Rehovot 7610001, Israel
[6]Lead contact
*Correspondence: rotem.sorek@weizmann.ac.il
https://doi.org/10.1016/j.cell.2024.12.035

SUMMARY

Viruses encode proteins that inhibit host defenses, but sifting through the millions of available viral sequences for immune-modulatory proteins has been so far impractical. Here, we develop a process to systematically screen virus-encoded proteins for inhibitors that physically bind host immune proteins. Focusing on Thoeris and CBASS, bacterial defense systems that are the ancestors of eukaryotic Toll/interleukin-1 receptor (TIR) and cyclic GMP-AMP synthase (cGAS) immunity, we discover seven families of Thoeris and CBASS inhibitors, encompassing thousands of genes widespread in phages. Verified inhibitors exhibit extensive physical interactions with the respective immune protein counterpart, with all inhibitors blocking the active site of the immune protein. Remarkably, a phage-encoded inhibitor of bacterial TIR proteins can bind and inhibit distantly related human and plant immune TIRs, and a phage-derived inhibitor of bacterial cGAS-like enzymes can inhibit the human cGAS. Our results demonstrate that phages are a reservoir for immune-modulatory proteins capable of inhibiting bacterial, animal, and plant immunity.

INTRODUCTION

Bacteria employ a highly diverse set of defense systems to resist phage infections,[1–5] some of which are the ancestors of innate immune pathways in animals and plants.[6–11] Phages, in return, evolved diverse anti-defense proteins to overcome bacterial defenses.[12–14] Phage proteins were demonstrated to inhibit immunity by binding immune complexes,[15–20] by sequestering or cleaving immune signaling molecules,[21–26] by enzymatically modifying immune proteins,[27–29] and via additional mechanisms.[13,14]

Current phage sequence databases contain millions of phage genomes, and in most of these sequenced genomes, the majority of genes are of unknown function.[30–32] Even in model phages, such as T4, T5, and T7, that have been extensively studied for many decades, the functions of 30%–60% of the genes remain obscure.[33–35] It was recently estimated that many of these genes may serve to inhibit bacterial defenses,[1,13,36] but there is currently no systematic methodology to predict which of the phage genes inhibits which defense system.

In this study, we present a computational pipeline that makes use of AlphaFold2 structural predictions to identify immune-modulating proteins within large phage protein databases. Our pipeline relies on the AlphaFold2-Multimer function, a software that predicts physical interactions between input proteins.[37] Combining a dimension-reduction process with a set of reliability measures we developed, we establish a pipeline that identifies potent inhibitors of bacterial immune proteins and show that some of these inhibitors are capable of binding and inhibiting human and plant immune proteins as well.

RESULTS

A computational pipeline for the discovery of anti-defense proteins

To access the vast diversity of protein functions encoded by phages, we analyzed ~32 million unique phage proteins from ~2 million phage genome scaffolds present in the Integrated Microbial Genomes/Virus (IMG/VR) v3 phage genome database.[38] We clustered these proteins based on sequence similarity and screened the resulting clusters via several filtering steps (STAR Methods), requiring sufficient sequence coverage, no known functional annotation, and considering only clusters with proteins shorter than 200 amino acids because most of the anti-defense proteins discovered to date are small.[14] We predicted the three-dimensional structure of a representative sequence from

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

each cluster and retained only clusters generating structures with reliable folding prediction (STAR Methods). This analysis retrieved ~38,700 clusters, each representing a phage protein family with at least 40 homologs (Figure 1A).

To examine whether AlphaFold2-Multimer[37] could be used to discover phage proteins that inhibit bacterial immunity, we considered the type I Thoeris defense system from *Bacillus cereus* MSX-D12, a two-gene system that protects against a broad array of phages and whose function is well understood.[7,22] The Thoeris system encodes ThsB, a TIR-domain protein that generates a signaling molecule once it detects phage infection, and ThsA, an effector NAD^+-cleaving protein that is activated by the signaling molecule. A representative sequence from each of the ~38,700 phage protein clusters was co-folded together with each of the 2 proteins of Thoeris, using an iterative process aimed to reduce computation loads (Figure 1B). Possible interactions between the Thoeris proteins and each phage protein were ranked based on the model confidence score of AlphaFold2-Multimer,[37] ultimately selecting phage proteins whose scores were consistently high across 25 co-folding predictions generated by AlphaFold2-Multimer (Figure 1B).

Phage proteins that inhibit type I Thoeris

We detected 16 phage proteins whose predicted co-folding scores with one of the Thoeris proteins passed the cutoff threshold. Fifteen of these proteins were predicted to bind ThsB, the TIR-domain-containing protein responsible for sensing phage infection,[7,22] and one was predicted to bind the Thoeris immune effector ThsA (Table S1). The DNA sequence for each of the candidate immune inhibitor proteins was synthesized together with an inducible promoter and integrated into the genome of a *Bacillus subtilis* strain also carrying the Thoeris system.[2] Using phage infection assays, we found that four out of the sixteen tested phage proteins inhibited the activity of Thoeris, as phage infection was no longer blocked by Thoeris in the presence of the anti-defense proteins (Figure 1C). We named these four Thoeris anti-defense proteins Tad3, Tad4, Tad5, and Tad6.

All four verified anti-Thoeris proteins were predicted to bind the TIR-domain protein ThsB. In the type I Thoeris defense system, ThsB senses phage infection and then produces the immune signaling molecule 1″-3′ glycocyclic ADPR (gcADPR, also called 3'cADPR).[22] This molecule binds and activates the NADase activity of ThsA, which then depletes the cell of the essential molecule NAD^+[7,22,39] (Figure 1D). To test whether the anti-Thoeris proteins specifically inhibit ThsB as predicted, we expressed each of these proteins in cells expressing ThsB only, without an active ThsA protein, and infected these cells with phage SBSphiJ. We then lysed the infected cells, filtered the lysates to include only small molecules, and incubated a purified ThsA protein with these lysates *in vitro*. Filtered lysates extracted from infected cells that expressed only ThsB were able to induce the NADase activity of ThsA, but lysates from cells also expressing Tad3, Tad4, Tad5, or Tad6 failed to activate ThsA (Figure 1E). These results indicate that the four anti-Thoeris proteins inhibit the activity of ThsB, as predicted from the AlphaFold2-Multimer binding predictions.

To test whether Tad3, Tad4, Tad5, and Tad6 indeed directly bind ThsB, we co-expressed each of these proteins with a 6× His-SUMO2-tagged ThsB and assessed interaction in a pull-down assay. In all four cases, ThsB co-purified with the respective anti-Thoeris protein (Figure 1F), and mass spectrometry (MS) analyses of the gel areas containing the pulled-down proteins confirmed the presence of the respective anti-Thoeris protein in all four cases (Tables S2, S3, S4, S5, and S6). Moreover, size-exclusion chromatography with multi-angle light scattering (SEC-MALS) confirmed that each of these proteins forms a complex with ThsB (Figure S1). These results confirm that the anti-Thoeris proteins discovered here form a protein complex to directly inhibit ThsB.

To further interrogate whether the proteins we detected represent bona fide phage inhibitors of Thoeris, we focused on one of the proteins, Tad3. We engineered *tad3* into the genome of SBSphiJ, a phage normally blocked by Thoeris, under the control of the native promoter of *tad1*, a previously identified Thoeris inhibitor.[22] SBSphiJ knocked-in for *tad3* became fully resistant to Thoeris, showing that *tad3* expression from the phage genome is sufficient for Thoeris inhibition (Figure 2A). We next determined a crystal structure of the ThsB-Tad3 complex at a resolution of 1.8 Å in order to examine the accuracy of the interactions predicted by AlphaFold2-Multimer (Figures 2B–2D). Analysis of the ThsB-Tad3 crystal structure confirmed that the AlphaFold2-predicted interactions were virtually identical to those in the experimentally determined structure (Figure 2E), but also surprisingly revealed that Tad3 forms a homodimer that simultaneously binds two ThsB monomers (Figures 2B and 2C). α helices at the N terminus of Tad3 interlock to mediate homodimer formation, and the C terminus of each Tad3 protomer makes extensive contacts with individual ThsB protomers (Figures 2B and 2C). The Tad3 interlocking N-terminal helices also contain a highly conserved serine residue (S48) that extends into the TIR-domain active site pocket within the opposing ThsB protomer and forms a hydrogen bond with the essential ThsB catalytic residue E85, explaining how gcADPR synthesis by ThsB is inhibited (Figure 2D).

Examining the predicted structural models for the interactions between ThsB and the other phage-derived binders, we observe that Tad4 and Tad6 also directly block the active site pocket of ThsB via a loop that interacts with the active site residues (Figure 2F). The amino acid residues that block the ThsB active site pocket are highly conserved among homologs of each of Tad3, Tad4, and Tad6, suggesting that blockage of the active site of ThsB is a conserved function of these anti-Thoeris proteins (Figure S2). Tad5 is also predicted to interact with E85 and other active site residues in ThsB but does not completely block the active site pocket (Figure 2F).

We mutated the loop residues in Tad3, Tad4, and Tad6 that are predicted to block the catalytic site of ThsB. Removal of the loop or a point mutation in the respective Tad residue predicted to directly interact with the E85 active site of ThsB impaired the function of Tad4 and Tad6 as anti-Thoeris proteins (Figure S3). Similar mutations in Tad3 were not sufficient to impair its anti-Thoeris activity (Figure S3). Possibly, the extended interface that the Tad3 dimer generates with ThsB is sufficient for ThsB inhibition even when the active site-interacting loop of Tad3 is mutated (Figure 2B–2D).

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

CellPress

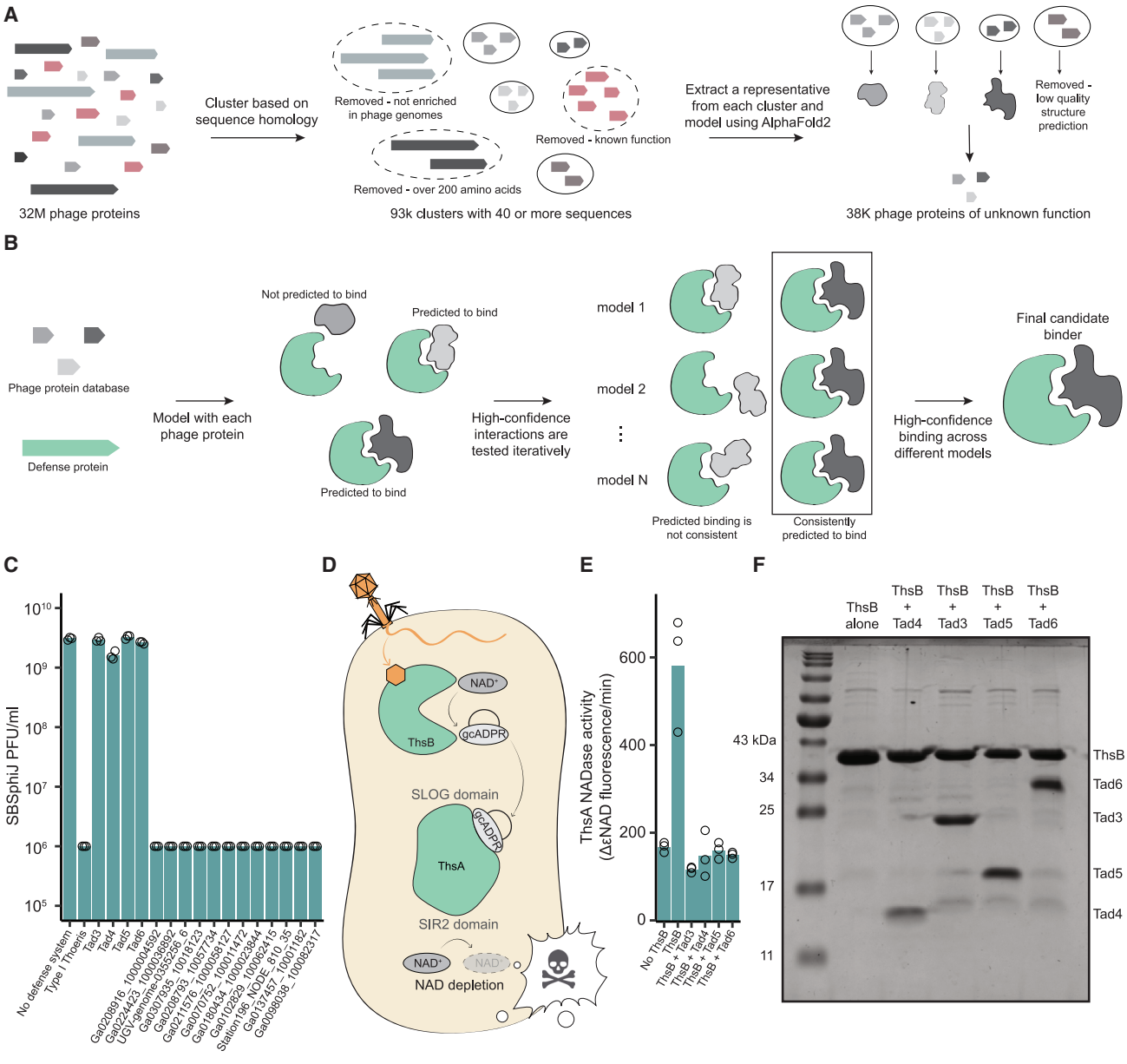


Figure 1. Structure-guided discovery of protein inhibitors of type I Thoeris

(A) Clustering the phage protein space and establishing a database of phage clusters representing short proteins of unknown function.

(B) An iterative pipeline for the prediction of interactions between a bacterial immune protein and phage proteins. High-scoring interactions based on a single AlphaFold2-Multimer model are further examined via multiple models.

(C) Anti-defense activity of the anti-Thoeris candidates. Data represent plaque-forming units per milliliter (PFU/mL) of phage SBSphiJ infecting control cells (no system), cells expressing the Thoeris system (Thoeris), and cells co-expressing the Thoeris system and each of the anti-Thoeris candidates. Shown is the average of three replicates, with individual data points overlaid.

(D) Schematic of the mechanism of the Thoeris defense system.

(E) Anti-defense proteins inhibit ThsB. Cells expressing ThsB, co-expressing ThsB and each of the anti-Thoeris proteins, or control cells that do not express ThsB, were infected with phage SBSphiJ at a multiplicity of infection (MOI) of 10. Lysates from infected cells were extracted and filtered, and NADase activity of purified ThsA incubated with filtered lysates was measured using a nicotinamide 1,N6-ethenoadenine dinucleotide (εNAD) cleavage fluorescence assay. Bars represent the mean of three experiments, with individual data points overlaid.

(F) Anti-defense proteins bind ThsB. Pull-down of a 6×His-SUMO2-tagged ThsB protein co-expressed with Tad3, Tad4, Tad5, or Tad6 retrieves the respective anti-Thoeris protein. Shown is SDS-PAGE analysis.

See also Figures S1 and S8.

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

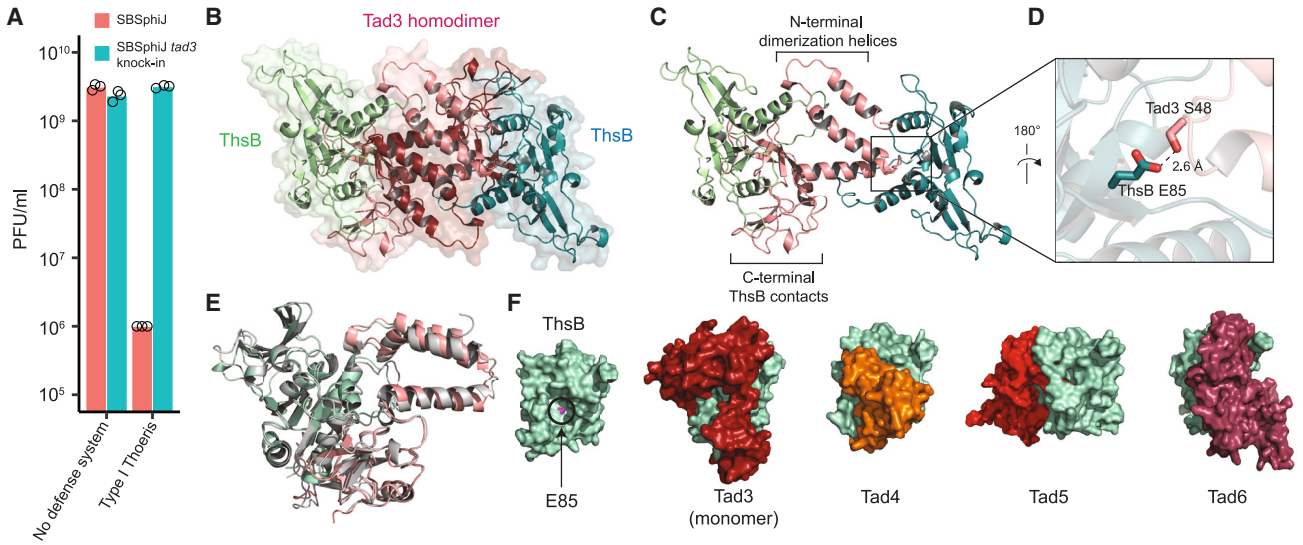


Figure 2. Anti-Thoeris proteins bind ThsB

(A) Knockin of *tad3* in phage SBSphiJ renders the phage resistant to Thoeris. Data represent PFUs/mL of phages infecting control cells (no defense system) and cells expressing the Thoeris defense system. Shown is the average of three replicates, with individual data points overlaid.

(B) Overview of the Tad3:ThsB crystal structure. Tad3 forms a homodimer (monomers shown in shades of red) that binds two ThsB enzymes (shown in shades of green and blue).

(C) Overview of a single Tad3 protomer bound to both ThsB protomers with interacting regions highlighted.

(D) Detailed view of Tad3 blocking the active site of ThsB. Tad3 forms a loop that penetrates the active site and forms a hydrogen bond with the catalytic E85 residue.

(E) Comparison of the solved crystal structure of one ThsB:Tad3 heterodimer (green and red) and the predicted structure generated by AlphaFold2-Multimer (in gray). Root-mean-square deviation (RMSD) values of the predicted and crystal structures are 1.33/1.4 Å for the ThsB/Tad3 structures, respectively.

(F) Predicted complex structures of ThsB bound to each one of the anti-Thoeris proteins Tad3, Tad4, Tad5, and Tad6. ThsB is presented in the same orientation, with the active site front-facing, in all cases.

See also Figures S2, S3, and S8.

Refinement of the computational pipeline

Examining viral protein homologs of each of the four verified inhibitors by AlphaFold2-Multimer modeling demonstrated that these homologs were also predicted to bind ThsB with high confidence (Figure 3A). Remarkably, when repeating this analysis for the 12 candidates that failed to inhibit Thoeris when tested experimentally, we noticed that in most cases homologs of the candidate phage protein were not predicted by AlphaFold2-Multimer to bind the immune protein (Figure 3A; Table S7). In some cases, even close homologs with >80% sequence identity when compared with the tested candidate showed poor binding predictions when analyzed by AlphaFold2-Multimer, suggesting that the original prediction for binding was spurious (Figure 3A). These results suggested that computational analysis of homologs could be used to increase chances for the discovery of biologically meaningful protein-protein interactions via AlphaFold2-Multimer analyses.

In an additional attempt to better define parameters predictive of true positive hits, we compared the predicted structures of the four verified inhibitors to those of the 12 tested candidates that did not inhibit Thoeris when tested experimentally. While all verified inhibitors presented an ample binding surface with the host ThsB binding partner, some of the non-verified candidates were predicted to bind only via a relatively small surface patch (Figures 3B and 3C; Table S8). These observations additionally suggested that potential interactions involving large surface areas between the immune protein and the candidate phage protein may be a predictive factor to increase true positive hit rate.

Phage proteins that inhibit type II Thoeris and the CBASS system

Based on the above observations, we updated our computational pipeline to consider the binding scores for homologs of each candidate, as well as the surface area of the interaction (Figure 4A). Within the updated pipeline, high-scoring phage proteins are tested experimentally as inhibitors only if co-folding of protein homologs with the target immune protein also results in high scores and only if the phage protein co-folds with its counterpart via extensive surface interactions.

To test whether the updated pipeline can discover phage-derived immunity inhibitors with a higher rate of success, we applied it to the type II Thoeris system from *Bacillus amyloliquefaciens* Y2. This system employs a ThsB TIR-domain protein that generates a histidine conjugated to ADPR (His-ADPR) as an immune signaling molecule.[40] The ThsA effector protein of type II Thoeris encodes a macro domain capable of binding His-ADPR and a transmembrane-spanning domain that likely impairs membrane integrity once activated by the signaling molecule.[2,40]

The ThsB TIR-domain protein of the type II Thoeris system is substantially different to that of type I Thoeris, with no detectable

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

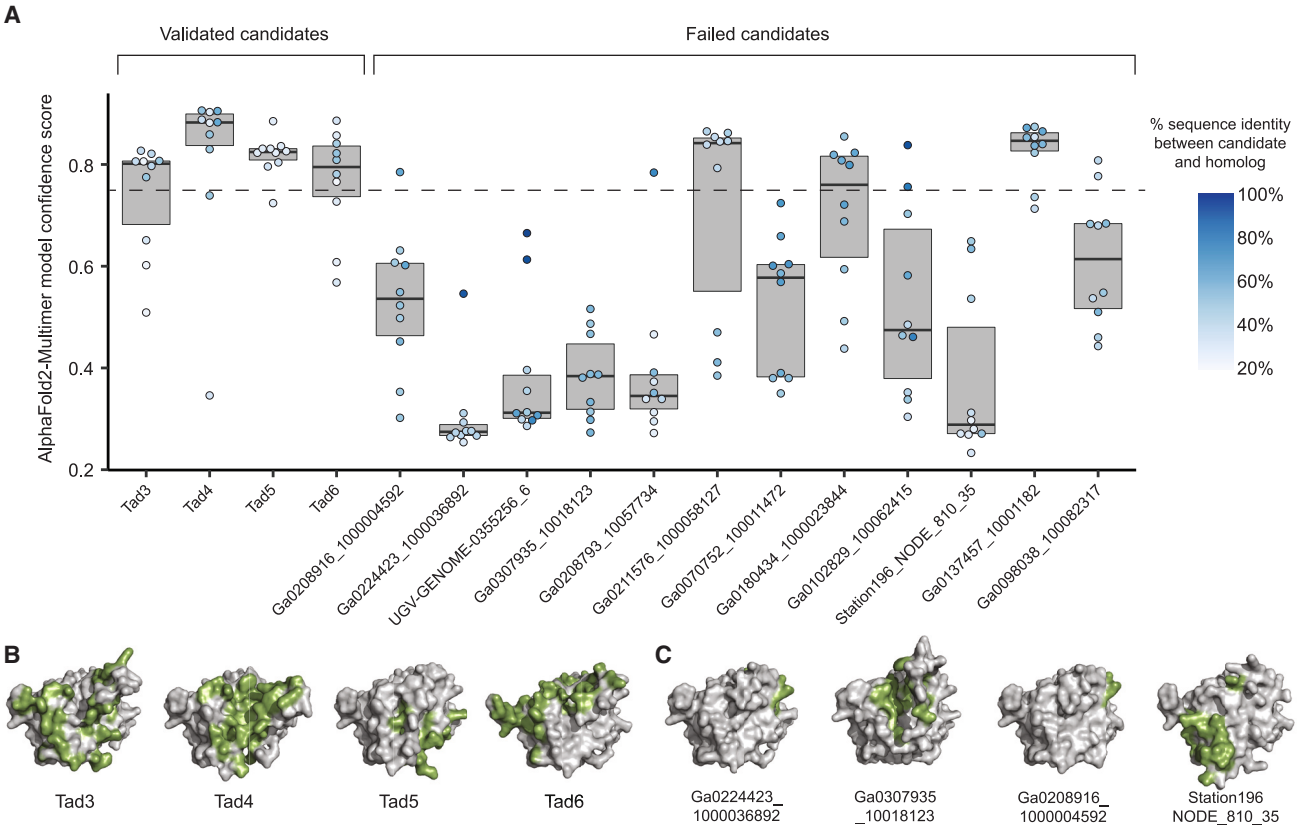


Figure 3. Comparing verified inhibitors to non-verified candidates

(A) Homologs of verified Thoeris inhibitors are also predicted to bind ThsB when analyzed via AlphaFold2-Multimer, but homologs of most of the non-verified candidates are not. For each of the 16 proteins tested as candidate anti-Thoeris inhibitors, 10 homologs ranging in sequence identity between 25% and 95% were analyzed by AlphaFold2-Multimer. Presented are the AlphaFold2-Multimer model confidence scores for protein-protein interactions of each homolog when co-folded with the respective immune protein. Scores are the averages of 25 co-folding predictions. Sequence homology between each candidate and its 10 homologs is presented in a white-to-blue scale.

(B and C) Depiction of ThsB surfaces predicted to interact with each of the anti-Thoeris proteins (B) or with false-positive candidate anti-Thoeris proteins that did not inhibit Thoeris when tested experimentally (C). Surface on ThsB predicted to interact with the respected anti-defense protein is marked in green. Structure of ThsB differs between models, as binding of the anti-Thoeris protein sometimes affects the predicted ThsB structure.

sequence similarity between the two and only little structural similarity (Figure S4). As expected from the structural divergence between the TIR-domain proteins, AlphaFold2-Multimer did not predict high-scoring interactions between ThsB of type II Thoeris and the anti-type I Thoeris proteins identified in this study, and, consistently, inhibitors of type I Thoeris did not inhibit type II Thoeris defense when tested experimentally (Figure S5).

We ran our computational pipeline to predict inhibitors of type II Thoeris, attempting to co-fold each of the 38,700 phage proteins together with either ThsA or ThsB of this system. This analysis retrieved three proteins predicted to bind ThsA and one predicted to bind ThsB (Table S9). Of these, two were verified *in vivo* as bona fide inhibitors of type II Thoeris defense (Figure 4B), one of them predicted to bind ThsB (Figures 4C and 4D) and the other one predicted to bind ThsA (Figures 4E–4G). We named these proteins Tad7 and Tad8, respectively. Examining the predicted structural interactions between the verified inhibitors and the type II Thoeris proteins, we found that similar to inhibitors of type I Thoeris, Tad7 binds the ThsB protein of type II Thoeris and inserts a loop into the active site pocket (Figures 4C and 4D). Tad8 is predicted to form a homodimer and insert loops into the macro domain of ThsA. These loops occupy the same space that would be otherwise occupied by the His-ADPR immune signaling molecule (Figures 4E–4G).

To further test the pipeline on a completely unrelated defense system, we next set out to detect phage proteins that inhibit the bacterial CBASS defense system.[41,42] For this, we applied the pipeline on the cGAS/DncV-like nucleotidyltransferase (CD-NTase) protein from the type III CBASS of *Escherichia coli* KTE188.[11] The computational pipeline retrieved only one phage protein predicted to interact with the bacterial CD-NTase (Table S10). When co-expressed with the *E. coli* KTE188 CBASS, the phage protein inhibited its ability to defend against phages, verifying the pipeline prediction (Figure 4H). We named this CBASS inhibitor anti-CBASS 3 (Acb3).

Analysis of the predicted structure shows that Acb3 wraps around the CD-NTase and makes extensive contacts with both the nucleotidyltransferase active site and the putative ligand

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

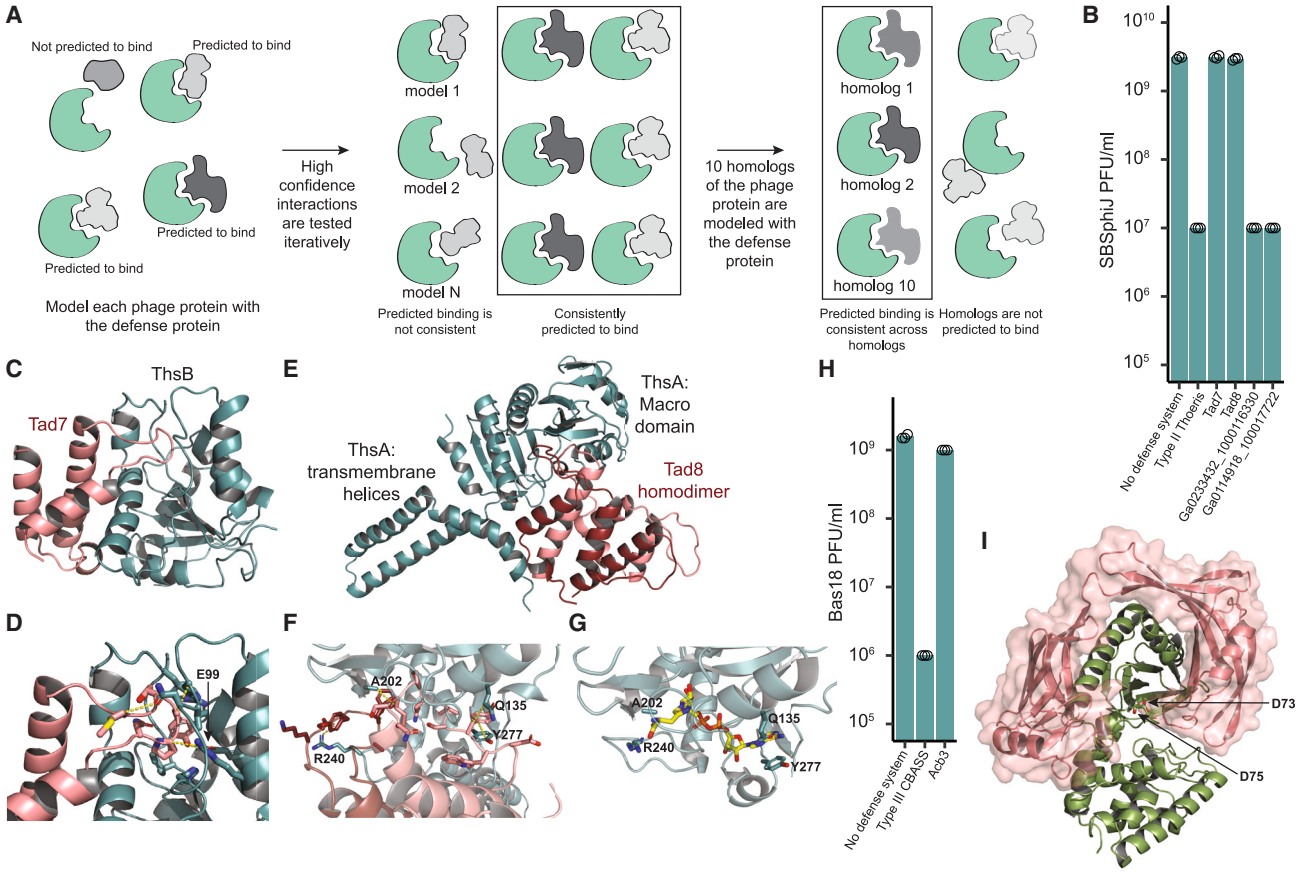


Figure 4. An improved computational pipeline predicts phage-encoded inhibitors of bacterial immune proteins

(A) The modified computational pipeline takes into account co-folding of protein homologs.

(B) Anti-defense activity of the anti-Thoeris type II candidates. Data represent PFUs/mL of phage SBSphiJ infecting control cells (no defense system), cells expressing the type II Thoeris system, and cells co-expressing type II Thoeris and each of the anti-Thoeris candidates.

(C) Structure of the AlphaFold2-Multimer predicted complex formed between Tad7 (red) and ThsB (dark green).

(D) Zoom in on the Tad7-ThsB complex, showing a detailed view of how the active site of ThsB is blocked by a loop of Tad7. E99 is the catalytic residue of the type II ThsB.

(E) Structure of the AlphaFold2-Multimer predicted complex formed by a Tad8 homodimer (shades of red) binding to ThsA (dark green).

(F) Detailed view of the macro domain pocket of ThsA blocked by Tad8. Residues in the macro domain that also interact with His-ADPR are indicated.

(G) Detailed view of the macro domain pocket in complex with the His-ADPR immune signal,[40] showing that His-ADPR occupies the same pocket that is blocked by Tad8.

(H) Acb3 inhibits type III CBASS. Data represent PFUs/mL of phage Bas18 infecting control cells (no defense system), cells expressing the type III CBASS system from *E. coli* KTE188, and cells co-expressing type III CBASS and Acb3.

(I) Structure of the AlphaFold2-Multimer predicted complex formed between Acb3 (red) and the CD-NTase enzyme of the *E. coli* KTE188 CBASS (green). D73 and D75 are two catalytic residues required for 2′3′-cyclic GMP-AMP (cGAMP) production.[41]

Bar graph data shown in (B) and (H) are the average of three replicates, with individual data points overlaid.

See also Figures S4–S7.

binding surface that is required for CD-NTase activation (Figure 4I). Despite extensive efforts, we were unable to purify Acb3 to perform biochemical assays. This is possibly due to a large solvent-exposed hydrophobic patch on Acb3 that likely decreases protein solubility and may further disrupt CD-NTase function (Figure S6A).

AlphaFold2-Multimer analysis showed that Acb3 is predicted to bind multiple CD-NTases from previously studied CBASS systems, despite substantial divergence in sequence between these CD-NTases (Figure S6B). To test whether Acb3 is indeed a broad-range inhibitor of CBASS, we co-expressed it in *Bacillus subtilis* together with the type I CBASS system from *B. cereus* VD146. Despite the low sequence similarity (21.7%) between the *B. cereus* VD146 CD-NTase enzyme and the *E. coli* KTE188 one, Acb3 was capable of inhibiting both systems (Figures 4H and S7). These results demonstrate that Acb3 can inhibit diverse CBASS systems.

Distribution of anti-Thoeris and anti-CBASS proteins in phage genomes

Sequence- and structure-based homology searches revealed over 7,000 homologs of the anti-defense proteins discovered

in this study in the IMG/VR v4 database of viral proteins.[30] These proteins were derived from phages predicted to infect over 40 families of bacteria from a diverse set of taxonomic phyla, including *Firmicutes*, *Bacteroidota*, *Proteobacteria*, and *Actinobacteriota* (Tables S11, S12, S13, S14, S15, S16, and S17). Tad3 was the most abundant anti-Thoeris protein family in this set, represented in more than 5,000 homologs; 600 homologs of Tad4 were found in this set; and the least abundant were the Tad5 and Tad7 protein families occurring in 96 and 92 homologs, respectively. About 1,000 homologs of Acb3 were detected in this set (Tables S11, S12, S13, S14, S15, S16, and S17).

We additionally examined the Metagenomic Gut Virus (MGV) database,[32] an independent database containing ~190,000 genome scaffolds representing sequenced and partially sequenced phage genomes. The anti-Thoeris proteins detected in this study were found in 4,356 of these genome scaffolds (2.3% of the scaffolds), with Tad3 being the most abundant anti-Thoeris protein also in this set, occurring in 3,627 scaffolds (Table S18). Acb3 was underrepresented in this set of phage genomes, detected in only 8 scaffolds (Table S18).

Tad3 and Tad4 are capable of binding diverse ThsB enzymes

The presence of Tad3 and Tad4 homologs in thousands of phage scaffolds (Tables S11, S12, and S18) suggests that these inhibitors are important for immune evasion in phages. To further assess the ability of these proteins to inhibit ThsB TIR-domain proteins, we collected 1,225 ThsB proteins from predicted type I Thoeris systems (Table S19). Phylogenetic analysis of these ThsB proteins revealed two major clades (Figure S8A). We then used AlphaFold2-Multimer to predict possible complex formation between each of the ThsB proteins on the tree and Tad3 or Tad4 and found that while Tad4 was predicted to bind ThsB proteins from all across the tree, Tad3 was limited to binding ThsB proteins from only one of the two clades (Figure S8A).

To test these predictions, we selected five ThsB homologs representing the phylogenetic diversity of ThsB proteins (in addition to the ThsB protein from *Bacillus cereus* MSX-D12 used as a positive control). These six ThsB proteins span substantial sequence diversity and share 25%–50% sequence identity (Figure S8B). We cloned each of these ThsB proteins with an N-terminal 6×His-SUMO2 tag and then co-expressed each of the tagged ThsBs with Tad3 or Tad4 and assessed interactions using co-purification assays. Consistent with our predictions, all six ThsB proteins co-purified with Tad4, whereas Tad3 co-purified with ThsB proteins from one clade only (Figures S8C and S8D). These results demonstrate that Tad4, and to some extent also Tad3, are broad Thoeris inhibitors and can bind and antagonize diverse ThsB proteins belonging to type I Thoeris systems.

Phage-encoded anti-defense proteins bind and antagonize eukaryotic immune proteins

Recent studies show that central components of the eukaryotic cell-autonomous innate immune system originated from bacterial defense systems.[6–11] Specifically, ample evidence suggests that the human cGAS (hcGAS)-STING pathway originated from the bacterial CBASS system, and it was shown that the hcGAS is structurally similar to its bacterial counterpart.[6,43] It was also shown that bacterial immune TIR domains are structurally and functionally similar to plant and human TIR domains that are involved in immunity and regulated cell death.[44,45] The structural and functional similarities between the bacterial immune proteins and related homologs in eukaryotes led us to hypothesize that some of the phage-encoded inhibitors we discovered might also inhibit human and plant TIR and cGAS proteins.

To test this hypothesis, we modeled interactions between eukaryotic immune proteins and the phage-encoded inhibitors to identify high-scoring predicted interactions (Figures S9A–S9D). First, we focused on the TIR-domain protein from the plant *Brachypodium distachyon* (BdTIR), which is known to produce 1″-2′ gcADPR (also called 2'cADPR) and 1″-3′ gcADPR molecules.[46] We modeled the interaction between each of the anti-TIR phage proteins we discovered (Tad3-Tad7) and the BdTIR protein using AlphaFold2-Multimer and found high-scoring predicted interactions between Tad4 and BdTIR, suggesting that Tad4 may bind BdTIR (Figure S9A). To test the predicted binding, we co-expressed Tad4 in cells also expressing a 6×His-SUMO2-tagged BdTIR. Co-purification experiments and MS analyses demonstrated that Tad4 co-precipitates with BdTIR (Figure 5A; Tables S20 and S21). Furthermore, purification via nickel-nitrilotriacetic acid (Ni-NTA) affinity and SEC demonstrated the formation of a stable complex between Tad4 and BdTIR, and analysis of the purified complex via SEC-MALS confirmed that BdTIR forms a 1:1 complex with Tad4 (Figures S10A–S10C) as expected from the AlphaFold2-Multimer model (Figure S9A). These results demonstrate that Tad4 forms a stable complex with BdTIR.

BdTIR is known to consume NAD^+ and constitutively produce 1″-2′ gcADPR and 1″-3′ gcADPR molecules when expressed in *E. coli*,[46] and indeed, filtered cell lysates derived from cells expressing BdTIR activated ThsA from type I Thoeris, a protein triggered by 1″-3′ gcADPR (Figure 5B). However, filtered lysates derived from cells in which BdTIR was co-expressed with Tad4 failed to activate ThsA *in vitro*, demonstrating that Tad4 inhibits the TIR-mediated enzymatic activity of BdTIR (Figure 5B). In further support of this observation, co-expression of BdTIR and Tad4 stabilized the expression of BdTIR, likely by inhibiting the toxic effects associated with BdTIR NAD^+ consumption (Figure 5A).

We next considered the human Sterile alpha and TIR motif-containing 1 protein (SARM1). SARM1 is an essential protein within a pathway that leads to axonal death in response to neuronal injury, and it was shown that the SARM1 TIR domain degrades cellular NAD^+ when activated following neuron insult.[47,48] Co-folding the TIR domain of human SARM1 ($SARM1_{TIR}$) with either Tad3, Tad4, Tad5, Tad6, and Tad7 using AlphaFold2-Multimer strongly predicted that Tad4, and to a lesser extent also Tad3, bind $SARM1_{TIR}$ (Figures S9B and S9C). We were not able to verify the predicted interactions between Tad3 and $SARM1_{TIR}$. However, co-expression of Tad4 with a 6×His-SUMO2-tagged $SARM1_{TIR}$ demonstrated that Tad4 co-purifies with $SARM1_{TIR}$ (Figure 5C), and analysis of the purified complex via MS and SEC-MALS verified the formation of a 1:1 Tad4:$SARM1_{TIR}$ complex (Figures S10A–S10C; Tables S22 and S23).

It was previously shown that $SARM1_{TIR}$ is constitutively active as an NADase when expressed in bacterial cells, depleting NAD^+ in these cells.[49] To test whether the binding of Tad4 to $SARM1_{TIR}$

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

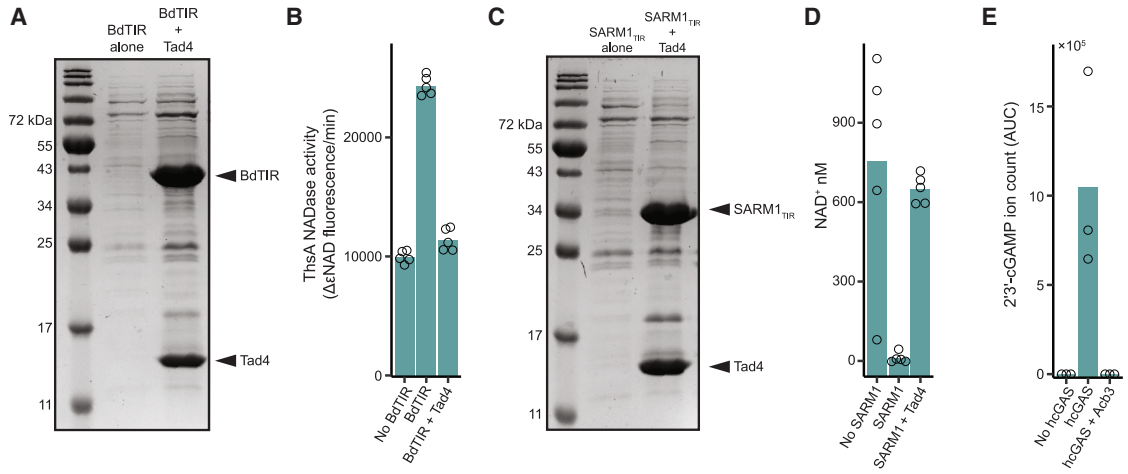


Figure 5. Phage-derived anti-defense proteins bind and inhibit human and plant immune proteins

(A) Pull-down of a 6×His-SUMO2-tagged BdTIR co-expressed with Tad4 demonstrates that these proteins co-elute together. Shown is an SDS-PAGE of proteins following pull-down.

(B) NADase activity of purified ThsA incubated with filtered lysates derived from cells expressing BdTIR alone, BdTIR together with Tad4, or control cells that do not express BdTIR. NADase activity was measured using a nicotinamide εNAD cleavage fluorescence assay. Bars represent the mean of five experiments, with individual data points overlaid.

(C) Pull-down of a His-tagged TIR domain of the human SARM1 protein ($SARM1_{TIR}$) that was co-expressed with Tad4 demonstrates that these proteins co-elute together. Shown is an SDS-PAGE of proteins following pull-down.

(D) NAD^+ levels in filtered lysates derived from cells expressing $hSARM1_{TIR}$, cells co-expressing $hSARM1_{TIR}$ and Tad4, or control cells that do not express $hSARM1_{TIR}$. Bars represent the mean of five experiments, with individual data points overlaid.

(E) Co-expression of Acb3 with hcGAS prevents 2′3′-cGAMP production. Lysates were analyzed by LC-MS. The y axis represents the area under the curve (AUC) of 2′3′-cGAMP ions detected in MS analysis.

See also Figures S9 and S10.

inhibits its enzymatic function, we measured the levels of NAD^+ in lysates derived from cells in which $SARM1_{TIR}$ was co-expressed with Tad4. We found that Tad4 inhibited $SARM1_{TIR}$ activity almost completely (Figure 5D). Taken together, these results demonstrate that the phage-derived anti-Thoeris protein Tad4 is capable of binding and inhibiting the human SARM1 TIR domain.

Based on the AlphaFold2-Multimer models, Tad4 interacts with the TIR domains of ThsB, BdTIR, and SARM1 in a nearly identical manner (Figure S10D). Analyses of the modeled structures suggest that this is achieved via an extensive network of molecular interactions between Tad4 and each of the TIRs, anchoring multiple conserved Tad4 side chains to residues that are broadly conserved between diverse TIR proteins and forming interactions both within and outside the TIR active site pocket (Figure S10D). In addition to the flexible loop of Tad4 that blocks the active site of the TIR domains, the C terminus of Tad4 forms a beta strand that interlaces with a conserved beta strand of each of the TIR domains, further stabilizing these interactions (Figure S10D).

Finally, we examined possible interactions between Acb3 and the hcGAS. AlphaFold2-Multimer analysis predicted interactions between Acb3 and hcGAS (Figure S9D), and we hence co-expressed these two proteins in bacterial cells. When expressed alone in bacteria, hcGAS constitutively produced 2′3′-cGAMP, as previously shown[50] (Figure 5E). By contrast, we could not detect 2′3′-cGAMP in lysates derived from cells in which hcGAS was co-expressed with Acb3 (Figure 5E). The AlphaFold2-Multimer models suggest that Acb3 binds both bacterial and hcGAS-like proteins in a surprisingly similar manner, forming a series of molecular interactions with residues that are conserved between the bacterial and the human enzymes, both within the nucleotidyltransferase active site pocket and in the putative ligand binding site (Figure S10E). Together, these results demonstrate that conservation of immune proteins between prokaryotes and eukaryotes renders the eukaryotic homologs susceptible to inhibition by viral proteins that primarily evolved to inhibit bacterial immunity.

DISCUSSION

The vast majority of the viral protein universe comprises proteins of unknown function,[30–32] and it was predicted that the purpose of many of these proteins is to inhibit host immunity.[1,13,36] In this study we demonstrated that AlphaFold2-Multimer co-folding analyses can sift through vast sets of viral proteins and identify those that bind and inhibit immune factors. While AlphaFold2-Multimer notoriously yields false-positive predictions,[51] we show here that false-positive rates can be substantially reduced by co-folding sets of protein homologs with the target immune protein. We demonstrate the utility of our discovery pipeline by detecting phage-derived inhibitors for the Thoeris and CBASS defense systems. We predict that future analyses using our approach will discover numerous families of phage proteins that inhibit the >150 defense systems known to be encoded by bacteria. Future analyses on viruses that infect humans, animals,

and plants are similarly predicted to discover viral proteins that inhibit eukaryotic immunity.

Intriguingly, in all but one of the phage proteins we detected, inhibition of host immunity is predicted to involve a flexible loop that enters the active site pocket of the immune protein and directly interacts with active site residues. Tad3, Tad4, and Tad6 each have a conserved loop that blocks the active site pocket and interacts with the catalytic E85 residue of the type I ThsB; the conserved loop of Tad7 binds the catalytic E99 residue of type II ThsB via a hydrogen bond; conserved loops in Tad8 occupy the pocket within ThsA responsible for His-ADPR binding; and Acb3 encodes a conserved loop that is predicted to block the active site of CD-NTases and interact with the nucleotidyltransferase catalytic residues via hydrogen and ionic bonds. These observations suggest that blocking access to active sites that produce or recognize immune signaling molecules is a common viral strategy for inhibition of defense systems that rely on small molecule signaling.

We find it remarkable that phage proteins that likely evolved to inhibit bacterial immune proteins also bind and inhibit the distant immune protein homologs in animals and plants. Presumably, these proteins were selected throughout billions of years of evolution to recognize and bind conserved structural features of the respective immune protein, features that remained conserved across domains of life. Since bacterial pathogens often benefit from virulence factors carried by prophages integrated in their genomes,[52,53] we predict that future studies will find cases in which TIR and cGAS inhibitors carried by prophages were adopted by bacterial pathogens of plants and animals to shut down host immunity. From a practical standpoint, our findings suggest that phage-derived inhibitors can be adopted in research and clinical applications as immune-modulatory factors to inhibit immune processes in animals and plants.

Limitations of the study

While the AlphaFold2-Multimer pipeline presented in this work successfully identified multiple viral proteins that bind and inhibit immune proteins, there are limitations to consider for future research. The reliance of AlphaFold2-Multimer on data from homologous proteins[37,51] restricts its ability to identify inhibitors that have only few homologs in sequence databases. Furthermore, our study focused on phage-derived inhibitors that bind immune proteins through a large surface area. Studies in the CRISPR-Cas field demonstrate that anti-CRISPR proteins may bind and inhibit Cas proteins using a relatively small surface area.[54–56] This highlights that a large interaction surface area may not always be crucial for binding immune proteins. The AlphaFold2-Multimer pipeline described in this study may therefore not be the optimal method for detecting immune inhibitors that bind defense proteins using a small surface area.

RESOURCE AVAILABILITY

Lead contact

Further information and requests for resources and reagents should be directed to the lead contact, Rotem Sorek (rotem.sorek@weizmann.ac.il).

Materials availability

This study did not generate new unique reagents.

Data and code availability

Data that support the findings of this study are available in the article and Tables S1, S2, S3, S4, S5, S6, S7, S8, S9, S10, S11, S12, S13, S14, S15, S16, S17, S18, S19, S20, S21, S22, S23, S24, and S25. Coordinates and structure factors for the structure of the ThsB-Tad3 complex have been deposited in the Protein Data Bank (PDB) under the accession code PDB: 9B7D. Sequences of Tad3, Tad4, Tad5, Tad6, Tad8, and Acb3 proteins were deposited as new protein families in the Protein families (Pfam) database under the accession codes Pfam: PF25185, PF25186, PF25190, PF25188, PF25189, and PF25187, respectively. Any additional information required to reanalyze the data reported in this paper is available from the lead contact upon reasonable request.

ACKNOWLEDGMENTS

We thank the Sorek laboratory members for comments on the manuscript and fruitful discussion. We thank A. Savidor and M. Kupervaser at the De Botton Protein Profiling Institute at the Weizmann Institute for help with mass spectrometry data generation and analysis. R.S. was supported, in part, by the European Research Council (grant ERC-AdG GA 101018520), the Israel Science Foundation (MAPATS grant 2720/22), the Deutsche Forschungsgcemeinschaft (SPP 2330 grant 464312965), the Ernest and Bonnie Beutler Research Program of Excellence in Genomic Medicine, a research grant from the Estate of Marjorie Plesset, and the Knell Family Center for Microbiology. E.Y. is supported by the Clore Scholars Program and, in part, by the Israeli Council for Higher Education (CHE) via the Weizmann Data Science Research Center. P.J.K. was supported, in part, by the Pew Biomedical Scholars program, The Mathers Foundation, and The Mark Foundation for Cancer Research. S.J.H. is supported through a Cancer Research Institute Irvington Postdoctoral Fellowship (CRI3996). X-ray data were collected at the Northeastern Collaborative Access Team beamlines 24-ID-C and 24-ID-E (P30 GM124165) using a Pilatus detector (S10RR029205), an Eiger detector (S10OD021527), and the Argonne National Laboratory Advanced Photon Source (DE-AC02-06CH11357).

AUTHOR CONTRIBUTIONS

E.Y. and R.S. conceptualized and led the study. E.Y. built and executed the computational pipeline and analyzed all the data. Structural and *in vitro* biochemical experiments were performed by S.J.H. with the help of J.M.J.T. and H.C.T. under the supervision of P.J.K. A.L. conducted all the phage infection experiments with the help of C.A., B.M., and M.S. I.O. designed and performed knockin experiments. D.H. helped with the identification of protein homologs in metagenomic databases. I.C. helped with setting up an infrastructure for computational analysis. M.I. and S.M. performed the LC-MS metabolite analysis. G.A. conducted the *in vivo* biochemical experiments with the help of A.L. The manuscript was written by E.Y. and R.S. The study was supervised by R.S. All authors contributed to editing the manuscript and support the conclusions.

DECLARATION OF INTERESTS

R.S. is a scientific co-founder of and adviser for BiomX and Ecophage. A patent application titled "a method for discovery of proteins that inhibit immunity" was submitted in parallel with the manuscript.

STAR★METHODS

Detailed methods are provided in the online version of this paper and include the following:

- KEY RESOURCES TABLE
- EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS
  - Bacterial strains
  - Phage strains
- METHOD DETAILS
  - A database of short phage proteins of unknown function
  - Prediction of phage-encoded proteins that inhibit type I Thoeris
  - Cloning of candidate anti-defense genes

- Plaque assays
- Protein co-expression for biochemical assessment of metabolites
- Preparation of filtered cell lysates
- Enzymatic assays
- Protein expression and purification
- Co-expression of ThsB homologs with Tad3 and Tad4
- SDS-PAGE analysis
- Protein mass spectrometry analysis
- SEC-MALS analysis
- Crystallization and structure determination
- Knockin of Tad3 into phage SBSphiJ
- Selection of homologs of anti-defense candidates for binding analysis
- Prediction of phage encoded proteins that inhibit type II Thoeris or CBASS
- Distribution of anti-defense proteins in phage genomes
- Identification of ThsB homologs and phylogenetic reconstruction
- LC-MS polar metabolite analysis

- QUANTIFICATION AND STATISTICAL ANALYSIS

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.cell.2024.12.035.

Received: April 5, 2024
Revised: October 29, 2024
Accepted: December 24, 2024
Published: January 23, 2025

REFERENCES

1. Hampton, H.G., Watson, B.N.J., and Fineran, P.C. (2020). The arms race between bacteria and their phage foes. Nature *577*, 327–336. https://doi.org/10.1038/s41586-019-1894-8.
2. Doron, S., Melamed, S., Ofir, G., Leavitt, A., Lopatina, A., Keren, M., Amitai, G., and Sorek, R. (2018). Systematic discovery of antiphage defense systems in the microbial pangenome. Science *359*, eaar4120. https://doi.org/10.1126/science.aar4120.
3. Gao, L., Altae-Tran, H., Böhning, F., Makarova, K.S., Segel, M., Schmid-Burgk, J.L., Koob, J., Wolf, Y.I., Koonin, E.V., and Zhang, F. (2020). Diverse Enzymatic Activities Mediate Antiviral Immunity in Prokaryotes. Science *369*, 1077–1084. https://doi.org/10.1126/science.aba0372.
4. Millman, A., Melamed, S., Leavitt, A., Doron, S., Bernheim, A., Hör, J., Garb, J., Bechon, N., Brandis, A., Lopatina, A., et al. (2022). An expanded arsenal of immune systems that protect bacteria from phages. Cell Host Microbe *30*, 1556–1569.e5. https://doi.org/10.1016/j.chom.2022.09.017.
5. Tal, N., and Sorek, R. (2022). SnapShot: Bacterial immunity. Cell *185*, 578–578.e1. https://doi.org/10.1016/j.cell.2021.12.029.
6. Morehouse, B.R., Govande, A.A., Millman, A., Keszei, A.F.A., Lowey, B., Ofir, G., Shao, S., Sorek, R., and Kranzusch, P.J. (2020). STING cyclic dinucleotide sensing originated in bacteria. Nature *586*, 429–433. https://doi.org/10.1038/s41586-020-2719-5.
7. Ofir, G., Herbst, E., Baroz, M., Cohen, D., Millman, A., Doron, S., Tal, N., Malheiro, D.B.A., Malitsky, S., Amitai, G., et al. (2021). Antiviral activity of bacterial TIR domains via immune signalling molecules. Nature *600*, 116–120. https://doi.org/10.1038/s41586-021-04098-7.
8. Bernheim, A., Millman, A., Ofir, G., Meitav, G., Avraham, C., Shomar, H., Rosenberg, M.M., Tal, N., Melamed, S., Amitai, G., et al. (2021). Prokaryotic viperins produce diverse antiviral molecules. Nature *589*, 120–124. https://doi.org/10.1038/s41586-020-2762-2.
9. Johnson, A.G., Wein, T., Mayer, M.L., Duncan-Lowey, B., Yirmiya, E., Oppenheimer-Shaanan, Y., Amitai, G., Sorek, R., and Kranzusch, P.J. (2022). Bacterial gasdermins reveal an ancient mechanism of cell death. Science *375*, 221–225. https://doi.org/10.1126/science.abj8432.
10. Wein, T., and Sorek, R. (2022). Bacterial origins of human cell-autonomous innate immune mechanisms. Nat. Rev. Immunol. *22*, 629–638. https://doi.org/10.1038/s41577-022-00705-4.
11. Rousset, F., Yirmiya, E., Nesher, S., Brandis, A., Mehlman, T., Itkin, M., Malitsky, S., Millman, A., Melamed, S., and Sorek, R. (2023). A conserved family of immune effectors cleaves cellular ATP upon viral infection. Cell *186*, 3619–3631.e13. https://doi.org/10.1016/j.cell.2023.07.020.
12. Yirmiya, E., Leavitt, A., Lu, A., Ragucci, A.E., Avraham, C., Osterman, I., Garb, J., Antine, S.P., Mooney, S.E., Hobbs, S.J., et al. (2024). Phages overcome bacterial immunity via diverse anti-defence proteins. Nature *625*, 352–359. https://doi.org/10.1038/s41586-023-06869-w.
13. Gao, Z., and Feng, Y. (2023). Bacteriophage strategies for overcoming host antiviral immunity. Front. Microbiol. *14*, 1211793. https://doi.org/10.3389/fmicb.2023.1211793.
14. Mayo-Muñoz, D., Pinilla-Redondo, R., Camara-Wilpert, S., Birkholz, N., and Fineran, P.C. (2024). Inhibitors of bacterial immune systems: discovery, mechanisms and applications. Nat. Rev. Genet. *25*, 237–254. https://doi.org/10.1038/s41576-023-00676-9.
15. Atanasiu, C., Su, T.J., Sturrock, S.S., and Dryden, D.T.F. (2002). Interaction of the ocr gene 0.3 protein of bacteriophage T7 with EcoKI restriction/modification enzyme. Nucleic Acids Res. *30*, 3936–3944. https://doi.org/10.1093/nar/gkf518.
16. Otsuka, Y., and Yonesaki, T. (2012). Dmd of bacteriophage T4 functions as an antitoxin against *Escherichia coli* LsoA and RnlA toxins. Mol. Microbiol. *83*, 669–681. https://doi.org/10.1111/j.1365-2958.2012.07975.x.
17. Bondy-Denomy, J., Garcia, B., Strum, S., Du, M., Rollins, M.F., Hidalgo-Reyes, Y., Wiedenheft, B., Maxwell, K.L., and Davidson, A.R. (2015). Multiple mechanisms for CRISPR–Cas inhibition by anti-CRISPR proteins. Nature *526*, 136–139. https://doi.org/10.1038/nature15254.
18. Davidson, A.R., Lu, W.-T., Stanley, S.Y., Wang, J., Mejdani, M., Trost, C.N., Hicks, B.T., Lee, J., and Sontheimer, E.J. (2020). Anti-CRISPRs: Protein Inhibitors of CRISPR-Cas Systems. Annu. Rev. Biochem. *89*, 309–332. https://doi.org/10.1146/annurev-biochem-011420-111224.
19. Garb, J., Lopatina, A., Bernheim, A., Zaremba, M., Siksnys, V., Melamed, S., Leavitt, A., Millman, A., Amitai, G., and Sorek, R. (2022). Multiple phage resistance systems inhibit infection via SIR2-dependent NAD^+ depletion. Nat. Microbiol. *7*, 1849–1856. https://doi.org/10.1038/s41564-022-01207-8.
20. Antine, S.P., Johnson, A.G., Mooney, S.E., Leavitt, A., Mayer, M.L., Yirmiya, E., Amitai, G., Sorek, R., and Kranzusch, P.J. (2024). Structural basis of Gabija anti-phage defence and viral immune evasion. Nature *625*, 360–365. https://doi.org/10.1038/s41586-023-06855-2.
21. Hobbs, S.J., Wein, T., Lu, A., Morehouse, B.R., Schnabel, J., Leavitt, A., Yirmiya, E., Sorek, R., and Kranzusch, P.J. (2022). Phage anti-CBASS and anti-Pycsar nucleases subvert bacterial immunity. Nature *605*, 522–526. https://doi.org/10.1038/s41586-022-04716-y.
22. Leavitt, A., Yirmiya, E., Amitai, G., Lu, A., Garb, J., Herbst, E., Morehouse, B.R., Hobbs, S.J., Antine, S.P., Sun, Z.J., et al. (2022). Viruses inhibit TIR gcADPR signalling to overcome bacterial defence. Nature *611*, 326–331. https://doi.org/10.1038/s41586-022-05375-9.
23. Huiting, E., Cao, X., Ren, J., Athukoralage, J.S., Luo, Z., Silas, S., An, N., Carion, H., Zhou, Y., Fraser, J.S., et al. (2023). Bacteriophages inhibit and evade cGAS-like immune function in bacteria. Cell *186*, 864–876.e21. https://doi.org/10.1016/j.cell.2022.12.041.
24. Li, D., Xiao, Y., Fedorova, I., Xiong, W., Wang, Y., Liu, X., Huiting, E., Ren, J., Gao, Z., Zhao, X., et al. (2024). Single phage proteins sequester signals from TIR and cGAS-like enzymes. Nature *635*, 719–727. https://doi.org/10.1038/s41586-024-08122-4.
25. Cao, X., Xiao, Y., Huiting, E., Cao, X., Li, D., Ren, J., Fedorova, I., Wang, H., Guan, L., Wang, Y., et al. (2024). Phage anti-CBASS protein simultaneously sequesters cyclic trinucleotides and dinucleotides. Mol. Cell *84*, 375–385.e7. https://doi.org/10.1016/j.molcel.2023.11.026.

26. Jenson, J.M., Li, T., Du, F., Ea, C.-K., and Chen, Z.J. (2023). Ubiquitin-like conjugation by bacterial cGAS enhances anti-phage defence. Nature *616*, 326–331. https://doi.org/10.1038/s41586-023-05862-7.

27. Alawneh, A.M., Qi, D., Yonesaki, T., and Otsuka, Y. (2016). An ADP-ribosyltransferase Alt of bacteriophage T4 negatively regulates the *Escherichia coli* MazF toxin of a toxin-antitoxin module. Mol. Microbiol. *99*, 188–198. https://doi.org/10.1111/mmi.13225.

28. Dong, L., Guan, X., Li, N., Zhang, F., Zhu, Y., Ren, K., Yu, L., Zhou, F., Han, Z., Gao, N., et al. (2019). An anti-CRISPR protein disables type V Cas12a by acetylation. Nat. Struct. Mol. Biol. *26*, 308–314. https://doi.org/10.1038/s41594-019-0206-1.

29. Niu, Y., Yang, L., Gao, T., Dong, C., Zhang, B., Yin, P., Hopp, A.-K., Li, D., Gan, R., Wang, H., et al. (2020). A Type I-F Anti-CRISPR Protein Inhibits the CRISPR-Cas Surveillance Complex by ADP-Ribosylation. Mol. Cell *80*, 512–524.e5. https://doi.org/10.1016/j.molcel.2020.09.015.

30. Camargo, A.P., Nayfach, S., Chen, I.A., Palaniappan, K., Ratner, A., Chu, K., Ritter, S.J., Reddy, T.B.K., Mukherjee, S., Schulz, F., et al. (2023). IMG/VR v4: an expanded database of uncultivated virus genomes within a framework of extensive functional, taxonomic, and ecological metadata. Nucleic Acids Res. *51*, D733–D743. https://doi.org/10.1093/nar/gkac1037.

31. Cook, R., Brown, N., Redgwell, T., Rihtman, B., Barnes, M., Clokie, M., Stekel, D.J., Hobman, J., Jones, M.A., and Millard, A. (2021). INfrastructure for a PHAge REference Database: Identification of Large-Scale Biases in the Current Collection of Cultured Phage Genomes. Phage (New Rochelle) *2*, 214–223. https://doi.org/10.1089/phage.2021.0007.

32. Nayfach, S., Páez-Espino, D., Call, L., Low, S.J., Sberro, H., Ivanova, N.N., Proal, A.D., Fischbach, M.A., Bhatt, A.S., Hugenholtz, P., et al. (2021). Metagenomic compendium of 189,680 DNA viruses from the human gut microbiome. Nat. Microbiol. *6*, 960–970. https://doi.org/10.1038/s41564-021-00928-6.

33. Miller, E.S., Kutter, E., Mosig, G., Arisaka, F., Kunisawa, T., and Rüger, W. (2003). Bacteriophage T4 Genome. Microbiol. Mol. Biol. Rev. *67*, 86–156. https://doi.org/10.1128/MMBR.67.1.86-156.2003.

34. Wang, J., Jiang, Y., Vincent, M., Sun, Y., Yu, H., Wang, J., Bao, Q., Kong, H., and Hu, S. (2005). Complete genome sequence of bacteriophage T5. Virology *332*, 45–65. https://doi.org/10.1016/j.virol.2004.10.049.

35. Molineux, I.J. (2005). The T7 Group. In The Bacteriophages, Second Edition, R. Calendar and S.T. Abedon, eds. (Oxford University Press), pp. 277–301. https://doi.org/10.1093/oso/9780195148503.003.0020.

36. Ofir, G., and Sorek, R. (2018). Contemporary Phage Biology: From Classic Models to New Insights. Cell *172*, 1260–1270. https://doi.org/10.1016/j.cell.2017.10.045.

37. Evans, R., O'Neill, M., Pritzel, A., Antropova, N., Senior, A., Green, T., Žídek, A., Bates, R., Blackwell, S., Yim, J., et al. (2022). Protein complex prediction with AlphaFold-Multimer. Preprint at bioRxiv. https://doi.org/10.1101/2021.10.04.463034.

38. Roux, S., Páez-Espino, D., Chen, I.A., Palaniappan, K., Ratner, A., Chu, K., Reddy, T.B.K., Nayfach, S., Schulz, F., Call, L., et al. (2021). IMG/VR v3: an integrated ecological and evolutionary framework for interrogating genomes of uncultivated viruses. Nucleic Acids Res. *49*, D764–D775. https://doi.org/10.1093/nar/gkaa946.

39. Tamulaitiene, G., Sabonis, D., Sasnauskas, G., Ruksenaite, A., Silanskas, A., Avraham, C., Ofir, G., Sorek, R., Zaremba, M., and Siksnys, V. (2024). Activation of Thoeris antiviral system via SIR2 effector filament assembly. Nature *627*, 431–436. https://doi.org/10.1038/s41586-024-07092-x.

40. Sabonis, D., Avraham, C., Lu, A., Herbst, E., Silanskas, A., Leavitt, A., Yirmiya, E., Zaremba, M., Amitai, G., Kranzusch, P.J., et al. (2024). TIR domains produce histidine-ADPR conjugates as immune signaling molecules in bacteria. Preprint at bioRxiv. https://doi.org/10.1101/2024.01.03.573942.

41. Cohen, D., Melamed, S., Millman, A., Shulman, G., Oppenheimer-Shaanan, Y., Kacen, A., Doron, S., Amitai, G., and Sorek, R. (2019). Cyclic GMP–AMP signalling protects bacteria against viral infection. Nature *574*, 691–695. https://doi.org/10.1038/s41586-019-1605-5.

42. Millman, A., Melamed, S., Amitai, G., and Sorek, R. (2020). Diversity and classification of cyclic-oligonucleotide-based anti-phage signalling systems. Nat. Microbiol. *5*, 1608–1615. https://doi.org/10.1038/s41564-020-0777-y.

43. Whiteley, A.T., Eaglesham, J.B., de Oliveira Mann, C.C., Morehouse, B.R., Lowey, B., Nieminen, E.A., Danilchanka, O., King, D.S., Lee, A.S.Y., Mekalanos, J.J., et al. (2019). Bacterial cGAS-like enzymes synthesize diverse nucleotide signals. Nature *567*, 194–199. https://doi.org/10.1038/s41586-019-0953-5.

44. Essuman, K., Milbrandt, J., Dangl, J.L., and Nishimura, M.T. (2022). Shared TIR enzymatic functions regulate cell death and immunity across the tree of life. Science *377*, eabo0001. https://doi.org/10.1126/science.abo0001.

45. Locci, F., Wang, J., and Parker, J.E. (2023). TIR-domain enzymatic activities at the heart of plant immunity. Curr. Opin. Plant Biol. *74*, 102373. https://doi.org/10.1016/j.pbi.2023.102373.

46. Bayless, A.M., Chen, S., Ogden, S.C., Xu, X., Sidda, J.D., Manik, M.K., Li, S., Kobe, B., Ve, T., Song, L., et al. (2023). Plant and prokaryotic TIR domains generate distinct cyclic ADPR NADase products. Sci. Adv. *9*, eade8487. https://doi.org/10.1126/sciadv.ade8487.

47. Essuman, K., Summers, D.W., Sasaki, Y., Mao, X., DiAntonio, A., and Milbrandt, J. (2017). The SARM1 Toll/Interleukin-1 Receptor Domain Possesses Intrinsic NAD^+ Cleavage Activity that Promotes Pathological Axonal Degeneration. Neuron *93*, 1334–1343.e5. https://doi.org/10.1016/j.neuron.2017.02.022.

48. Figley, M.D., and DiAntonio, A. (2020). The SARM1 axon degeneration pathway: control of the NAD^+ metabolome regulates axon survival in health and disease. Curr. Opin. Neurobiol. *63*, 59–66. https://doi.org/10.1016/j.conb.2020.02.012.

49. Garb, J., Amitai, G., Lu, A., Ofir, G., Brandis, A., Mehlman, T., Kranzusch, P.J., and Sorek, R. (2024). The SARM1 TIR domain produces glycocyclic ADPR molecules as minor products. PLoS One *19*, e0302251. https://doi.org/10.1371/journal.pone.0302251.

50. Zhou, W., Whiteley, A.T., de Oliveira Mann, C.C., Morehouse, B.R., Nowak, R.P., Fischer, E.S., Gray, N.S., Mekalanos, J.J., and Kranzusch, P.J. (2018). Structure of the Human cGAS–DNA Complex Reveals Enhanced Control of Immune Surveillance. Cell *174*, 300–311.e11. https://doi.org/10.1016/j.cell.2018.06.026.

51. Zhu, W., Shenoy, A., Kundrotas, P., and Elofsson, A. (2023). Evaluation of AlphaFold-Multimer prediction on multi-chain protein complexes. Bioinformatics *39*, btad424. https://doi.org/10.1093/bioinformatics/btad424.

52. Brüssow, H., Canchaya, C., and Hardt, W.-D. (2004). Phages and the Evolution of Bacterial Pathogens: from Genomic Rearrangements to Lysogenic Conversion. Microbiol. Mol. Biol. Rev. *68*, 560–602. https://doi.org/10.1128/mmbr.68.3.560-602.2004.

53. Fortier, L.-C., and Sekulovic, O. (2013). Importance of prophages to evolution and virulence of bacterial pathogens. Virulence *4*, 354–365. https://doi.org/10.4161/viru.24498.

54. Jia, N., and Patel, D.J. (2021). Structure-based functional mechanisms and biotechnology applications of anti-CRISPR proteins. Nat. Rev. Mol. Cell Biol. *22*, 563–579. https://doi.org/10.1038/s41580-021-00371-9.

55. Fuchsbauer, O., Swuec, P., Zimberger, C., Amigues, B., Levesque, S., Agudelo, D., Duringer, A., Chaves-Sanjuan, A., Spinelli, S., Rousseau, G.M., et al. (2019). Cas9 Allosteric Inhibition by the Anti-CRISPR Protein AcrIIA6. Mol. Cell *76*, 922–937.e7. https://doi.org/10.1016/j.molcel.2019.09.012.

56. Deng, X., Sun, W., Li, X., Wang, J., Cheng, Z., Sheng, G., and Wang, Y. (2024). An anti-CRISPR that represses its own transcription while blocking Cas9-target DNA binding. Nat. Commun. *15*, 1806. https://doi.org/10.1038/s41467-024-45987-5.

57. Maffei, E., Shaidullina, A., Burkolter, M., Heyer, Y., Estermann, F., Druelle, V., Sauer, P., Willi, L., Michaelis, S., Hilbi, H., et al. (2021). Systematic exploration of Escherichia coli phage–host interactions with the BASEL phage collection. PLoS Biol. *19*, e3001424. https://doi.org/10.1371/journal.pbio.3001424.

58. Steinegger, M., and Söding, J. (2017). MMseqs2 enables sensitive protein sequence searching for the analysis of massive data sets. Nat. Biotechnol. *35*, 1026–1028. https://doi.org/10.1038/nbt.3988.

59. Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., Tunyasuvunakool, K., Bates, R., Žídek, A., Potapenko, A., et al. (2021). Highly accurate protein structure prediction with AlphaFold. Nature *596*, 583–589. https://doi.org/10.1038/s41586-021-03819-2.

60. Katoh, K., and Standley, D.M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol. Biol. Evol. *30*, 772–780. https://doi.org/10.1093/molbev/mst010.

61. Deatherage, D.E., and Barrick, J.E. (2014). Identification of mutations in laboratory-evolved microbes from next-generation sequencing data using breseq. Methods Mol. Biol. *1151*, 165–188. https://doi.org/10.1007/978-1-4939-0554-6_12.

62. McCoy, A.J., Grosse-Kunstleve, R.W., Adams, P.D., Winn, M.D., Storoni, L.C., and Read, R.J. (2007). Phaser crystallographic software. J. Appl. Crystallogr. *40*, 658–674. https://doi.org/10.1107/S0021889807021206.

63. Emsley, P., and Cowtan, K. (2004). Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. *60*, 2126–2132. https://doi.org/10.1107/S0907444904019158.

64. Clementel, D., Del Conte, A., Monzon, A.M., Camagni, G.F., Minervini, G., Piovesan, D., and Tosatto, S.C.E. (2022). RING 3.0: fast generation of probabilistic residue interaction networks from structural ensembles. Nucleic Acids Res. *50*, W651–W656. https://doi.org/10.1093/nar/gkac365.

65. van Kempen, M., Kim, S.S., Tumescheit, C., Mirdita, M., Lee, J., Gilchrist, C.L.M., Söding, J., and Steinegger, M. (2024). Fast and accurate protein structure search with Foldseek. Nat. Biotechnol. *42*, 243–246. https://doi.org/10.1038/s41587-023-01773-0.

66. Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. *7*, 539. https://doi.org/10.1038/msb.2011.75.

67. Minh, B.Q., Schmidt, H.A., Chernomor, O., Schrempf, D., Woodhams, M.D., von Haeseler, A., and Lanfear, R. (2020). IQ-TREE 2: New Models and Efficient Methods for Phylogenetic Inference in the Genomic Era. Mol. Biol. Evol. *37*, 1530–1534. https://doi.org/10.1093/molbev/msaa015.

68. Letunic, I., and Bork, P. (2021). Interactive Tree Of Life (iTOL) v5: an online tool for phylogenetic tree display and annotation. Nucleic Acids Res. *49*, W293–W296. https://doi.org/10.1093/nar/gkab301.

69. Mitchell, A.L., Almeida, A., Beracochea, M., Boland, M., Burgin, J., Cochrane, G., Crusoe, M.R., Kale, V., Potter, S.C., Richardson, L.J., et al. (2020). MGnify: the microbiome analysis resource in 2020. Nucleic Acids Res. *48*, D570–D578. https://doi.org/10.1093/nar/gkz1035.

70. Waterhouse, A.M., Procter, J.B., Martin, D.M.A., Clamp, M., and Barton, G.J. (2009). Jalview Version 2—a multiple sequence alignment editor and analysis workbench. Bioinformatics *25*, 1189–1191. https://doi.org/10.1093/bioinformatics/btp033.

71. Mazzocco, A., Waddell, T.E., Lingohr, E., and Johnson, R.P. (2009). Enumeration of Bacteriophages Using the Small Drop Plaque Assay System. In Bacteriophages: Methods and Protocols, Volume 1: Isolation, Characterization, and Interactions, M.R.J. Clokie and A.M. Kropinski, eds. (Humana Press), pp. 81–85. https://doi.org/10.1007/978-1-60327-164-6_9.

72. Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., and Richardson, D.C. (2010). MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. *66*, 12–21. https://doi.org/10.1107/S0907444909042073.

73. Karplus, P.A., and Diederichs, K. (2012). Linking crystallographic model and data quality. Science *336*, 1030–1033. https://doi.org/10.1126/science.1218231.

74. Zheng, L., Cardaci, S., Jerby, L., MacKenzie, E.D., Sciacovelli, M., Johnson, T.I., Gaude, E., King, A., Leach, J.D.G., Edrada-Ebel, R., et al. (2015). Fumarate induces redox-dependent senescence by modifying glutathione metabolism. Nat. Commun. *6*, 6001. https://doi.org/10.1038/ncomms7001.

75. Gnainsky, Y., Itkin, M., Mehlman, T., Brandis, A., Malitsky, S., and Soen, Y. (2022). Protocol for studying microbiome impact on host energy and reproduction in *Drosophila*. Star Protoc. *3*, 101253. https://doi.org/10.1016/j.xpro.2022.101253.

76. Chen, I.A., Chu, K., Palaniappan, K., Ratner, A., Huang, J., Huntemann, M., Hajek, P., Ritter, S., Varghese, N., Seshadri, R., et al. (2021). The IMG/M data management and analysis system v.6.0: new tools and advanced capabilities. Nucleic Acids Res. *49*, D751–D763. https://doi.org/10.1093/nar/gkaa939.

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

STAR★METHODS

KEY RESOURCES TABLE

REAGENT or RESOURCE	SOURCE	IDENTIFIER
Bacterial and virus strains		
E. coli K-12 MG1655	American Type Culture Collection (ATCC)	ATCC 47076; Genbank: NC_000913
NEB 5-alpha Competent *E. coli*	NEW ENGLAND BioLabs	Cat#C2987H
B. subtilis BEST7003	M. Itaya	Genbank: AP012496
Phage SBSphiJ	Doron et al.[2]	Genbank: LT960608.1
Phage SBSphiC	Doron et al.[2]	Genbank: LT960610.1
Phage Bas18	Maffei et al.[57]	NCBI Reference Sequence: NC_073089.1
Critical commercial assays		
DNeasy blood and tissue kit	QIAGEN	Cat#69504
Nextera DNA Library Prep Kit Illumina	Illumina	Cat#15027865; Cat#15027866
DNase-I	Merck	Cat#11284932001
QIAprep Spin Miniprep Kit	QIAGEN	Cat#27106
KLD enzyme mix	New England Biolabs	Cat#M0554S
NEBuilder® HiFi DNA Assembly Master Mix	New England Biolabs	Cat#E2621L
NAD/NADH-Glo	Promega	Cat#G9071; Cat#G9072
Deposited data		
Structure of the ThsB-Tad3 complex	This paper	PDB: 9B7D
Tad3 family PFAM accession	This paper	Pfam: PF25185
Tad4 family PFAM accession	This paper	Pfam: PF25186
Tad5 family PFAM accession	This paper	Pfam: PF25190
Tad6 family PFAM accession	This paper	Pfam: PF25188
Tad8 family PFAM accession	This paper	Pfam: PF25189
Acb3 family PFAM accession	This paper	Pfam: PF25187
Oligonucleotides		
Primers, see Table S24	This paper	N/A
Recombinant DNA		
pSG1	Doron et al.[2]	N/A
pSG-thrC-Phspank	Leavitt et al.[22]	N/A
pBbS8k-RFP	Addgene	Cat#35276
pBbA6c-RFP	Addgene	Cat#35290
Software and algorithms		
MMseqs2 release 12-113e3	Steinegger and Söding[58]	https://github.com/soedinglab/MMseqs2
AlphaFold2 version 2.2	Jumper et al.[59]	https://github.com/deepmind/alphafold
AlphaFold2-Multimer version 2.3	Evans et al.[37]	https://github.com/deepmind/alphafold
MAFFT version 7.490	Katoh and Standley[60]	https://mafft.cbrc.jp/alignment/software/
Breseq version 0.34.1	Deatherage and Barrick[61]	https://barricklab.org/twiki/bin/view/Lab/ToolsBacterialGenomeResequencing
Phenix 1.13-2998	McCoy et al.[62]	https://www.phenix-online.org/
Coot 0.8.9	Emsley and Cowtan[63]	https://www2.mrc-lmb.cam.ac.uk/personal/pemsley/coot/
RING version 4 web server	Clementel et al.[64]	https://ring.biocomputingup.it/
Foldseek release 5.53465f0	van Kempen et al.[65]	https://github.com/steineggerlab/foldseek
Clustal-Omega version 1.2.4	Sievers et al.[66]	http://www.clustal.org/omega/
IQ-TREE version 2.2.0	Minh et al.[67]	http://www.iqtree.org/
iTOL version 5 web server	Letunic and Bork[68]	https://itol.embl.de/

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

EXPERIMENTAL MODEL AND STUDY PARTICIPANT DETAILS

Bacterial strains

E. coli and *B. subtilis* strains were grown in magnesium manganese broth (MMB; LB + 0.1 mM $MnCl_2$ + 5 mM $MgCl_2$) at 37°C shaking at 200 RPM. Whenever applicable, the appropriate antibiotics were added at the following concentrations: For *B. subtilis* strains spectinomycin (100 μg ml^{-1}) and chloramphenicol (5 μg ml^{-1}), and for *E. coli* strains ampicillin (100 μg ml^{-1}) and kanamycin (50 μg ml^{-1}).

The type I and type II Thoeris systems, as well as the type I CBASS system, were cloned previously under their native promoters into the *amyE* locus of the *B. subtilis* BEST7003 genome.[2,41] The type III CBASS system was synthetized and cloned with its native promoter into plasmid pSG1-CBASS as described previously.[11]

Phage strains

The *B. subtilis* phages SBSphiJ (Genbank: LT960608.1) and SBSphiC (Genbank: LT960610.1) were isolated by us as described in a previous study.[2] The *E. coli* phage BAS18 from the BASEL phage collection was kindly contributed by Prof. Alexander Harms.[57] Phages were propagated on either *E. coli* MG1655 or *B. subtilis* BEST7003 by picking a single phage plaque into a liquid culture grown at 37°C to an optical density at 600 nm (OD_{600nm}) of 0.3 in MMB broth until culture collapse (or 3 h in the case of no lysis). The culture was then centrifuged for 10 min at 3,200 *g* and the supernatant was filtered through a 0.2-μm filter to get rid of remaining bacteria and bacterial debris.

METHOD DETAILS

A database of short phage proteins of unknown function

To construct a database of short phage proteins of unknown function, the IMG/VR v3 database[38] was downloaded. Identical protein sequences were removed, resulting in ~32 million non-identical protein sequences. Next, the 32 million proteins were clustered into groups of homologs via an iterative clustering process. First, the protein sequences were clustered using the "cluster" option of MMseqs2 release 12-113e3[58] with default parameters. Then, a representative sequence from each cluster was extracted using the "createsubdb" option of MMseqs2, and the representative sequences were aggregated into groups using the "cluster" option of MMseqs2 with the parameter "-c 0". Next, the 2 sets of clusters were merged using the "mergeclusters" option of MMseqs2. Merged clusters that include at least 40 members and at least one group of homologs from the first step of clustering with an MMseqs2-assigned representative sequence shorter than 200 amino acids were retained for downstream analysis. The MMseqs2-assigned sequence representing the highest number of proteins in the first step of clustering that is also shorter than 200 amino acids was selected as the representative sequence of each merged cluster. Next, the structure of each representative protein was modeled using AlphaFold2 version 2.2[59] with default parameters, with the addition of the IMG/VR v3 database to the default protein databases searched through the AlphaFold2 pipeline to collect homologous sequences. Proteins that were modeled with an average predicted local distance difference test (pLDDT) score lower than 80 were removed from the analysis. Additionally, proteins with a higher number of homologs detected in the MGnify database (a general metagenomic database)[69] than the IMG/VR v3 database (a database of phage proteins) were removed, as these proteins were considered as not primarily carried on phage genomes. The retaining proteins were searched against an annotated database of proteins constructed by us previously[4] using the "search" option of MMseqs2 with default parameters, and the top hit of each phage protein was extracted. Phage proteins similar to a protein with a known function were removed, resulting in a final database of ~38,700 protein sequences, each representing a family of short phage proteins, enriched for proteins with unknown functions. Each one of the ~38,700 proteins was modeled as a homodimer using AlphaFold2-Multimer version 2.2,[37] and predictions with an average predicted aligend error (PAE) score lower than 5 were considered as homodimers for downstream analyses.

Prediction of phage-encoded proteins that inhibit type I Thoeris

To discover phage proteins that bind the type I Thoeris system, each one of the ~38,700 phage protein sequences was predicted as a protein complex with the ThsA and ThsB proteins of type I Thoeris in an iterative process using AlphaFold2-Multimer version 2.3.[37] First, each one of the phage proteins was modeled together with each immune protein generating one predicted complex using the first model of AlphaFold2-Multimer ran with default parameters. Complexes that were generated with a co-folding model confidence score higher than 0.8 were collected for further analysis. Each of the collected interactions was predicted with AlphaFold2-Multimer again, this time generating five predictions based on the five different AlphaFold2-Multimer models. Predicted interactions with a co-folding model confidence score above 0.8 in at least three of the five predicted complexes were collected. Finally, the collected complexes were modeled again using AlphaFold2-Multimer, generating five predictions per each one of the five AlphaFold2-Multimer models, resulting in 25 predicted protein complexes for each pair of immune and phage proteins. Predicted interactions with a co-folding model confidence score above 0.8 in at least 15 of the 25 predicted complexes were defined as final candidate anti-defense proteins, and were taken for experimental verification. Phage proteins that were predicted to form a homodimer were presented in two copies when modeled as a protein complex with an immune protein. Multiple sequence alignments presented in Figures S2 and S8B were computed with MAFFT version 7.490[60] and visualized using Jalview.[70]

Cloning of candidate anti-defense genes

Anti-defense genes were synthesized and cloned by Genscript Corp. The anti-defense candidates for type I and type II Thoeris and type I CBASS were cloned into the pSG-thrC-Phspank vector[22] and transformed into NEB 5-alpha competent cells. The cloned vector was subsequently transformed into *B. subtilis* BEST7003 cells containing the respective defense system integrated into the *amyE* locus,[2,41] resulting in cultures expressing both a defense system and the corresponding anti-defense gene candidate, integrated into the *amyE* and *thrC* loci, respectively. As a negative control, a transformant with an identical plasmid containing sfGFP instead of the anti-defense gene was used. Transformation to *B. subtilis* was carried out using MC medium as previously described[2] and transformants were plated on LB agar plates supplemented with 5 μg ml^{-1} chloramphenicol and incubated overnight at 30°C. Whole-genome sequencing was then applied to all transformed *B. subtilis* strains, and Breseq (v0.34.1) analysis[61] was used to verify the integrity of the inserts and lack of mutations. Substitutions and indels in anti-defense proteins were made using the KLD Enzyme Mix (NEB) with the primers listed in Table S24.

The anti-defense candidate for type III CBASS was cloned into the pBbS8k vector (Addgene #35276) and subsequently transformed into NEB 5-alpha competent cells. The cloned vector was further transformed into *E. coli* MG1655 cells containing the pSG1-CBASS system.[11] As controls, transformants with an identical plasmid containing RFP instead of the anti-defense gene were used.

Plaque assays

Phage titer was determined using the small drop plaque assay method.[71] A 300 μl (*E. coli*) or 400 ul (*B. subtilis*) volume of overnight culture of bacteria was mixed with 0.5% agar and 30 ml MMB and poured into a 10-cm square plate followed by incubation for 1 h at room temperature. In cases of bacteria expressing anti-defense candidates, 1 mM IPTG (*B. subtilis*) or 0.2% arabinose (*E. coli*) was added to the 30 ml MMB 0.5% agar. Tenfold serial dilutions in MMB were carried out for each of the tested phages and 10-μl drops were put on the bacterial layer. After the drops had dried up, the plates were inverted and incubated at 25°C overnight. PFUs were determined by counting the derived plaques after overnight incubation and lysate titer was determined by calculating PFUs per milliliter. When no individual plaques could be identified, a faint lysis zone across the drop area was considered to be 10 plaques. The efficiency of plating was measured by comparing plaque assay results for control bacteria and those for bacteria containing the defense system and/or a candidate anti-defense gene.

Protein co-expression for biochemical assessment of metabolites

B. subtilis BEST7003 cultures, co-expressing a genomically-integrated *B. cereus* MSX-D12 Thoeris system (ThsA$_{N112A}$B)[7] under native promoter and either Tad3, Tad4, Tad5, or Tad6 under the Physpank promoter were grown in 50 ml MMB media supplemented with 1mM IPTG. Cultures were grown for ~2 h at 37°C, 200 RPM until reaching an OD$_{600}$ of 0.3. Cultures were then infected with phage SBSphiJ at MOI = 10. Cultures were collected 120 min after infection and centrifuged (3200 *g*) for 10 min at 4°C. The pellet was flash-frozen and stored at −80°C. Control cultures included *B. subtilis* BEST7003 expressing GFP without the Thoeris system, and *B. subtilis* BEST7003 co-expressing the Thoeris system (ThsB + ThsA$_{N112A}$) with GFP under the Physpank promoter.

Human SARM1 TIR domain (position 561–724, NCBI Ref seq: NP_055892.2) and BdTIR (*Brachypodium distachyon* TIR, NCBI ref seq: XP_003560074.3) were co-expressed with Tad4 in *E .coli* MG1655. Cultures were grown in 50 ml MMB media with ampicillin (100 μg ml^{-1}) and chloramphenicol (30 μg ml^{-1}). Initially, cultures were grown for ~2 h at 37°C, 200 RPM until reaching an OD$_{600}$ of 0.3. This was followed by inducing expression of Tad4 with 1mM IPTG and lowering growth temperature to 30°C. After 30 min, 0.2% arabinose was added to induce the expression of either the hSARM1 TIR domain or BdTIR. Cells were harvested after 3 h by centrifugation at 3200 *g* for 15 min in 4°C followed by flash freezing the pellets and storing them in -80°C. Control cultures included *E. coli* MG1655 co-expressing hSARM1$_{TIR}$ or BdTIR with RFP and a negative control expressing only RFP (Addgene #35290).

Preparation of filtered cell lysates

To extract cell metabolites from frozen pellets, 1 ml of 100 mM Na-phosphate buffer (pH 8.0) was added to each pellet. Tubes were then incubated for 10 min at 25°C and returned to ice. Samples were transferred to FastPrep Lysing Matrix B in a 2 ml tube (MP Biomedicals, no. 116911100) and lysed at 4°C using a FastPrep bead beater for 2 × 40 s at 6 m s^{-1}. Tubes were then centrifuged at 4°C for 10 min at 15,000 *g*. Supernatant was then transferred to an Amicon Ultra-0.5 Centrifugal Filter Unit 3 kDa (Merck Millipore, no. UFC500396) and centrifuged for 45 min at 4°C, 12,000 *g*.

Enzymatic assays

The ThsA protein was expressed and purified as previously described,[12] and ThsA-based NADase activity assay for the detection gcADPR was carried out as described previously.[12] The NAD/NADH-Glo (Promega) kit was used for directly measuring the NAD$^+$ levels in filtered cell lysates. The lysates were diluted 1:150 in 0.1 M Na-phosphate buffer, pH 8.0. Reactions were performed in a volume of 10 μl (5 μl of sample + 5 μl of reaction mixture) according to ratios recommended in the manufacturer's instructions. Luciferin signal, proportional to the amount of NAD$^+$, was detected using the kit luciferase enzyme. Tecan Infinite 200 PRO plate reader was used to monitor the developing luminescent signal. NAD$^+$ concentrations were calculated from the calibration curve using a set of NAD$^+$ standards with known concentrations.

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

Protein expression and purification

ThsB, BdTIR, and SARM1$_{TIR}$ sequences were codon optimized for *E. coli*, synthesized (Integrated DNA Technologies), and cloned into a custom pET-based expression vector with an N-terminal 6×His-SUMO2 tag. A spacer (AAAGAGGAGAAATTAACT) containing a second ribosome binding site was inserted directly downstream of ThsB, BdTIR, and SARM1$_{TIR}$, and codon-optimized sequences for Tad3-6 were cloned into the second open reading frame for co-expression studies. Expression plasmids were transformed into RIL cells (Agilent) and plated on MDG plates (1.5% Bacto agar, 0.5% glucose, 25 mM Na_2HPO_4, 25 mM KH_2PO_4, 50 mM NH_4Cl, 5 mM Na_2SO_4, 0.25% aspartic acid, 2–50 μM trace metals, 100 μg ml^{-1} ampicillin, 34 μg ml^{-1} chloramphenicol). Colonies were picked into 30 ml MDG liquid media and grown overnight at 37°C with shaking. Overnight cultures were diluted 1:100 into 2 L of M9ZB media (47.8 mM Na_2HPO_4, 22 mM KH_2PO_4, 18.7 mM NH_4Cl, 85.6 mM NaCl, 1% casamino acids, 0.5% glycerol, 2 mM $MgSO_4$, 2–50 μM trace metals, 100 μg ml^{-1} ampicillin, 34 μg ml^{-1} chloramphenicol) and grown at 37°C with shaking for 5–6 hours until OD = 1.5-2.0. Cultures were cooled on ice for 15 minutes before addition of IPTG to a final concentration of 500 μM. Induced cultures were then incubated at 16°C with shaking for 16 hours before centrifugation (4000*g* for 20 minutes at 4°C). Cell pellets were resuspended in 120 ml lysis buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 10% glycerol, 30 mM imidazole, 1 mM TCEP) and lysed by sonication using 10"on/20"off pulses at 70% power for a total sonication time of 5 minutes (Qsonica). Lysates were then centrifuged (50,000*g* for 30 min at 4°C) and clarified supernatants were poured over 8 ml of Ni-NTA resin twice (Qiagen). Resin was then washed with 20 ml lysis buffer, 50 ml wash buffer (20 mM HEPES-KOH pH 7.5, 1 M NaCl, 10% glycerol, 30 mM imidazole, 1 mM TCEP). Bound protein was eluted in 20 ml elution buffer (20 mM HEPES-KOH pH 7.5, 400 mM NaCl, 10% glycerol, 300 mM imidazole, 1 mM TCEP). Samples were dialysed overnight in dialysis tubing with a 20 kDa molecular weight cutoff (Ward's Science), and SUMO2 tag cleavage was carried out with recombinant human SENP2 protease as previously described.[50] Proteins were further purified by size-exclusion chromatography using a 16/600 Superdex 200 (Cytiva) and concentrated to a final concentration of >15 mg ml^{-1} and aliquots were flash frozen in liquid nitrogen and stored at −80°C.

Co-expression of ThsB homologs with Tad3 and Tad4

ThsB homologs were codon-optimized for *E. coli*, synthesized (Twist Bioscience), and cloned into a custom pET-based expression vector with an N-terminal 6×His-SUMO2 tag. A spacer sequence (AAAGAGGAGAAATTAACT) containing a second ribosome binding site directly followed each *thsB* homolog gene, enabling co-expression of downstream *tad3* or *tad4* genes. Expression plasmids were transformed into LOBSTR-BL21(DE3)-RIL cells (Kerafast). Cultures were grown in MMB media (200 RPM, 37°C) until reaching an OD of 0.5, followed by the addition of 0.5 mM IPTG and a temperature shift to 30°C. Cells were harvested after 3 hours by centrifugation at 3200*g*, 4°C, for 15 minutes, and pellets were stored at -80°C.

SDS-PAGE analysis

Protein purity and complex formation were assessed by SDS-PAGE by mixing 10 μl of protein (containing 1-10 μg of protein sample) with 4 ul of loading buffer. The samples were run on a 15% gel for 45 minutes at 200 V. Proteins were stained using BrilliantBlue coomassie stain (VWR) and visualized using a ChemiDoc MP Imaging system (BioRad).

Protein mass spectrometry analysis

Gel bands were excised from the gel, sliced into 1-2 mm pieces and placed in a microcentrifuge tube. Gel bands were destained with 25mM NH4HCO3 in 50% acetonitrile (ACN) and then vacuum dried. Protein disulfide bonds were reduced by saturating the dry gel bands with 10 mM dithiothreitol in 25 mM NH4HCO3 at 56°C for 1 h and alkylated with 55 mM iodoacetamide in 25 mM NH4HCO3 in the dark for 45 min at room temperature. Bands were washed twice with 25 mM NH4HCO3 and twice with 25 mM NH4HCO3 in 50% ACN. Bands were vacuum dried followed by rehydrating the gel pieces with 500 ng trypsin in 25 mM NH4HCO3 overnight at 37°C. Peptides were then extracted by addition of 50% ACN/5% formic acid, vortexed, centrifuged, and the supernatant collected. The digestions were stopped by addition of trifluroacetic acid (1% final concentration). The samples were stored in -80C until further analysis.

Dry digested samples were dissolved in 97:3% H2O/acetonitrile + 0.1% formic acid. Each sample was loaded using split-less nano-Ultra Performance Liquid Chromatography (10 kpsi nanoAcquity; Waters, Milford, MA, USA). Desalting of the samples was performed using a reversed-phase Symmetry C18 trapping column (180 μm internal diameter, 20 mm length, 5 μm particle size; Waters). The peptides were then separated using a T3 HSS nano-column (75 μm internal diameter, 250 mm length, 1.8 μm particle size; Waters) at 0.35 μl/min. The mobile phase consisted of: A) H2O + 0.1% formic acid and B) acetonitrile + 0.1% formic acid. Peptides were eluted from the column into the mass spectrometer using the following gradient: 4% to 27% phase B in 55 min, 27% to 90% phase B in 5 min, maintained at 90% for 5 min and then back to the initial conditions.

Peptides were detected using a nanoUPLC that was coupled through a nanoESI emitter (10 μm tip; New Objective; Woburn, MA, USA) to a quadrupole orbitrap mass spectrometer (Q Exactive HF, Thermo Scientific) using a FlexIon nanospray apparatus (Proxeon). Data was acquired in data dependent acquisition (DDA) mode, using a Top10 method. MS1 resolution was set to 120,000 (at 200 m/z), mass range of 375-1650 m/z, AGC of 3e6 and maximum injection time was set to 60 msec. MS2 resolution was set to 15,000, quadrupole isolation 1.7 m/z, AGC of 1e5, dynamic exclusion of 20 sec and maximum injection time of 60 msec.

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

SEC-MALS analysis

Purified proteins complexes were diluted in SEC-MALS running buffer (250 mM KCl, 20 mM HEPES-KOH pH 7.5, 1 mM TCEP) to 2 mg ml^{-1} and 100 μl was analyzed on an SRT SEC-150 column with 5 μm bead size, 300 Å pore, 7.8x300 mm column dimensions (Sepax). Protein concentration was determined using a Wyatt Optilab T-rEX Refractive Index Detector assuming dn dc–1 of 0.185 and the molar mass of the selected peaks was determined using a Wyatt Dawn Heleos II Multi-Angle Light Scattering detector and ASTRA software. Data were visualized using Prism9 (GraphPad).

Crystallization and structure determination

Crystals of the ThsB-Tad3 complex were grown by mixing 1 μl purified protein (10 mg ml^{-1}, 20 mM HEPES-KOH pH 7.5, 80 mM KCl, 1 mM TCEP) with 1 μl reservoir solution (0.1 M Tris-HCl pH 8.0, 0.2 M $MgCl_2$, 25% PEG3350) and incubating at 18°C in 15 well trays (NeXtal) in hanging drop format. Crystals were harvested in reservoir solution supplemented with 10% ethylene glycol for cryoprotection after 1–2 days of growth. X-ray diffraction data were collected at the Advanced Photon Source (beamlines 24-ID-C and 24-ID-E), and data were processed using the SSRL autoxds script (A. Gonzalez, Stanford SSRL). For phase determination, AlphaFold2 models of ThsB and Tad3 were used for molecular replacement search and an initial map was generated using phaserMR (Phenix).[62] Model building was performed using Coot,[63] and then refined in Phenix. Statistics were analyzed as described in Table S25.[72,73] Final structures were refined to stereochemistry statistics for Ramachandran plot (favoured/allowed), rotamer outliers and MolProbity score as follows: (97.84% / 2.02%), 0.59%, 1.38. See Table S25 and the Data and code availability statement for the deposited PDB codes. All structure figures were generated with PyMOL 2.4.2.

Knockin of Tad3 into phage SBSphiJ

The DNA sequence of *tad3* was amplified from the *tad3*-containing pSG-thrC-Phspank plasmid using KAPA HiFi HotStart ReadyMix (Roche, catalogue number KK2601) with the primer pair IO95Tad3_9F and IO96Tad3_9R (Table S24). The backbone fragment with the upstream and downstream genomic arms (±1.2 kb) for the integration site of *tad3* was amplified from the plasmid used previously for knockin of the *tad1* gene,[22] with the primer pair IO34_JKIR2 and IO22_JKIF (Table S24). Cloning was carried out using the NEBuilder HiFi DNA Assembly cloning kit (NEB, number E5520S) and the cloned vector was transformed into NEB 5-alpha competent cells. The cloned vector was subsequently transformed into the *thrC* site of *B. subtilis* BEST7003.

The *tad3*-containing B. subtilis BEST7003 strain was then infected with phage SBSphiJ with a multiplicity of infection (MOI) of 0.1 and cell lysate was collected. Tad3 lysate was used to infect a Thoeris-containing *B. subtilis* culture in two consecutive rounds with an MOI of 2 in each round at 30°C. Several plaques were collected and screened using PCR for the desired insertion within the phage genome. Phages with *tad3* were purified three times on *B. subtilis* BEST7003. Purified phages were verified again for the presence of *tad3* using PCR amplifications.

Selection of homologs of anti-defense candidates for binding analysis

Homologs of candidate anti-defense proteins were identified in the IMG/VR v3 database using the "search" option of MMseqs2 release 12-113e3 with the parameter "-c 0.8". Then, the homologs of each candidate anti-defense protein were separated into 10 bins based on their sequence identity percentage to the query anti-defense candidate. A random sequence was selected from each bin and predicted as a complex with the relevant immune protein using AlphaFold2-Multimer version 2.3, generating five predictions per each one of the five AlphaFold2-Multimer models.

Prediction of phage encoded proteins that inhibit type II Thoeris or CBASS

Inhibitors of type II Thoeris and CBASS were predicted by analyzing the ThsB and ThsA proteins of type II Thoeris and the CD-NTase protein from *E. coli* KTE188 together with each of the phage proteins as described above, except that two additional filtering steps were applied to the results. First, interactions that were predicted to have less than 25 residues of the immune protein interacting with the candidate inhibitor based on an analysis in the RING version 4 server[64] were removed. Additionally, 10 homologs of each predicted binder were selected and modeled as a complex with the immune protein using AlphaFold2-Multimer as described above. Candidate inhibitors that did not have diverse homologs predicted to bind the immune protein with an average co-folding model confidence score higher than 0.75 were removed. For each of the retaining candidates, the homolog having the highest co-folding model confidence score was selected for experimental verification.

Distribution of anti-defense proteins in phage genomes

Homologs of anti-defense genes were identified in the IMG/VR v4 database[30] by conducting sequence-based and structure-based homology searches. For this, ~5.5 million phage scaffolds labeled as "high-confidence virus" were downloaded from the IMG/VR v4 database. Homologous sequences of the anti-defense proteins detected in this study were identified using the "search" option of MMseqs2 release 12-113e3 with the parameters "-c 0.8 –cov-mode 2". To identify structural homologs, the downloaded proteins from IMG/VR v4 were clustered using the "cluster" option of MMseqs2 with default parameters. Next, a representative sequence was extracted from each cluster containing at least 30 non-identical members, and its structure was predicted using AlphaFold2 version 2.3 with default parameters, resulting in 182,179 phage protein structures. Structures of the anti-defense proteins were searched against this set of 182,179 phage protein structures using foldseek release 5.53465f0[65] with default parameters. Hits

Please cite this article in press as: Yirmiya et al., Structure-guided discovery of viral proteins that inhibit host immunity, Cell (2025), https://doi.org/10.1016/j.cell.2024.12.035

with probability of 1.0 were collected with all their cluster members as structural homologs of the anti-defense proteins. Finally, hits longer than 800 amino acids were discarded. Homologs of anti-defense proteins were detected in the Metagenomic Gut Virus (MGV) database[32] using the "search" option of MMseqs2 with the parameter "-c 0.8", using all of the anti-defense homologs detected in IMG/VR v4 as queries.

Identification of ThsB homologs and phylogenetic reconstruction

Protein sequences of all genes in 38,167 bacterial and archaeal genomes were downloaded from the IMG database in October 2017 and clustered into groups of homologs as previously described.[4] Protein sequences of genes from metagenome scaffolds were downloaded from the IMG database in April 2020 and merged into the set of protein clusters from isolate genomes as previously described.[9] Protein clusters were assigned with a Pfam annotation based on the most common Pfam annotation of individual proteins in the cluster. Protein clusters were analyzed to predict defense systems as previously described.[2,4] 1,225 protein sequences belonging to clusters annotated with one of the TIR Pfam annotations (PF18567, PF01582, PF08937, PF10137, and PF13676) and found within type I Thoeris operons were defined as ThsB homologs and collected for downstream analysis. The protein sequences were aligned using Clustal-Omega version 1.2.4[66] with default parameters. The phylogenetic tree was constructed using IQ-TREE v2.2.0[67] with the parameters -m LG and -B 1000, and visualized by the iTOL web application (v5).[68] Each one of the 1,225 ThsB proteins was modeled together with Tad3 and Tad4 as a protein complex using AlphaFold2-Multimer, generating one prediction per each one of the five AlphaFold2-Multimer models. The average model confidence score across the five models was defined as the score of the interaction.

LC-MS polar metabolite analysis

Samples were centrifuged twice (20,800*g*) and transferred to HPLC vials. Sample evaluation was conducted according to the method described by Zheng et al.[74] with minor modifications as described below. Briefly, analysis was performed using Acquity I class UPLC System combined with mass spectrometer Q Exactive Plus Orbitrap™ (Thermo Fisher Scientific), operated in a negative ionization mode with a scan range of 70 – 1050 m/z. The LC separation was done using the SeQuant Zic-pHilic (150 mm × 2.1 mm) with the SeQuant guard column (20 mm × 2.1 mm) (Merck). Mobile phase consisted of two different mobile phases: The first phase (A) consisting of acetonitrile, and the second phase (B) consisting of 20 mM ammonium carbonate, 0.1% ammonia hydroxide in DDW:acetonitrile (80:20, v/v). The chromatographic conditions were as described in Gnainsky et al.[75] Data processing was done using the TraceFinder version 5.1 (Thermo Fisher Scientific) software. Identification of 2′3′-cGAMP was done by MS-MS and retention time, and validated by the injection of a commercially available standard (Biolog).

QUANTIFICATION AND STATISTICAL ANALYSIS

The average of three triplicates is shown throughout with individual points overlaid, unless stated otherwise.

Supplemental figures

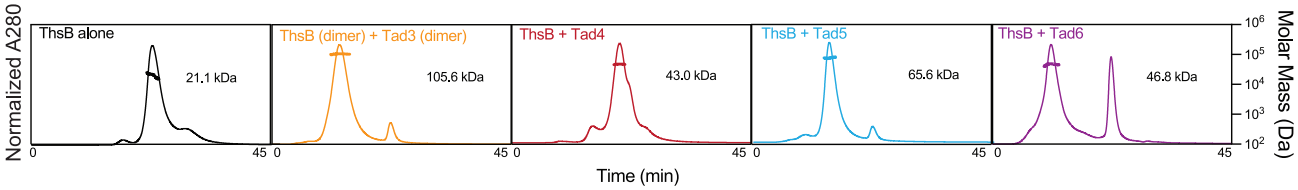


Figure S1. Phage-derived anti-Thoeris proteins bind ThsB, related to Figure 1

SEC-MALS analyses demonstrate that each of the Thoeris inhibitors elutes together with ThsB as a complex when co-expressed with it.

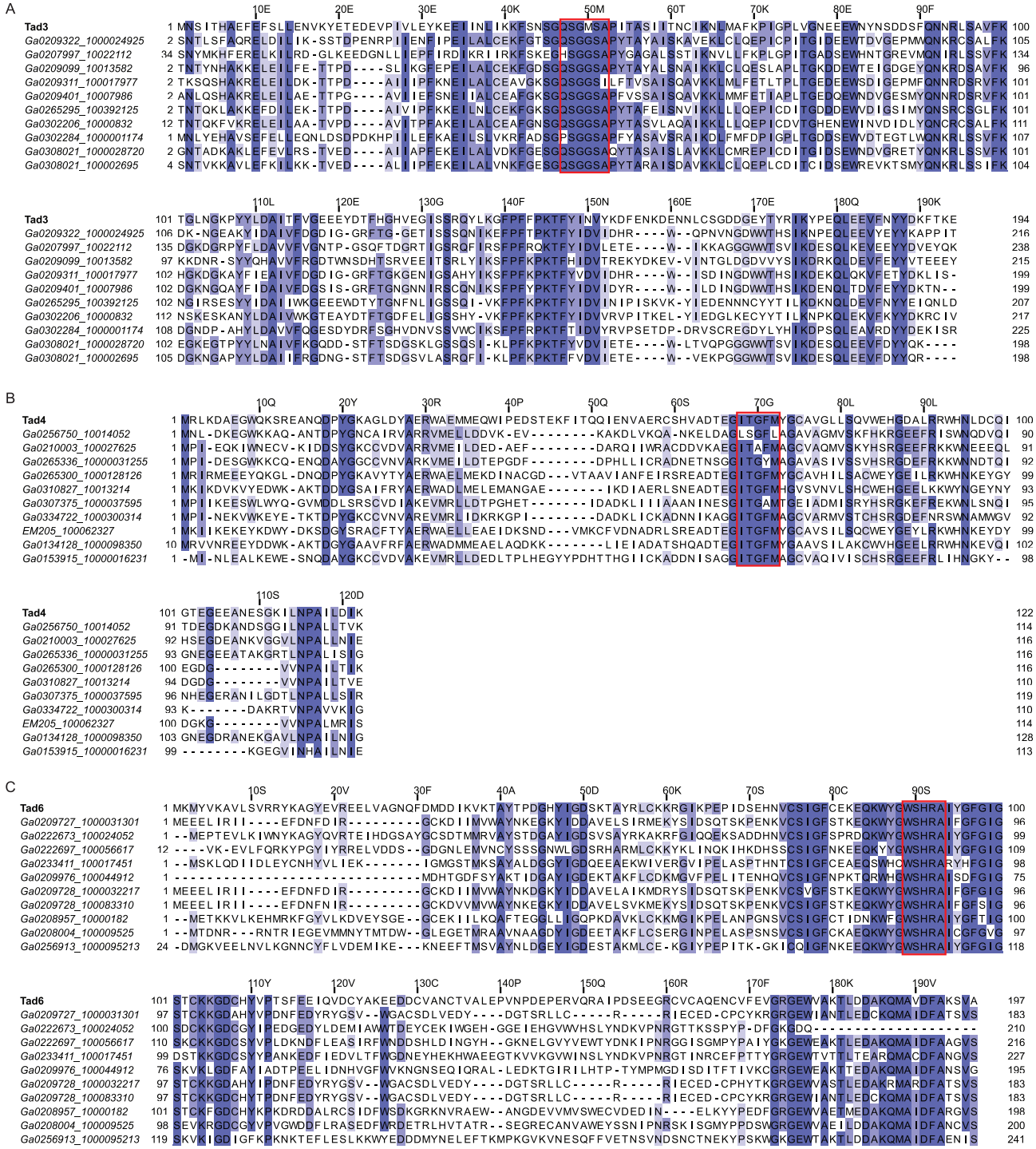


Figure S2. Multiple sequence alignment of anti-Thoeris proteins, related to Figure 2

Sequence alignment of (A) Tad3, (B) Tad4, and (C) Tad6. Alignment was generated using 10 randomly selected homologous sequences of each anti-Thoeris protein. Conserved residues are colored in purple. Red boxes indicate residues that comprise a loop that blocks the ThsB pocket. Positions in the alignments not occupied in the verified anti-Thoeris protein sequences are not shown.

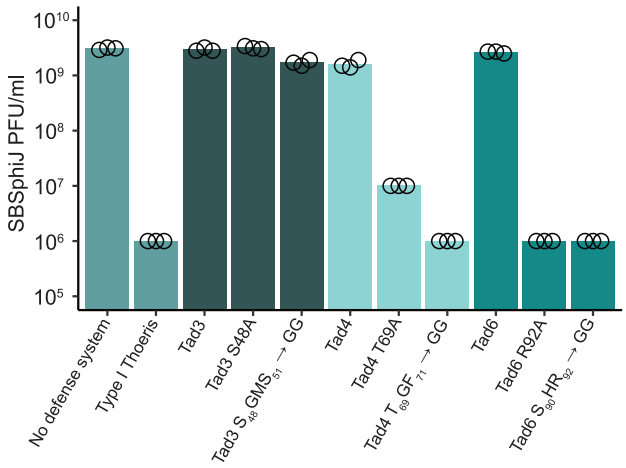


Figure S3. Substitutions in the loops predicted to block the ThsB pocket impact anti-Thoeris function, related to Figure 2

Data represent PFUs/mL of phage SBSphiJ infecting cells co-expressing the Thoeris system and a wild-type (WT) or mutated anti-Thoeris protein, as well as control cells expressing no defense system and cells expressing the Thoeris system alone ("type I Thoeris"). Shown is the average of three replicates, with individual data points overlaid. Data for type I Thoeris and the WT anti-defense proteins are the same as those presented in Figure 1C.

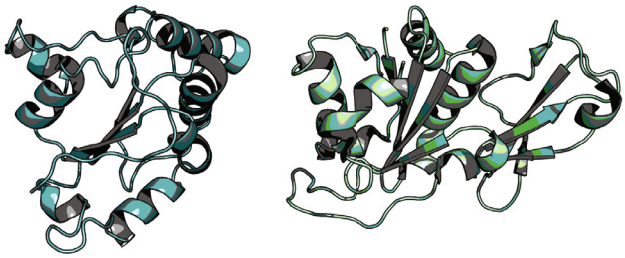


Figure S4. Structural comparison between the ThsB proteins of type I and type II Thoeris systems, related to Figures 1 and 4

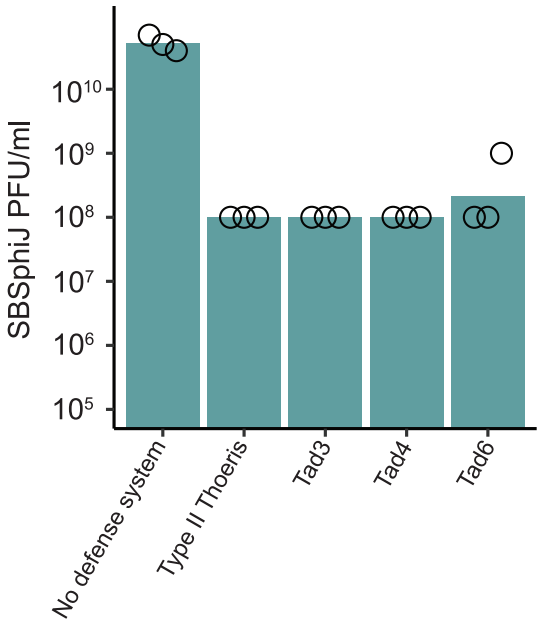


Figure S5. Type I Thoeris inhibitors do not inhibit type II Thoeris, related to Figure 4

Data represent PFU/mL of phage SBSphiJ infecting cells co-expressing the type II Thoeris system and a type I Thoeris inhibitor, as well as control cells expressing no defense system and cells expressing the type II Thoeris system alone ("type II Thoeris"). Shown is the average of three replicates, with individual data points overlaid.

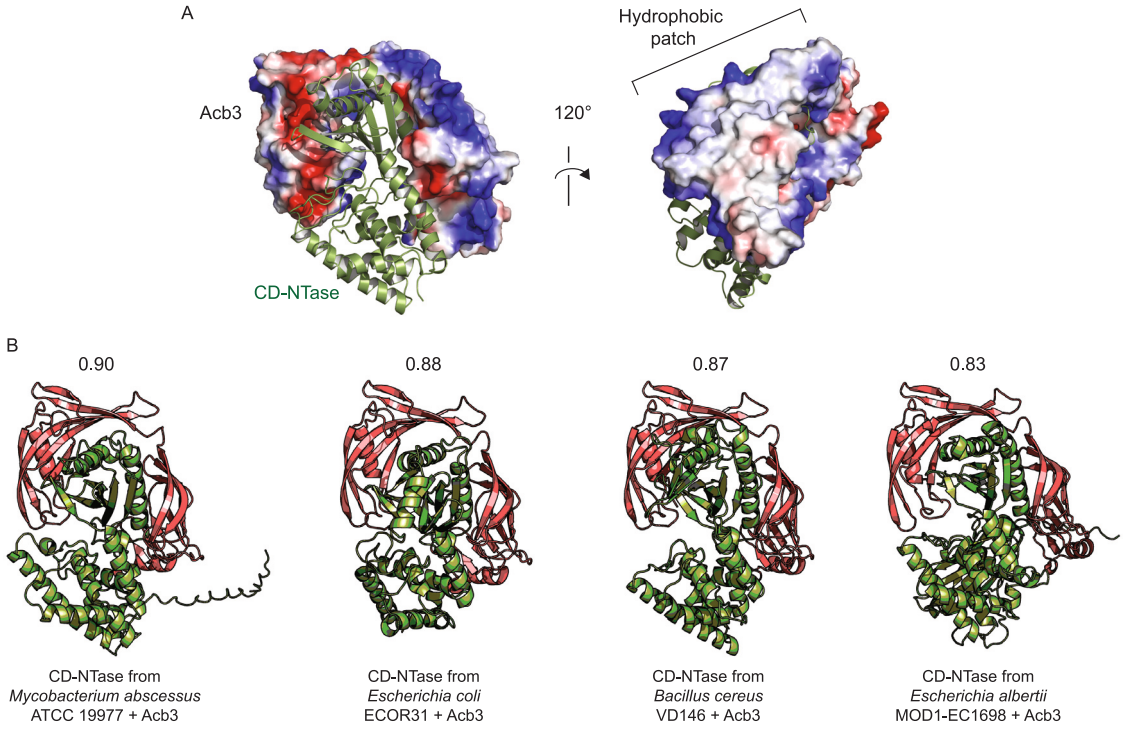


Figure S6. Acb3 is predicted to bind diverse CD-NTase enzymes, related to Figure 4

(A) Overview of the predicted complex between Acb3 and the CD-NTase of the *E. coli* KTE188 CBASS. Acb3 surface electrostatics are shown to highlight the hydrophobic patch that covers the CD-NTase ligand binding groove.

(B) Predicted complexes formed between Acb3 (red) and four bacterial CD-NTase enzymes (green). AlphaFold2-Multimer model co-folding confidence scores are indicated above the models.

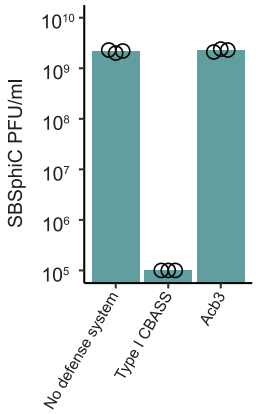
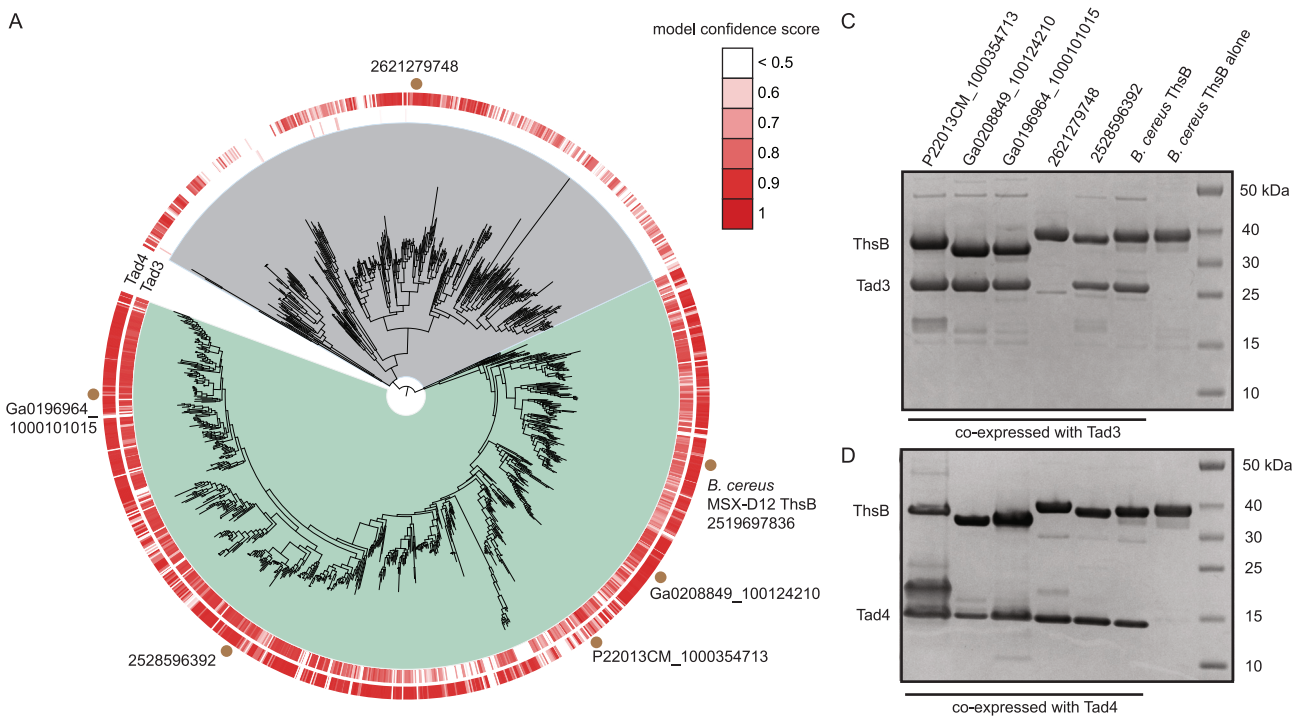


Figure S7. Acb3 inhibits type I CBASS from *B. cereus* VD146, related to Figure 4

Data represent PFU/mL of phage SBSphiC infecting cells co-expressing the type I CBASS system and Acb3, as well as control cells expressing no defense system and cells expressing the type I CBASS system alone (“type I CBASS”). Shown is the average of three replicates, with individual data points overlaid.



B

```
2621279748          1 MAIFTKQEFESIARIKAGYKGLSSVLNESRFYSKASALTSLFLSHSHYDKAI------VEQAKIF---FENLGISIYVDWAD  73
B. cereus ThsB      1 ----------------------------MAKRVFF-----------SFHYQDVIDFRVNVVRNHWVTKLNQSAAGVFDASLWED  45
Ga0208849_100124210 1 ----------------------------MARKVFF-----------SFHYERDI-WRTNVVRNSGVVE-GSAAAGFHDASLWEE  43
P22013CM_1000354713 1 ----------------------------MARRAFF-----------SFHYENDV-WRANIVRNSWVTKEDREAAGFVDAADFEE  44
2528596392          1 --------------------------MSKRQVFF-----------SFHYTNDN-WRAGQIRNMGKVDGSST----FSDNDWEE  41
Ga0196964_1000101015 1 ----------------------------MARRVFY-----------SFHYDNDS-WRTSQVRNIGVIEGNRP----APDNDWEA  40

2621279748          74 ETMPEKTNGITAT-RIKSKIITNDKFVLLATNSAVISRWCNWFVGIGDTYKLNSDKICILPLSDN--K--------------  138
B. cereus ThsB      46 AK---KTSDIALKRLINGGLNNTSVTCVLIGSQTFNRRWVRYEIMKSIEKGNKIIGIHINAFKDKYGNIKSKGPNPFDYLGY  124
Ga0208849_100124210 44 AK---RKGDAEVKKLIDKGLIGTSVTVVLIGAETSTRKFVDYEIEQSIARGNGLLGIKISAIRDRDGKTDVQGAVPAR----  118
P22013CM_1000354713 45 VK---KGGDAAIKRWIDNQLVGTSVTVVLIGLETSNRPYVQYELQQSYSRGNGLPGIYIHQLKDRNGNTSSKGSNAFGEIGK  123
2528596392          42 VK---EKTDNKIKKWIDKQLNLRSCLVVLIGQKTANRKWINYEIKRAYELNKGIVGIYIHKLENSSGIQDKKGNNPFDYIPL  120
Ga0196964_1000101015 41 VK---KGGDPAIKRWIDNQMYGRSCTIVLVGSNTAGRKWIEYEIKKSWEEGMGVVGIRIHNLKNYAQQQSYWGNNPFALSAN  119

2621279748          139 ----------NDWNGNEYLQIYPRIEP----------------VYKDSYSFYDNIFKVIYPDGHSTWLDDWLKK-------  186
B. cereus ThsB      125 QYSSDGKQLHLYEWTGGKWEE-YKDLAPYRVNQIAPESLRGKFYSLSSVYRVYDWVADDGY-NKFSSWVN------------  192
Ga0208849_100124210 119 ----------------------------------------LTNAGAPSYTWDRD-----KFGDWVEAAYKKAHPQG--  149
P22013CM_1000354713 124 DSN-----------------------------------GNTVYFSVNYPCYDWVDDDGY-NNLGKWIEEAAKKAGK----  163
2528596392          121 NKNDEK----------------------------LSKYVKCFDSIYPFSKDVYNDIK-NNIENLIEDCIKNKPSTWQN  169
Ga0196964_1000101015 120 ---GIW----------------------------LEDKISLHNPPYVDSKDVYNYIA-DNIGNWTEEAVRNRNRY---  162
```

Figure S8. Tad3 and Tad4 antagonize diverse ThsB enzymes, related to Figures 1 and 2

(A) Phylogenetic analysis of ThsB homologs from type I Thoeris systems reveals two major clades, represented in gray and green. AlphaFold2-Multimer model confidence scores of co-folding ThsB homologs with Tad3 and Tad4 are displayed with white-to-red scales surrounding the phylogenetic tree. ThsB homologs tested experimentally for binding to Tad3 and Tad4 are depicted on the tree periphery with brown circles, with the IMG IDs[76] indicated of each tested gene.

(B) Multiple sequence alignment of six ThsB homologs that were tested experimentally for Tad3 and Tad4 binding. These sequences represent the phylogenetic diversity of ThsB proteins from type I Thoeris systems. Conserved residues are colored in purple.

(C and D) Pull-down assays of 6×His-SUMO2-tagged ThsB homologs co-expressed with Tad3 (C) and Tad4 (D) validate the co-folding predictions. Shown are SDS-PAGE analyses.

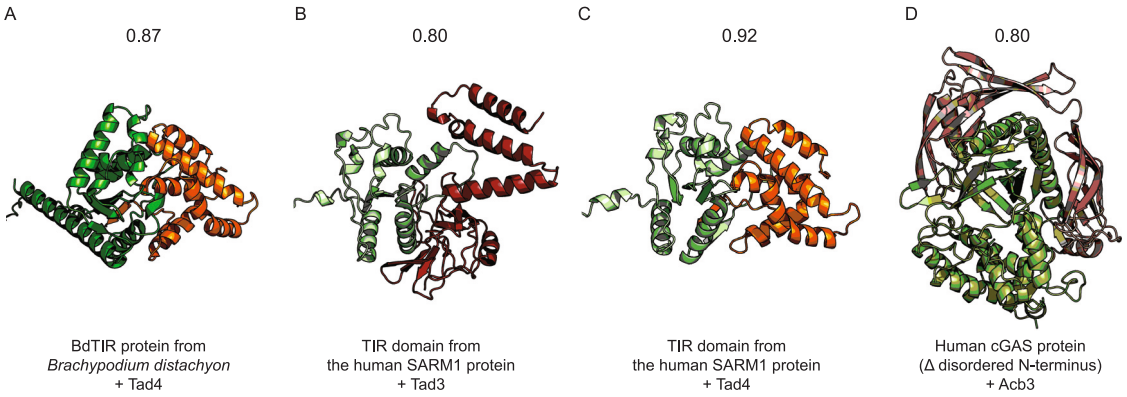
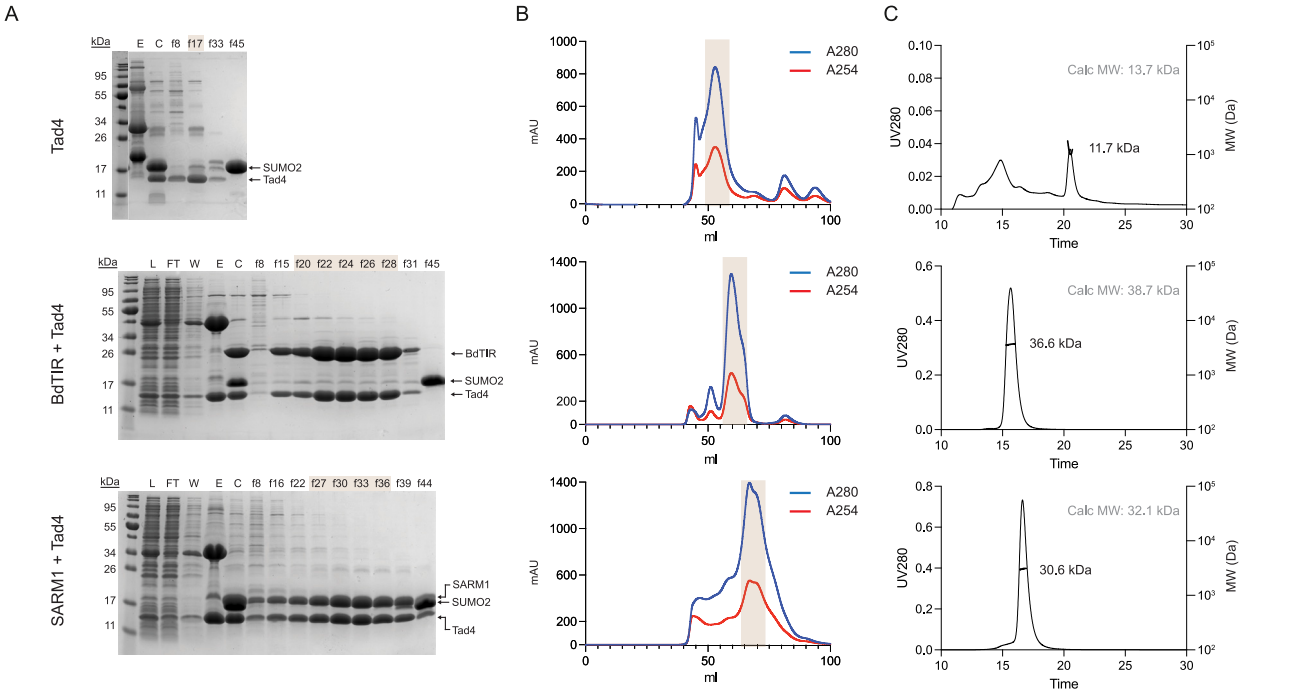
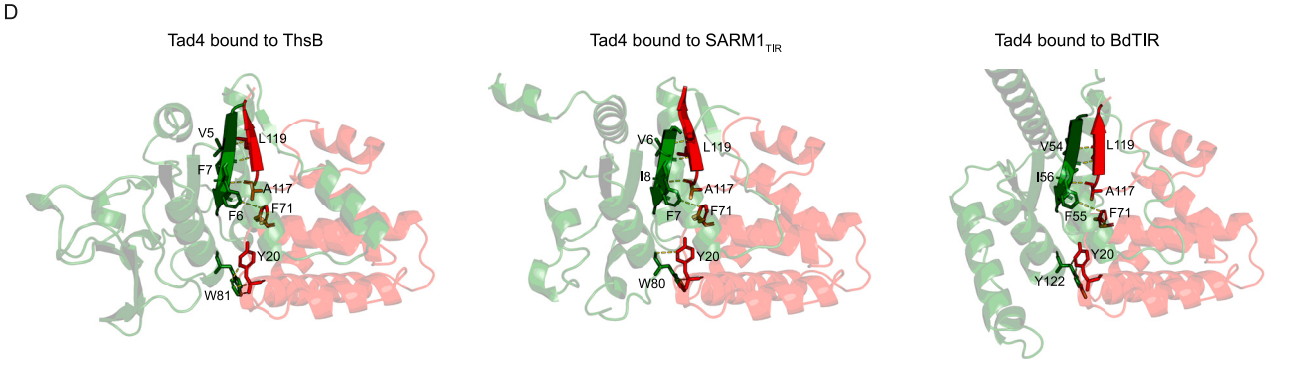
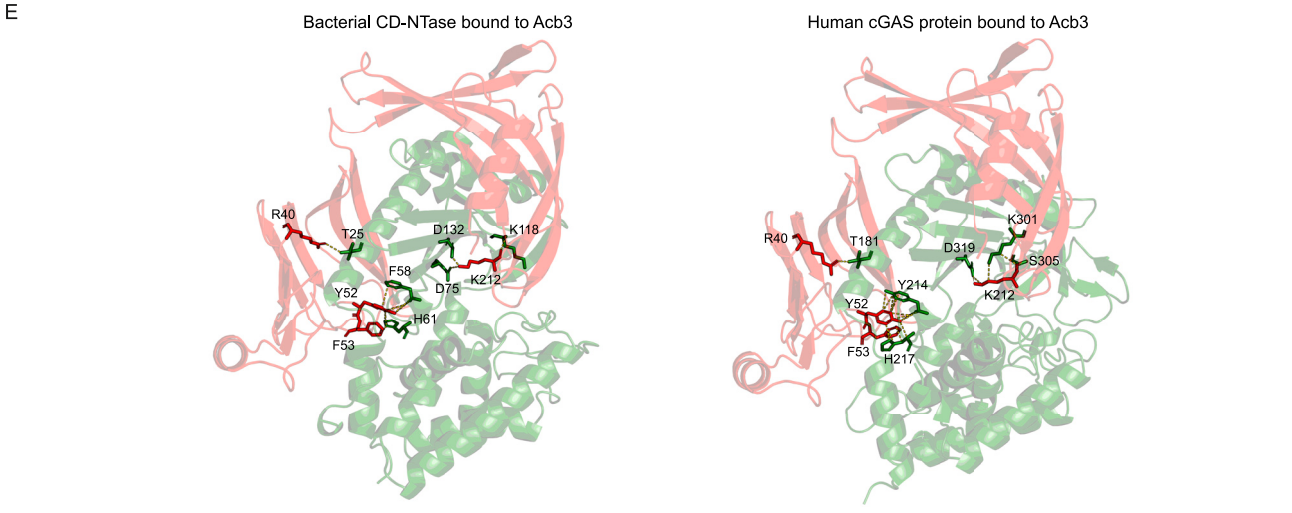


Figure S9. Phage-derived anti-defense proteins (red/orange) are predicted to bind eukaryotic immune proteins (green), related to Figure 5
Predicted complexes formed between (A) BdTIR and Tad4, (B) the TIR domain of SARM1 and Tad3, (C) the TIR domain of SARM1 and Tad4, and (D) cGAS and Acb3. AlphaFold2-Multimer model co-folding confidence scores are indicated above the models.







(legend on next page)

Figure S10. Conservation between eukaryotic and prokaryotic immune proteins renders the eukaryotic homologs susceptible to phage-encoded inhibitors, related to Figure 5

(A) SDS-PAGE gels stained with Coomassie blue demonstrating purification of the indicated protein or protein complexes. L, total cell lysate; FT, Ni-NTA flowthrough; E, Ni-NTA elution; C, 6xHis-SUMO2 tag cleavage; W, wash; f, fractions from SEC.

(B) UV traces from SEC of the indicated protein or protein complex. Shaded area represents the cleanest fractions that were collected and analyzed by SEC-MALS.

(C) SEC-MALS traces of the shaded peaks of the indicated protein or protein complexes demonstrate that both BdTIR and $SARM1_{TIR}$ elute together with Tad4 as a complex when co-expressed with it.

(D) Predicted complexes formed between Tad4 (red) and the TIR domains of ThsB, hSARM1, and BdTIR (green). Tad4 forms molecular interactions with residues conserved among the TIR domains, both within and outside the active site pocket. A conserved beta strand, which interlaces with the C terminus of Tad4, is highlighted in each of the TIR-domain proteins.

(E) Predicted complexes formed between Acb3 (red) and the CD-NTase enzyme of *E. coli* KTE188 (green) and between Acb3 (red) and the hcGAS enzyme (green). Acb3 interacts with residues conserved between the two immune proteins, both within the nucleotidyltransferase active site pocket and in the putative ligand binding site.