

# **A lanthipeptide library used to identify a protein-protein interaction inhibitor**

Xiao Yang,<sup>1,§</sup> Katherine R. Lennard,<sup>2,§</sup> Chang He,<sup>1</sup> Mark C. Walker,<sup>1</sup> Andrew T. Ball,<sup>2</sup> Cyrielle Doigneaux,<sup>2</sup> Ali Tavassoli<sup>2,\*</sup> and Wilfred A. van der Donk<sup>1,\*</sup>

<sup>1</sup> Department of Chemistry and Howard Hughes Medical Institute, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801, USA.

<sup>2</sup> Department of Chemistry and Institute for Life Sciences, University of Southampton, Southampton, SO17 1BJ, United Kingdom.

§ These authors contributed equally to this study

\* To whom correspondence should be addressed:

Wilfred A. van der Donk, 600 S. Mathews Avenue, Urbana, Illinois 61801, United States; vddonk@illinois.edu; phone: (217) 244-5360; fax: (217) 244-8533.

Ali Tavassoli, Chemistry, University of Southampton, Southampton, SO17 1BJ, United Kingdom. a.tavassoli@soton.ac.uk; phone: +44 2380 592395.

## Abstract

We describe the production and screening of a genetically encoded library of  $10^6$  lanthipeptides in *Escherichia coli* using the substrate tolerant lanthipeptide synthetase ProcM. This plasmid-encoded library was combined with a bacterial reverse two-hybrid system for the interaction of the HIV p6 protein with the UEV domain of the human TSG101 protein, a critical protein-protein interaction for HIV budding from infected cells. Using this approach, we identified an inhibitor of this interaction from the lanthipeptide library, whose activity was verified *in vitro* and in cell-based virus-like particle budding assays. Given the variety of lanthipeptide backbone scaffolds that may be produced with ProcM, this method may be used for the generation of genetically encoded libraries of natural product-like lanthipeptides containing significant structural diversity. Such libraries may be combined with any cell-based assay for the identification of lanthipeptides with new biological activities.

Macrocycles have proven to be an excellent scaffold for therapeutic agents; these molecules occupy unique chemical space, bridging the gap between small molecules and biologics, and have been successfully deployed against challenging drug targets such as protein-protein interactions (PPIs)<sup>1</sup>. Whilst natural products and their synthetic derivatives have historically been the main source for therapeutic macrocycles<sup>2</sup>, cyclic peptides are increasingly viewed as holding much potential in this area<sup>3-5</sup>. Macrocyclic peptides share many of the physical properties of natural product macrocycles, such as extended binding sites and limited conformational freedom, with the added advantage of being amenable to the creation of large and diverse genetically encoded libraries<sup>5-11</sup>.

Ribosomally-synthesized and post-translationally modified peptides (RiPPs) are a major group of natural products that are biosynthesized from a genetically-encoded precursor peptide generally containing an N-terminal leader sequence and a C-terminal core peptide<sup>12, 13</sup>. After various post-translational modifications on the core peptide catalyzed by the biosynthetic enzymes and leader peptide removal, the mature RiPP is produced (e.g. **Fig. 1a**). The biosynthetic enzymes of many RiPPs have

43 been shown to display high substrate tolerance<sup>14-25</sup>, and thus have the potential to be used for production  
44 of large libraries of non-natural macrocyclic peptides.

45 Lanthipeptides are a major group of RiPPs characterized by intramolecular thioether bridges  
46 (termed lanthionine or methyllanthionine) generated via two post-translational modification reactions<sup>26</sup>:  
47 dehydration of Ser/Thr residues followed by cyclization of Cys residues onto the dehydrated amino  
48 acids (**Fig. 1b**). Recent genome mining efforts lead to the discovery of a substrate-tolerant synthetase  
49 (ProcM) in *Prochlorococcus* that dehydrates and cyclizes up to 30 different linear precursor peptides  
50 encoded in the genome (designated as ProcAs; **Supplementary Fig. 1**). Remarkably these peptides are  
51 transformed into single polycyclic lanthipeptide products (prochlorosins) with highly diverse ring  
52 topologies<sup>27-29</sup>.

53 Macrocyclic peptides hold much promise for recognition of protein surfaces and inhibition of  
54 PPIs because of their ability to act as structural mimics of native ligands<sup>1,30</sup>. In addition to restricted  
55 conformational freedom, cyclization also offers increased stability against cellular catabolism compared  
56 to linear peptides. Among the multitude of methods of generating cyclic peptides<sup>31</sup>, genetic approaches  
57 offer advantages of large library size and the possibility to couple peptide generation with cell-based  
58 screening. Bicyclic peptides have also shown much promise, but other than disulfide crosslinked  
59 peptides, they have thus far been limited to peptides generated by chemical methods<sup>6,9,32</sup>. In this study,  
60 we used enzymatic methods to generate a bicyclic peptide library in *Escherichia coli*. The potential  
61 utility of this approach is demonstrated by coupling lanthipeptide library generation to a bacterial  
62 reverse two-hybrid system (RTHS) for the PPI between the HIV p6 protein and the ubiquitin E2 variant  
63 (UEV) domain of the human TSG101 protein<sup>33</sup>. This PPI is a potential target for antiviral therapy as it  
64 has been shown to be essential for the budding of HIV from infected cells<sup>34-36</sup>.

## 65 66 **RESULTS**

### 67 **Generation of a lanthipeptide library in *E. coli***

68 Seeking to take advantage of the remarkable substrate tolerance of ProcM, as well as technology  
69 that allows successful production of lanthipeptides in *E. coli*<sup>37</sup>, we tested the use of ProcM for the

70 generation of non-natural lanthipeptide libraries. Since the highly conserved prochlorosin leader peptide  
71 serves as recognition motif for ProcM to carry out dehydration and cyclization on the diverse core  
72 peptides<sup>27</sup>, the library of substrates all contained the same leader peptide. The core peptides attached to  
73 the leader peptide were diversified using a degenerate oligonucleotide with nucleotide diversity  
74 introduced in predetermined positions during oligonucleotide synthesis. For proof-of-concept studies,  
75 we selected one of the 30 naturally encoded substrates, ProcA2.8<sup>27</sup>. ProcM converts ProcA2.8 into a  
76 product containing two non-overlapping, 7-residue thioether-crosslinked rings (**Fig. 1a**)<sup>28</sup>. We first tested  
77 the robustness of the ProcM catalyst for faithful installation of the two-ring scaffold when the ten amino  
78 acids within the rings were randomized (**Fig. 1c**). We set several criteria with respect to the randomized  
79 sequence. First, we wanted to avoid incorporation of Ser, Thr, and Cys residues as these might disturb  
80 the original ring scaffold by forming additional or alternative rings. Second, we wanted to avoid  
81 incorporation of stop codons. Third, we preferred minimal bias among the amino acid variants. And  
82 lastly, we felt that incorporation of negative charge would be more important than positive charge given  
83 the sequence of the p6 peptide that binds to UEV (PEPTAPPEE). The NWY codon usage fulfills these  
84 criteria as it introduces two codons each for the amino acid residues Asp, Phe, His, Ile, Leu, Asn, Val,  
85 and Tyr. For ProcA2.8, a library of  $1.07 \times 10^9$  members is theoretically possible by randomizing 10  
86 positions using the NWY codons that can introduce one of eight amino acids at each position (i.e.  $8^{10}$ ).

87 In order to test ProcM's catalytic fidelity on these non-native substrates, the library was first  
88 constructed in a pRSFDuet vector for co-expression of ProcM with individual members of the  
89 randomized N-terminally hexahistidine-tagged substrate library. The library was constructed starting  
90 from PCR amplification using degenerate primers that cover the  $1.07 \times 10^9$  amino acid diversity based  
91 on the NWY codons. After digestion and ligation, the plasmid library isolated from transformed *E. coli*  
92 cells had a size of  $10^6$  as determined by the transformation efficiency, which currently limits the size of  
93 the library that can be functionally analyzed with our method. Deep-sequencing was used to assess the  
94 diversity in the pRSF-Duet-derived plasmid library. As anticipated (based on the observed  
95 transformation efficiency) the size of the library was around  $10^6$  as 1,075,925 unique peptide sequences  
96 were obtained out of 2,008,108 total reads. Unique peptides that were sequenced 10 or fewer times

account for 99.7% of the unique peptides and 97.3 % of the total library, with a relatively small number of sequences highly represented (**Supplementary Fig. 2**). As desired, all eight amino acids encoded by the NWY codons were nearly evenly encoded at all ten randomized positions (**Supplementary Fig. 3**). Thirty three clones were randomly chosen and the encoded peptides overexpressed with ProcM in *E. coli*. In the absence of a leader peptide protease and transporter these peptides remain in the cytoplasm<sup>37</sup>, and hence the peptides were purified by Ni<sup>2+</sup> affinity chromatography from lysed cells. The extent of dehydration for each core peptide was determined by matrix-assisted laser-desorption ionization time-of-flight (MALDI-TOF) mass spectrometry (MS) after removing most of the leader peptide by proteolysis with endoproteinase Glu-C. All 33 randomly chosen peptides were dehydrated twice (**Supplementary Fig. 4**). To interrogate the cyclization state, the core peptides were then treated with the thiol-alkylating reagents iodoacetamide (IAA) or N-ethylmaleimide (NEM) that would report on any non-cyclized Cys residues, and the ring topology was determined by tandem mass spectrometry. Remarkably, ProcM catalyzed full cyclization of all 33 samples (**Supplementary Table 1, Supplementary Fig. 4 and 5**). Hence, based on the results of these 33 randomly chosen clones, ProcM produces a high quality library of bicyclic peptides containing two non-overlapping lanthionine rings in *E. coli*.

### Screening the library for inhibitors of a PPI

We next assessed the feasibility of coupling the above lanthipeptide library with a cell-based selection assay to enable the identification of PPI inhibitors. A bacterial RTHS was used that couples the successful disruption of the p6/UEV PPI (**Fig. 2a**)<sup>38</sup> to the transcription and expression of reporter genes critical for cell survival on selection medium (**Fig. 2b**)<sup>33</sup>. The p6/UEV RTHS was previously used to screen a library of  $3.2 \times 10^6$  monocyclic peptides, generated via split-intein circular ligation of peptides and proteins (SICLOPPS); *cyclo*-SGWIYWNV was identified in this screen and shown to inhibit the formation of virus-like particles with an IC<sub>50</sub> of 7  $\mu$ M in a cell-based viral budding assay<sup>33</sup>. Given that the RTHS uses a PTac promoter to induce expression of the p6 and UEV proteins upon IPTG treatment, a pACYCDuet vector was engineered to contain two arabinose promoters to drive the expression of the library of substrates and ProcM (pARDuet; **Supplementary Fig. 6**) and the co-expression conditions

were optimized (**Supplementary Fig. 7**). A pARDuet library encoding lanthipeptides based on ProcA2.8 was generated as before but without the His<sub>6</sub>-tag, and used to transform the p6/UEV RTHS. In bacteria containing a plasmid encoding a p6/UEV PPI inhibitor, disruption of the p6/UEV PPI would also disrupt the P22/434 repressor complex to enable transcription of the kanamycin resistance and *HIS3* genes (**Fig. 2b** and **Supplementary Fig. 8**), leading to a growth advantage on selective media containing kanamycin and lacking histidine. In contrast, cells that contain a lanthipeptide that does not disrupt the p6/UEV interaction will not survive or grow with a strong growth defect on selective media (**Supplementary Fig. 8a**). The p6/UEV RTHS cells transformed with the plasmid library were plated onto selective media containing IPTG and arabinose, and incubated at 33 °C until colonies were observed. Eighty colonies exhibited a growth advantage and were picked for further assessment. Individual colonies were subjected to secondary screening by drop spotting for IPTG-dependent growth inhibition and restoration of growth with arabinose. Colonies with the expected phenotype (**Fig. 2c**) were subjected to a tertiary screen by isolating the plasmid and re-transforming the p6/UEV RTHS cells with the plasmid. This procedure would remove any false positives with chromosomal mutations that conferred the growth advantage. The substrate specificity of the hits was also assessed using another RTHS strain that was identical to the p6/UEV strain, except that p6 and UEV were replaced with an unrelated PPI (human CMG2 interacting with *B. anthracis* PA)<sup>39</sup>; peptides showing inhibition of this unrelated PPI were discarded as they were either non-specific inhibitors, or targeted other components of the RTHS (**Supplementary Fig. 8b**). One peptide (named XY3-3 (**1**), **Fig. 2d**) remained after secondary and tertiary screening of the 80 initial hits. Analysis of our deep-sequencing data of the plasmid library used in the p6/UEV RTHS cells revealed that nucleotide sequences encoding XY3-3 were present fifteen times in the sequenced library that contained 8,673,291 unique peptides (**Supplementary Fig. 9**). Thus, the representation of XY3-3 in the pARDuet screening library is about the same as in the pRSFDuet-based library used to test the fidelity of ProcM discussed above.

#### Assessing the activity of XY3-3 *in vitro*

XY3-3 was prepared in *E. coli* on a larger scale and purified to assess its activity *in vitro*. We initially established a sandwich enzyme-linked immunosorbent assay (ELISA) to monitor the p6/UEV

PPI. We first titrated p6 onto plate-bound UEV, and found an  $EC_{50}$  of  $948 \pm 21$  nM for the p6/UEV interaction in this assay (**Supplementary Fig. 10a**). We next titrated XY3-3 onto plate-bound UEV, followed by p6, and observed disruption of the p6/UEV interaction by XY3-3 with an  $IC_{50}$  of  $3.6 \pm 0.3$   $\mu$ M (**Fig. 3a**). For comparison, a previously reported cyclic peptide inhibitor of this PPI (CP11, *cyclo*-SGWIYWNV)<sup>33</sup> was also used in this assay, and disrupted the p6/UEV interaction with an  $IC_{50}$  of  $30.2 \pm 2.2$   $\mu$ M (**Supplementary Fig. 10b**). The contribution of the leader peptide to the activity of XY3-3 was assessed via two independent approaches. First, we determined the activity of the parent ProcA2.8 scaffold in the above assay; as this molecule contains the same leader peptide as XY3-3, ProcA2.8 would also disrupt the interaction if the observed p6/UEV inhibitory activity was a consequence of the leader peptide. However, neither the cyclic nor linear forms of ProcA2.8 inhibited the p6/UEV PPI (**Supplementary Fig. 11**). Second, we used the Lys-C and Glu-C endoproteases to remove portions of the leader peptide from XY3-3; Lys-C cleaves ~50% of the leader sequence to generate peptide **2**, whereas Glu-C removes all but five residues of the leader sequence to afford peptide **3** (**Supplementary Fig. 12**). Cleavage of the leader sequence increased the potency of the lanthipeptide against the p6/UEV PPI; Peptide **2** disrupted the p6/UEV interaction with an  $IC_{50}$  of  $2.7 \pm 0.6$   $\mu$ M (**Fig. 3b**), while peptide **3** disrupted this PPI with an  $IC_{50}$  of  $1.9 \pm 0.8$   $\mu$ M (**Fig. 3c**). The necessity of XY3-3's lanthipeptide backbone for its activity was assessed using linear variants of XY3-3 (full-length and Glu-C-cleaved) (**Supplementary Fig. 12**). Neither molecule was able to disrupt the p6/UEV PPI by ELISA (**Supplementary Fig. 13**). It may therefore be concluded that the lanthipeptide core of XY3-3 (rather than its leader sequence) is responsible for disruption of the p6/UEV PPI, and that the cyclic thioether backbone is required for its activity.

In order to determine whether XY3-3 binds p6 or UEV, we generated an N-terminal fluorescein-labeled derivative of XY3-3 (**Supplementary Fig. 14**). The ability of p6 or UEV to pull down this molecule was assessed by immobilizing either protein onto  $Ni^{2+}$ -coated 96-well plates, and washing the fluorescent derivative of XY3-3 over these wells. We observed a 4-fold increase in fluorescence in wells containing UEV, with no change over background in wells containing p6 (**Fig. 3d**), indicating that XY3-3 binds to UEV. The binding affinity for UEV was determined using microscale thermophoresis (MST).

179 XY3-3 bound to UEV with a  $K_d$  of  $16.0 \pm 1.1 \mu\text{M}$  (**Fig. 3e**), while peptide **3** bound UEV with a  $K_d$  of  $4.0$   
 180  $\mu\text{M} \pm 3.3 \mu\text{M}$  (**Fig. 3f**). Thus, as with the ELISA data, removal of the leader sequence of XY3-3 resulted  
 181 in a more potent molecule, indicating that the lanthipeptide core of this molecule is responsible for its  
 182 activity. The binding constants obtained by MST were corroborated using a fluorescence polarization  
 183 (FP) assay; a  $K_d$  of  $5.5 \pm 0.5 \mu\text{M}$  was measured for the binding of N-terminal FITC-labeled peptide **3** to  
 184 UEV (**Fig. 3g**). For comparison, we also determined the affinity of UEV for the nonapeptide from p6  
 185 (PEPTAPPEE) that contains the PTAP motif that is essential for the binding of p6 to UEV (**Fig. 2a**)<sup>40</sup>. A  
 186 fluorescently-labeled PTAP nonapeptide bound to UEV with a  $K_d$  of  $16.6 \pm 0.2 \mu\text{M}$  (**Fig. 3h**). This value  
 187 is within the range of previously reported values for this interaction, measured by various methods ( $3$   
 188  $\mu\text{M}$ <sup>41</sup>,  $27 \mu\text{M}$ <sup>42</sup>, and  $50 \mu\text{M}$ <sup>40</sup>). Thus, peptide **3** binds to UEV with ~3-fold better affinity than the PTAP  
 189 nonapeptide when measured with the same assay. We also conducted competition FP experiments to  
 190 assess whether FITC-labeled peptide **3** displaces the PTAP peptide; we did not observe displacement  
 191 from UEV of PTAP by peptide **3**, nor peptide **3** by PTAP (**Supplementary Fig. 15**). These results  
 192 indicated that PTAP and peptide **3** bind to non-overlapping regions of UEV. In contrast, peptide **3** was  
 193 displaced from UEV by increasing concentrations of full length p6 peptide in the FP assay (**Fig. 3i**).  
 194 Taken together, the FP data indicates that while peptide **3** does not bind to the PTAP binding pocket, it  
 195 competes with p6 for binding to UEV.

196 The contribution from each of the two lanthionine rings to the binding of XY3-3 to UEV was  
 197 next assessed. Two variants of peptide **3** were generated, with the amino acid sequence of each ring  
 198 inverted one at a time. Inverting the amino acid sequence of ring A of peptide **3** (from LHFFL to LFFHL)  
 199 caused the loss of UEV binding activity (**Supplementary Fig. 16**). Similarly, inverting the amino acid  
 200 sequence of ring B (from HVLDI to IDLVH) in peptide **3** resulted in a lanthipeptide that did not bind  
 201 UEV (**Supplementary Fig. 16**). These observations demonstrate the requirement of both rings A and B  
 202 of peptide **3** in binding to UEV.

## 203 204 **Assessing the activity of XY3-3 in cells**

205 We next evaluated the potential of peptide **3** to disrupt viral budding in a cell-based assay<sup>33</sup>. This



peptide did not affect the GAG-mediated budding of virus-like particles, which we attributed to lack of cell permeability. We therefore designed an analog of XY3-3 in which a Cys replaced the N-terminal Gly (**Supplementary Fig. 12**). Co-expression of this peptide with ProcM in *E. coli* resulted in a fully dehydrated and cyclized product and tandem ESI-MS showed that the same ring topology was formed as the parent peptide (**Supplementary Fig. 17**). Leader peptide removal with endoproteinase Glu-C produced an analogue of peptide **3** with an N-terminal Cys, that was conjugated to a Cys-containing variant of the cell-permeable Tat peptide via a disulfide bond (XY3-3-Tat, **Supplementary Fig. 12**)<sup>33</sup>. Previous studies have shown that loss of TSG101 function is lethal in cells and *in vivo*<sup>43, 44</sup>. We therefore initially assessed the cytotoxicity of the PPI inhibitor and found XY3-3-Tat to be toxic to HEK293T cells at doses of 1  $\mu$ M and above (**Fig. 4a**). However, the compound was not toxic at 100 nM or 500 nM, and therefore these doses were used in the subsequent experiments. The effect of XY3-3-Tat on GAG-mediated viral budding was assessed in an established cell-based assay that utilizes GFP-tagged GAG polyprotein, and measures the formation of viral-like particles (VLPs) in the supernatant<sup>33, 45</sup>. XY3-3-Tat (100 nM) showed ~65% inhibition of viral budding from HEK293T cells in this assay, whereas no effect was observed when using 25  $\mu$ M Tat alone (**Fig. 4b**). The UEV domain of TSG101 also mediates degradation of Epidermal Growth Factor Receptor (EGFR). The binding of Epidermal Growth Factor (EGF) to its receptor (EGFR) triggers EGFR degradation, a process mediated by the interaction between TSG101 and hepatocyte growth factor-regulated tyrosine kinase substrate (HRS)<sup>46</sup>. This interaction involves a PSAP motif on HRS that engages TSG101 via the same binding site as the p6 PTAP motif<sup>40</sup>. As such, a compound that disrupts the binding of p6 to UEV is likely to also interfere with HRS binding to TSG101 and therefore inhibit EGFR degradation. HeLa cells were treated with 500 nM Tat as a control and cultured either in the presence, or absence of EGF. As expected, we observed a reduction in EGFR levels in cells cultured with EGF (lane 2 vs lane 1, **Fig. 4c**). We next assessed the effect of XY3-3-Tat on this process by repeating the above experiment with cells treated with 500 nM XY3-3-Tat and cultured either in the presence or absence of EGF. Cells treated with both XY3-3-Tat and EGF contained higher levels of EGFR than those treated with Tat and EGF (lane 4 v lane 2, **Fig. 4c**), with EGFR levels approaching that of cells treated with Tat alone (lane 4 vs lane 1, **Fig. 4c**). Cells treated with XY3-3-Tat

only (no EGF) showed significantly elevated levels of EGFR compared to Tat-treated equivalents (lane 3 v lane 1, **Fig. 4c**). Together, these data demonstrate inhibition of EGFR degradation by XY3-3-Tat, in line with the proposed mechanism of action of XY3-3 binding to UEV and inhibiting its interaction with P(T/S)AP containing proteins.

## DISCUSSION

Our results demonstrate the feasibility of using a promiscuous lanthipeptide synthetase (ProcM) for the intracellular generation of lanthipeptide libraries of about a million members. While the tolerance of lanthipeptide synthetases has been previously demonstrated through Ala-scanning and saturation mutagenesis at single sites<sup>15, 17, 21</sup>, this study represents the first example of a library of variants of a natural lanthipeptide by simultaneous randomization of multiple positions. We illustrate this possibility by generating a library with a theoretical size of  $1.07 \times 10^9$  members by randomizing 10 amino acid positions within a bicyclic lanthipeptide scaffold. The library was screened for inhibitors of the p6/UEV PPI by transforming a previously reported RTHS for this PPI<sup>33</sup> with the library-encoding plasmids. At present, the transformation efficiency of the selection strain and the relatively large size of the plasmid limited the screening for inhibitors to  $\sim 10^6$  unique peptides. In future efforts, it may be possible to engineer a chromosomal copy of ProcM to decrease the plasmid size and increase the transformation efficiency. After secondary screening, a single inhibitor (XY3-3) was identified that selectively prevented the targeted PPI in the RTHS. The activity of XY3-3 was verified in a series of *in vitro* experiments. Removal of the leader sequence from XY3-3 improved its activity, demonstrating that the leader sequence plays no part in target engagement, and instead appears to hinder optimal binding. It should be noted that the sequence of XY3-3 is very different from that of the PTAP motif of p6 that is critical for its binding to UEV<sup>40</sup>. In fact, none of the amino acids in this nonapeptide (Pro, Glu, Thr, and Ala) are encoded by the NXY codons used in the randomized regions in our library. Given the central role played by this nonapeptide in the p6/UEV PPI<sup>40, 41</sup>, and considering the data from our competition FP assays, XY3-3 likely binds at a different position than the PTAP peptide, the “hotspot” for the p6/UEV PPI. We were unable to determine the precise site of binding by crystallography or NMR

spectroscopy, in part because of the relatively low solubility of the peptide. Nonetheless, our *in vitro* and cell-based data suggest that XY3-3 is a functional inhibitor of the p6/UEV PPI. These findings have implications for inhibitor discovery approaches that focus on a single, known hotspot on a PPI, and illustrates the advantages of unbiased, functional screening. The results with the inverted sequences of the A and B rings suggest that both rings are making contact. Compared to a previously identified monocyclic inhibitor,<sup>33</sup> the bicyclic XY3-3 has about a 10-fold improved affinity, which is modest compared to the expected increased affinity upon having two binding moieties<sup>47</sup>. It suggests that the sequence of XY3-3 is not optimal. Typical affinity maturation methods are not readily applied to RiPPs, but are under development to identify variants that take full advantage of two high affinity interactions.

Not unexpectedly, XY3-3 was not cell permeable, but conjugation of a Tat-tag enabled the activity of this molecule to be assessed in human cells. In line with the mode of action displayed *in vitro*, XY3-3-Tat inhibited virus-like particle formation in an HIV GAG-mediated, cell-based viral budding assay. The compound also prevented TSG101-mediated degradation of EGFR in HeLa cells, in line with our *in vitro* data showing XY3-3 acts by binding the UEV domain of TSG101. TSG101 function is critical for cell survival<sup>43, 44</sup> providing a likely explanation for the observed toxicity of XY3-3-Tat to cells at higher doses. However, toxicity resulting from off-target effects cannot be ruled out. Regardless, this study illustrates the successful generation and screening of a lanthipeptide library in a cell-based assay, with the activity of the identified hit being verified *in vitro* and in cells. Given the range of other scaffolds that can be generated from the 30 different peptide substrates of ProcM (**Supplementary Fig. 1**) and the high quality of the library generated from ProcA2.8 (including XY3-3 and the two inverted ring sequences 36 out of 36 examined peptides were fully dehydrated and cyclized), our studies suggest the possibility of generating lanthipeptide libraries with a wide range of structural and chemical diversity. In addition to its use with a RTHS, this lanthipeptide library may be interfaced with a variety of other cell-based assays for the identification of biologically active lanthipeptides. Whereas previous studies have used epitope grafting to bestow new biological activities onto existing RiPPs<sup>48, 49</sup> or installation of RiPP-type crosslinks in known biologically active peptides<sup>50</sup>, this study utilizes the RiPP biosynthetic machinery for the discovery of entirely new biological activities, further underscoring the enormous

287 possibilities of the enzymes that produce this class of natural products.

288

## 289 **Acknowledgments**

290 The authors thank Prof. Sven Eyckerman for pMET7-GAG-EGFP (via Addgene, plasmid # 80605),  
291 and Dr. Diego Gomez-Nicola for HEK239T cells. This work was supported by the National Institutes of  
292 Health (R37 GM 058822 to W.A.V.; F32 1 F32 GM0112284 to M.C.W.), Cancer Research UK (A20185  
293 to A.T.), and the Engineering and Physical Sciences Research Council and C4X Drug Discovery  
294 (EP/L505067/1, Ph.D. studentship for K.R.L. to A.T.). Data underlying these experiments are available  
295 from the University of Southampton data repository at <http://doi.org/10.5258/SOTON/D0133>.

296

## 297 **Author Contributions**

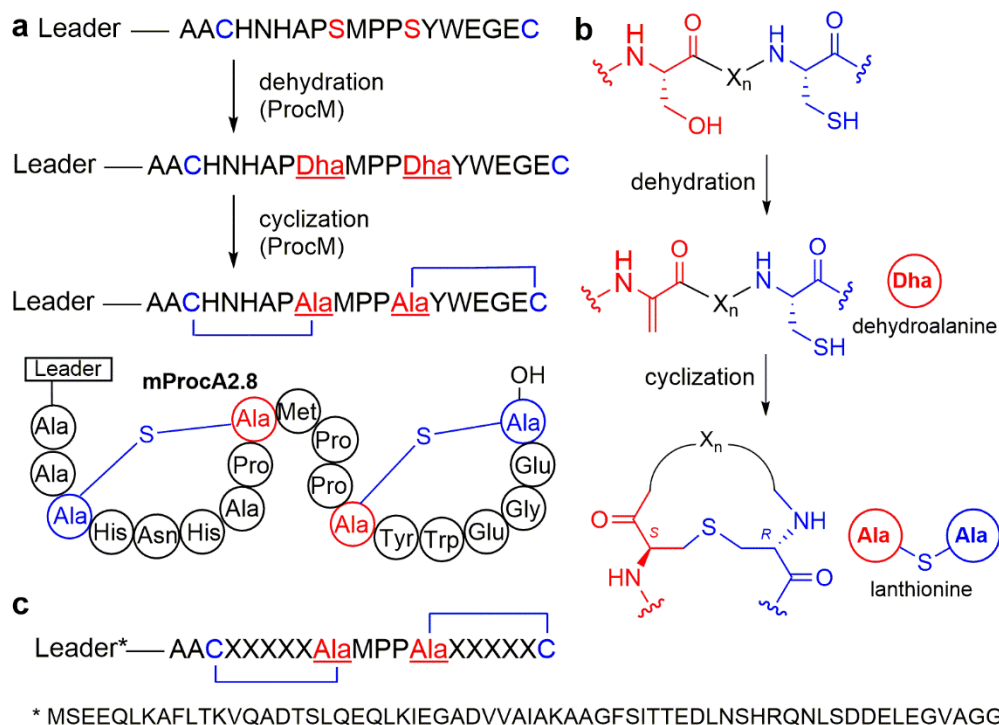
298 X.Y., A.T. and W.A.v.d.D. designed the study. X.Y., K.R.L., C.H., A.T.B. and C.D. performed all  
299 experiments. M.C.W. provided bioinformatics analysis of the plasmid library. All authors analyzed the  
300 data. X.Y., A.T., and W.A.v.d.D. wrote the manuscript.

301

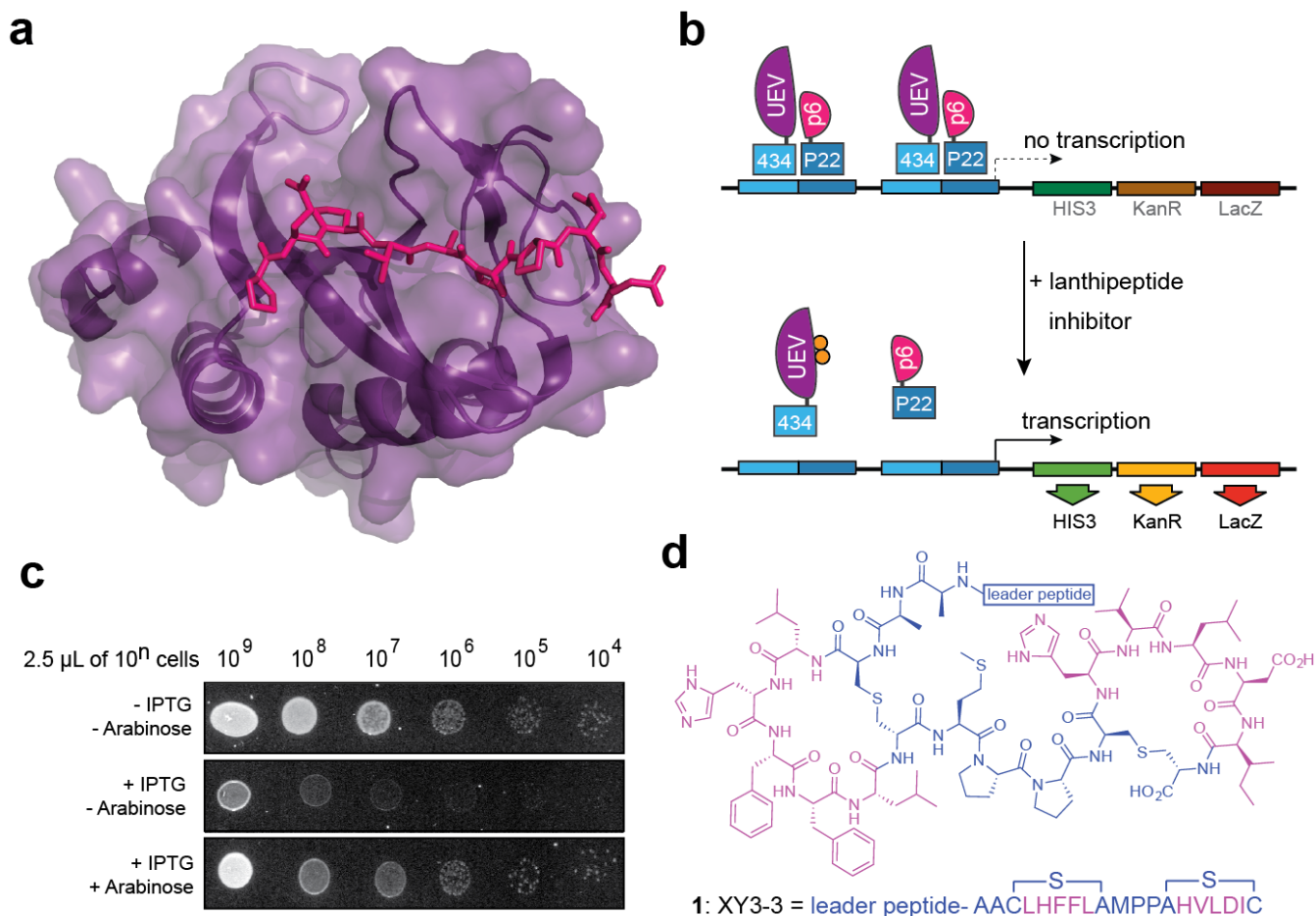
## 302 **Competing financial interests**

303 The authors declare no competing financial interests.

## Figures and legends

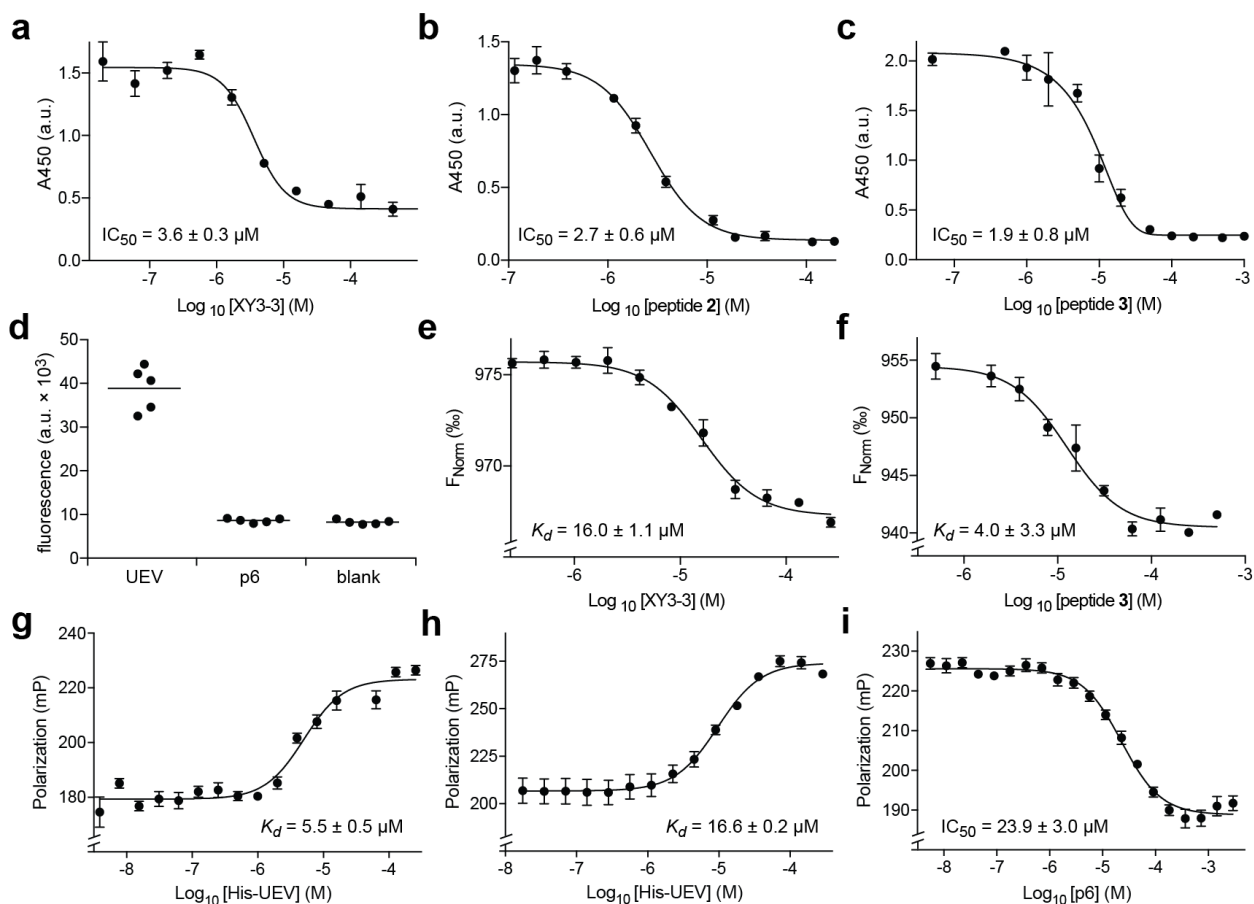


**Figure 1 | Representative illustration of lanthipeptide biosynthesis. (a)** Two Ser residues in the ribosomally synthesized linear precursor peptide ProcA2.8 are dehydrated by ProcM to generate two dehydroalanine (Dha) residues. The cyclization domain of ProcM then catalyzes the regioselective addition of two thiols of Cys residues to the Dha residues to generate modified ProcA2.8 (mProcA2.8). **(b)** Chemical structures showing the products of the dehydration and cyclization processes. **(c)** Generic structure demonstrating the randomization of the residues within the two rings of mProcA2.8 (X = D, F, H, I, L, N, V, or Y). The sequence of the leader peptide is also depicted. In all panels, structures derived from Ser are in red and from Cys in blue.



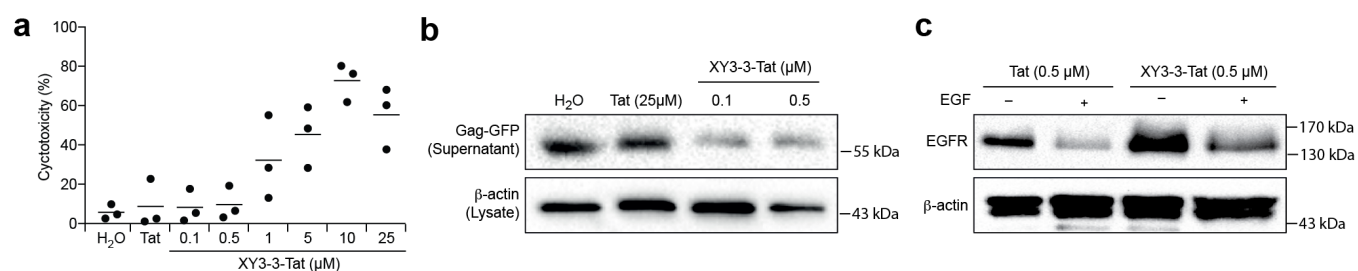
**Figure 2 | Identifying a lanthipeptide inhibitor of the p6/UEV PPI.** (a) Crystal structure of the PTAP peptide (PEPTAPPEE; pink) bound to UEV (purple); PDB ID = 3OBU<sup>40</sup>. (b) (Top) IPTG-induced expression of the 434-UEV and P22-p6 fusion proteins results in a functional repressor that inhibits expression of the reporter genes; HIS3 and KanR are required for survival on selective media. (Bottom) A lanthipeptide inhibitor of the p6/UEV interaction relieves the repression of the reporter genes and a growth advantage on selective media. (c) Drop spotting 10-fold serial dilutions of the p6/UEV RTHS transformed with the plasmid encoding XY3-3 in the absence of IPTG and arabinose (top row); expression of neither the repressors nor XY3-3 is induced, and full growth is observed. Upon adding IPTG (30  $\mu$ M; middle row), a functional repressor prevents expression of His3/KanR, resulting in inhibited cell growth. Upon addition of IPTG (30  $\mu$ M) and arabinose (65  $\mu$ M; bottom row), the repressor and XY3-3 are produced; XY3-3 disrupts the p6/UEV interaction, preventing formation of a functional

327 repressor and providing a growth advantage. All plates were incubated at 37 °C for 58 h. **(d)** Structure of  
328 XY3-3; residues in set positions are in blue, while randomized positions are in pink.  
329



**Figure 3 | Assessing the activity of XY3-3 *in vitro*.** (a-c) ELISA data showing that (a) XY3-3 disrupts the p6/UEV interaction with an IC<sub>50</sub> of  $3.6 \pm 0.3 \mu\text{M}$ , (b) Peptide 2 disrupts the p6/UEV interaction with an IC<sub>50</sub> of  $2.7 \pm 0.6 \mu\text{M}$ , and (c) Peptide 3 disrupts the p6/UEV interaction with an IC<sub>50</sub> of  $1.9 \pm 0.8 \mu\text{M}$ . (d) UEV selectively pulls down a fluorescent-labeled XY3-3 whereas p6 does not, indicating that XY3-3 binds to UEV. Line shown is mean, n=5. (e-f) MST analysis reveals that (e) XY3-3 binds to UEV with a  $K_d$  of  $16.0 \pm 1.1 \mu\text{M}$  and (f) Peptide 3 binds to UEV with a  $K_d$  of  $4.0 \pm 3.3 \mu\text{M}$ . (g-i) FP analysis reveals that (g) a FITC-tagged derivative of peptide 3 binds to UEV with a  $K_d$  of  $5.5 \pm 0.5 \mu\text{M}$ , (h) a fluorescein-tagged PTAP nonapeptide binds to UEV with a  $K_d$  of  $16.6 \pm 0.2 \mu\text{M}$ , and (i) a competition FP assay demonstrates that full length p6 displaces peptide 3 from UEV with an IC<sub>50</sub> of  $23.9 \mu\text{M}$ . Errors on  $K_d$  and IC<sub>50</sub> values are the standard error of the mean (S.E.M.) given by regression analysis from three independent experiments.





**Figure 4. Cellular activity of XY3-3-Tat.** (a) Cytotoxicity assay of XY3-3-Tat in HEK293T cells shows the compound is toxic to cells as doses above 1 μM. Line shown is mean, n =3. (b) Virus-like particle budding assay in HEK293T cells shows that a 100 nM dose of XY3-3-Tat inhibits virus-like particle production by ~65%. For uncropped blots see Supplementary Figure 18. (c) EGFR downregulation assay in HeLa cells shows XY3-3-Tat inhibits degradation of EGFR, in line with its mode of action of binding to UEV. For uncropped blots see Supplementary Figure 19.

## On line Methods

### General methods

For lanthipeptide residue labelling, positive residue numbers are used for amino acids in the core peptide counting forwards from the (putative) leader peptide cleavage site. Negative residue numbers are used for amino acids in the leader peptide counting backwards from the leader peptide cleavage site. For instance, in the ProcA2.8 precursor peptide, the C-terminal Gly residue of the leader sequence is -1, and when substituted by Lys the mutant is denoted as G-1K. Polymerase chain reaction (PCR) amplifications were carried out using an automated thermocycler (C1000, BioRAD). DNA sequencing was performed using appropriate primers by ACGT Inc. LC-ESI-Q/TOF MS analyses were conducted using a Synapt MS system equipped with an Acquity Ultra Performance Liquid Chromatography (UPLC) system (Waters). Matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF MS) analyses were conducted at the Mass Spectrometry Facility at UIUC using an UltrafleXtreme spectrometer (Bruker Daltonics). For MALDI-TOF MS analysis, samples were desalted using ZipTipC18 (Millipore), and spotted onto a MALDI target plate with a matrix solution containing 35 mg/mL 2,5-dihydroxybenzoic acid (DHB) in 3:2 MeCN/H<sub>2</sub>O with 0.1% TFA. Peptides were desalted by a C4 solid-phase extraction (SPE) column and further purified by preparative reversed-phase high performance liquid chromatography (RP-HPLC) on a Delta 600 instrument (Waters) equipped with a Phenomenex C18 column at a flow rate of 8 mL/min. For RP-HPLC, solvent A was 0.1% TFA in H<sub>2</sub>O and solvent B was 4:1 MeCN/H<sub>2</sub>O containing 0.086% TFA. An elution gradient from 0% solvent B to 74% solvent B over 40 min was used unless specified otherwise.

### Materials

All chemicals used were purchased from Sigma Aldrich or Fisher Scientific unless noted otherwise. Endoproteinases Glu-C and Lys-C were purchased from Roche Applied Science. Oligonucleotide primers used for molecular cloning were purchased from Integrated DNA Technologies. The pBAD/HisA vector was purchased from Invitrogen. Electrocompetent *E. coli* DH5 $\alpha$  cells, Phusion

High-Fidelity DNA polymerase, Taq ligase, dNTP solutions, T4 DNA ligase and all restriction endonucleases were purchased from New England Biolabs. Gel extraction, plasmid miniprep, and PCR purification kits were purchased from QIAGEN. Protein Calibration Standard I and Peptide Calibration Standard II for MALDI-TOF MS were purchased from Bruker. *E. coli* DH5 $\alpha$  was used as host for cloning and plasmid propagation, and *E. coli* BL21 (DE3) was used as a host for overexpression.

### Statistical analysis

Data acquired from enzyme-linked immunosorbent assay (ELISA), Microscale Thermophoresis (MST) and Fluorescence Polarization (FP) experiments were processed by Prism 6 (Graphpad) unless otherwise indicated. EC<sub>50</sub>, IC<sub>50</sub> and dissociation constants were calculated using a non-linear regression model from three independent replicates. Errors are the standard error of the mean (S.E.M.) given by regression analysis. Data shown are mean  $\pm$  S.E.M. from the three replicates.

### Construction of lanthipeptide library in pRSFDuet-1 vector

Library primers were designed that added an EcoRI restriction site to the 5' end (Lib\_EcoRI\_FP, Supplementary Table 2) and a NotI restriction site to the 3' end of the construct. The reverse primer (Lib2.8NWY\_NotI\_RP, Supplementary Table 2) containing degenerate codons was synthesized by IDT using hand-mixed bases and purified by PAGE. The PCR template was constructed by inserting the procA2.8 leader sequence (amplified with primer Lib\_EcoRI\_FP and 2.8L\_NotI\_RP, Supplementary Table 2) into MCS-1 of pRSFDuet-1 plasmid between EcoRI and NotI restriction sites. This plasmid was used as template for PCR generation and amplification of the library. PCR was carried out using Phusion polymerase by 24 cycles of denaturing (94 °C for 30 s), annealing (67 °C for 30 s) and extending (72 °C for 20 s). The PCR products (appearing as a strong band between 250-300 bp) were gel-purified, digested with EcoRI and NotI restriction enzymes and purified using a Qiagen PCR purification kit.

The pRSFDuet-1 vector with *procM* inserted into MCS2 between NdeI and KpnI<sup>37</sup> was used for library vector construction, the vector was digested with the same restriction enzymes as the PCR

products and an additional enzyme SbfI to minimize vector self-ligation, dephosphorylated with alkaline phosphatase and gel-purified. The digested vector (50 ng) was ligated with the insert DNA by T4 ligase at a molecular ratio of 1/6 (vector/insert), and the reactions were incubated at 15 °C overnight. The DNA product was desalted by butanol precipitation and used to transform electro-competent *E. coli* DH5 $\alpha$  cells (New England Biolabs). After recovery in 1 mL of SOC medium (for each transformation) for one hour, 5  $\mu$ L of the combined culture was plated as serial dilutions on LB-agar containing kanamycin (50 mg/L) to calculate the library size. The rest of the recovery culture was diluted in 10-fold volume of LB containing kanamycin (50 mg/L) and the library plasmids were isolated using a Qiagen Miniprep Kit after overnight incubation at 37 °C. Five ligation reactions described above were combined in order to scale up the library size. After transformation, the library size was approximately 10<sup>6</sup>, calculated by dilution plating.

#### **Construction of a library vector compatible with the RTHS strain**

The dual-arabinose promoter vector pARDuet was constructed based on the pACYCDuet-1 backbone. pBAD/HisA plasmid was used as template for PCR amplification of the *araC* and *P<sub>ara</sub>* region (primer *araC\_pBAD\_NarI\_FP* and *araC\_pBAD\_NcoI\_RP*, Supplementary Table 2). The PCR products were digested with NarI and NcoI and ligated into digested pACYCDuet-1 vector, which replaced the *lacI* with *araC* and the *T7* promoter in MCS-1 with *P<sub>ara</sub>*. Next, the *P<sub>ara</sub>* sequence was PCR-amplified from pBAD/HisA (primer *Para\_pBADHis\_BsrGI\_FP* and *Para\_pBADHis\_NdeI\_FP*, Supplementary Table 2), the PCR product was digested with BsrGI and NdeI, and cloned into the vector obtained in the first step, which replaced the *T7* promoter in MCS-2 with *P<sub>ara</sub>*. Finally, the *rrnB* terminator sequence from pBAD/HisA was PCR-amplified (primer *araTer\_XhoI\_FP* and *araTer\_Bsu36I\_RP*, Supplementary Table 2), digested with XhoI and Bsu36I and cloned into the resulting vector, giving the empty library vector named pARDuet (Supplementary Fig. 6, Supplementary Table 3).

#### **Construction of lanthipeptide libraries using the pARDuet vector**

The libraries for genetic selection do not require the N-terminal His<sub>6</sub>-tag. Therefore, a new forward library primer was designed that added an NcoI restriction site (closest to the start codon to eliminate unnecessary N-terminal residues) at the 5' end (primer Lib\_NcoI\_FP, Supplementary Table 2). The reverse library primer remained the same as that used for the library in pRSFDuet. *procM* was cloned into the MCS-2 of pARDuet between NdeI and KpnI. The two NcoI cut sites inside the *procM* gene were deleted by silent mutations via site-directed mutagenesis (primer pMQC\_1t2577c\_FP, pMQC\_1t2577c\_RP, pMQC\_2c2772t\_FP and pMQC\_2c2772t\_RP, Supplementary Table 2), resulting in library vector ProcM(t2577c/t2775c)\_pARDuet. The plasmid library with ProcA2.8 containing randomized core peptides was constructed as described above, except that NcoI and NotI were used as the restriction sites and chloramphenicol was used instead of kanamycin as selective marker. After PCR, digestion with restriction endonucleases, ligation, and transformation of *E. coli* DH5 $\alpha$  cells (NEB), the plasmid library was isolated and plasmids were eluted in water during plasmid preparation for maximum electroporation efficiency in the next step.

#### **Construction of pARDuet derivative for co-expression of His<sub>6</sub>-tagged peptides**

To test if the pARDuet vector could also produce a modified peptide library like the pRSFDuet vector, the gene encoding His<sub>6</sub>-tagged ProcA2.8<sup>37</sup> was amplified and the PCR product digested with NcoI and NotI and inserted into MCS-1 of the ProcM(t2577c/t2775c)\_pARDuet vector between the same restriction sites. This plasmid was tested using the expression conditions described below.

#### **Over-expression of modified library peptides in *E. coli***

In order to test the modification efficiency of the library peptides, *E. coli* BL21 (DE3) cells were transformed with the previously-mentioned pRSFDuet derivatives encoding His<sub>6</sub>-tagged substrates and ProcM, and colonies were randomly chosen from the transformation plate and inoculated into fresh medium for over-expression. The cell culture (50-100 mL LB) was induced with 0.25 mM IPTG at an OD<sub>600</sub> between 0.6-0.8, and incubated at 18 °C for 18-20 h. MALDI-TOF MS analysis demonstrated that fully modified peptides were produced (Supplementary Fig. 4 and 5). For determining the quality of

ProcM catalysis, in total 36 non-native ProcA derivatives were analyzed and all 36 were dehydrated and cyclized. With this sample size, at the 95% confidence level, the Clopper-Pearson exact lower confidence limit regarding the library quality is 90%<sup>51</sup>. In order to test the expression conditions for genetic selection, the co-expression plasmid containing genes encoding WT ProcA2.8 and ProcM<sup>37</sup> was then used at different temperatures in 0.5-2 L minimal medium A, which showed that expressed ProcA2.8 could be fully dehydrated at temperatures up to 30 °C (Supplementary Fig. 7a).

To test the overexpression and modification using the pARDuet vector, after induction with 6.5 mM arabinose (note this concentration is higher than that used for selection, see below, in order to obtain sufficient amounts of peptide for analysis), the cells were cultured in LB or minimal medium A containing 25 mg/L chloramphenicol at 18 °C in the *E. coli* RTHS selection strain or in *E. coli* BL21 (DE3) cells. MALDI-TOF MS analysis demonstrated that fully modified peptides were produced (Supplementary Fig. 7b and c).

### Sequencing of the plasmid library

Sequencing primers were designed to amplify the randomized portion of the pARDuet-derived plasmid library with Illumina adapter overhang nucleotide sequences per Illumina's protocol (Lib2.8NWYseq\_FP and Lib2.8NWYseq\_RP, Supplementary Table 2). PCR was carried out using Phusion polymerase by 15 cycles of denaturing (94 °C for 30 s), annealing (66 °C for 1 min) and extending (72 °C for 15 s). The PCR products were purified by Qiagen PCR purification kit and subjected to the following Index PCR. Nextera XT Index N701 and S517 (Nextera XT index kit FC-131-1001) were used to amplify the products of the first PCR using Phusion polymerase with 8 cycles of denaturing (94 °C for 30 s), annealing (66 °C for 1 min) and extending (72 °C for 30 s). The PCR products (appearing as a strong band at 200-250 bp) were gel-purified with Qiagen gel purification kit. DNA concentration was quantified with a Qubit® 2.0 fluorimeter (Invitrogen) using a Qubit HS dsDNA HS assay kit. Deep sequencing was performed in the Roy J. Carver Biotechnology Center (University of Illinois at Champaign-Urbana) on an Illumina HiSeq 2500.

The Illumina reads obtained were processed with in-house Python scripts (freely available

484 directly from the authors upon request). Reads containing the sequence GGAGCG, which encodes the  
485 last residue of the leader and the first residue of the core, were identified and the quality scores for those  
486 bases and the following 57 (the rest of the core peptide including the stop codon) were analysed. Any  
487 read not containing the sequence or with a base with a quality score less than 30 in the examined region  
488 were discarded. For paired end sequencing, the reverse reads were searched for the reverse complement  
489 of the above sequence and the same quality score cut off was applied to the preceding 57 bases. If any  
490 base in the examined region of the forward or reverse reads dropped below the quality cut off, or if the  
491 sequence of examined region was not identical in the forward and reverse reads, both reads were  
492 discarded. The remaining reads were translated and the number of times a given peptide sequence was  
493 represented in the library were tabulated to determine the diversity of the library.

494

#### 495 **Library selection in the UEV-P6 RTHS strain**

496 Each selection plate (150 mm diameter × 25 mm) contained 40 mL of M9 minimal media  
497 supplemented with 25 mg/L kanamycin, 2.5 mM 3-AT (3-amino-1,2,4-triazole), 30 mg/L carbenicillin,  
498 25 mg/L spectinomycin, 35 mg/L chloramphenicol, 1 mM MgSO<sub>4</sub>, 65 μM arabinose, 30 μM IPTG, 2%  
499 glycerol and 1.5% agar. A negative control plate was also made that omitted IPTG. UEV-p6 RTHS cells  
500 were cultured overnight in LB medium containing 30 mg/L carbenicillin and 25 mg/L spectinomycin, an  
501 aliquot of 1 mL cell culture was inoculated into 100 mL of the same medium and the cells were grown  
502 until the OD<sub>600</sub> reached 0.5. The cells were quickly chilled in an ice-water bath and incubated on ice for  
503 40 min before harvesting (3000 g, 15 min, 2 °C). The supernatant was discarded and the cell pellets  
504 were gently resuspended in 20 mL of ice-cold 10% glycerol and harvested again. The washing step was  
505 repeated two more times and the resulting electro-competent cell pellets were resuspended with a final  
506 total volume of 200-250 μL and split into four tubes. These freshly made competent cells were  
507 transformed (1 mm cuvette, 1.8 kV) with the library plasmids (4 × 25 ng), and 0.5 mL of SOC medium  
508 pre-warmed at 37 °C was immediately added into each transformation. After recovery for 1 h at 37 °C,  
509 the culture was centrifuged and the supernatant discarded. An aliquot of 0.5 mL of minimal medium A  
510 was added into each tube and the cell pellets were thoroughly resuspended by gentle pipetting. An

511 aliquot of 5  $\mu$ L of the combined culture was plated as serial dilutions on LB-agar containing  
512 chloramphenicol (35 mg/L) to calculate the library size. The rest of the culture was evenly spread onto  
513 the selection plates until all liquid was completely absorbed into the plates.

514 The selection plates were sealed with parafilm and incubated at 30-33  $^{\circ}$ C for 2-4 days. The  
515 growth was compared with cells on the control plate without IPTG, and a growth delay was observed on  
516 plates with IPTG during this time window. Visible colonies started to show up starting on the second  
517 day, and normally around the fourth day the background colonies started to grow due to IPTG  
518 degradation, antibiotic resistance, or mutations.

519 Eighty identified colonies that exhibited faster growth than the background colonies were  
520 inoculated into individual wells of a 96-well culture plate with 1 mL LB medium containing 30 mg/L  
521 carbenicillin, 25 mg/L spectinomycin and 35 mg/L chloramphenicol and the plate was incubated with  
522 constant shaking at 37  $^{\circ}$ C overnight. Then, using a multi-channel pipette, each culture was sequentially  
523 diluted (10-fold each time) into M9 minimal media.

524 Four sets of plates were made from identical selection medium as described previously, but with  
525 or without IPTG or arabinose. Then 2.5  $\mu$ L of each dilution culture (containing roughly  $10^6$ ,  $10^5$ ,  $10^4$  and  
526  $10^3$  cells) from each sample were drop-spotted onto each plate. The plates were sealed with parafilm and  
527 incubated at 30-33  $^{\circ}$ C for 2-3 days. Growth difference was monitored by comparing the samples on the  
528 four plates at each dilution level, and two colonies that exhibited both IPTG-dependent inhibition and  
529 arabinose-dependent growth advantage were selected. The UEV-p6 strain and another heterodimeric  
530 RTHS strain (in which UEV and p6 were replaced with an unrelated protein-protein interaction of  
531 CMG2 with anthrax toxin protective antigen, PA) were transformed with these plasmids to confirm the  
532 specificity of the observed arabinose-dependent growth advantage. Each culture was sequentially diluted  
533 and drop-spotted onto the same four sets of selection plates with or without IPTG or arabinose as  
534 described above. A single plasmid was active in the UEV-p6 RTHS with no effect in the CMG2/PA  
535 RTHS. Sequencing of the entire plasmid revealed the identity of the isolated hit (XY3-3, Supplementary  
536 Table 4) without any other mutations in the plasmid that could have conferred the growth advantage.

537



### **Overexpression of the XY3-3 peptide**

The pRSFDuet vector system was chosen for over-expression and characterization of the purified hit peptide. The gene was amplified by PCR from the pARDuet vector isolated from the hit colony described above and cloned into the MSC-1 of procM\_pRSFDuet<sup>37</sup> between the EcoRI and NotI sites. A thrombin cleavage site was also introduced between the His<sub>6</sub>-tag and the peptide sequence in order to remove the tag after overexpression (primer 2.8Thrombin\_EcoRI\_FP and Lib\_NotI\_RP, Supplementary Table 2).

### **Purification of modified His<sub>6</sub>-tagged ProcA2.8 derived peptides**

After harvesting, the cell pellet was resuspended at 0.1 g/mL in LanA Start Buffer (20 mM NaH<sub>2</sub>PO<sub>4</sub>, 500 mM NaCl, 0.5 mM imidazole, 20% glycerol, pH=7.5). The cell paste was subjected to sonication to lyse the cells. Insoluble cell debris was removed by centrifugation at 16,500 × g for 45 min. The supernatant was purified by immobilized metal affinity chromatography (IMAC) using 2-4 mL of His60 Ni Superflow Resin (Clontech). Following 1 h incubation at room temperature, the resin was washed with wash buffer (Start Buffer + 30 mM imidazole), and the peptide was eluted from the resin using elution buffer (start buffer + 4 M guanidine hydrochloride + 0.5 M imidazole). The elution fractions were desalted by C4 SPE column and purified on a C18 column using the solvent gradient described in the General Methods.

### **Expression and Purification of His<sub>6</sub>-UEV and GST-p6**

The gene encoding human UEV (ubiquitin E2 variant) was cloned into a pET28a vector using the NheI and XhoI restriction sites. The viral p6 gene was cloned into PGEX-2TK using the EcoRI and BamHI restriction sites. Both recombinant proteins were expressed in *Escherichia coli* BL21 (DE3) cells. Bacterial cultures were incubated at 37 °C with shaking until an OD<sub>600</sub> of between 0.5 and 0.7 was reached. Protein expression was induced with 0.1 mM IPTG overnight at 16 °C. After harvesting, the cell pellets were suspended at 1 g/mL in Buffer A (20 mM Tris-HCl, 100 mM NaCl, 10% glycerol, pH 7.4) supplemented with 1% (v/v) Triton X-100, Pierce Protease Inhibitor tablets (ThermoFisher

Scientific, 1 tablet per 20 mL) and lysozyme (100 µg/mL), and the cells were incubated on ice for 30 min. The lysate was sonicated (12 cycles of 10 s on followed by 10 s off) on ice to ensure complete lysis. Insoluble cell debris was pelleted via centrifugation at  $22,700 \times g$  for 45 min at 4 °C. The supernatant was filtered using 0.45 µm syringe filters. For His<sub>6</sub>-UEV, the supernatant was loaded onto a 1 mL HisTrap column (GE Healthcare) equilibrated in Buffer A coupled to an ÄKTA Prime fast protein liquid chromatography (FPLC) system (GE Healthcare). The protein was loaded onto the column at a flow rate of 1 mL/min. The column was washed with 30 mL of Buffer A and then with a gradient to 8% Buffer B (500 mM imidazole, 20 mM Tris-HCl, 100 mM NaCl, pH 7.4) in Buffer A at a flow rate of 1 mL/min, then the protein was eluted in 100% Buffer B. UV absorbance at 280 nm was monitored and fractions were collected and analysed by SDS-PAGE.

For GST-p6, the filtered supernatant after cell lysis was loaded onto a 1 mL GSTrap column (GE Healthcare) coupled to an ÄKTA Prime FPLC system (GE Healthcare) pre-equilibrated in Buffer A at a flow rate of 1 mL/min. The protein was loaded onto the column at a flow rate of 0.5 mL/min. The column was washed with 30 mL of Buffer A, and the GST-p6 was eluted with 10 mL GST Elution buffer (50 mM Tris-HCl, 150 mM NaCl, 15 mM reduced glutathione, pH 7.4) at a flow rate of 1 mL/min.

For both proteins, the fractions containing the desired protein were combined and dialysed overnight at 4 °C using SnakeSkin™ Dialysis Tubing (10K MWCO 22 mm; ThermoFisher Scientific) into Buffer A (1 L/ 2 mL protein). The buffer was exchanged for fresh buffer and allowed to dialyse for another 3 h. His<sub>6</sub>-UEV was obtained with a yield of 20 mg/L cell culture, and GST-p6 was obtained with a yield of 10 mg/L cell culture. GST-p6 was subjected to thrombin cleavage at 4 °C overnight and purified with an FPLC system with a Superdex 200 10/300 GL size exclusion column (GE Healthcare) using buffer A. Protein aliquots were flash-frozen in liquid nitrogen and stored at -80 °C.

588

### 589 **Endoproteinase digestion and purification of unmodified or modified peptides**

For peptides prepared by *E. coli* co-expression, the majority of the leader peptide needs to be removed in order to accurately characterize the dehydrations in the core peptide by MS. Library samples

were digested with Glu-C or Lys-C. XY3-3 (peptide **1**) was digested with thrombin to remove the His<sub>6</sub> tag, or with Lys-C, or Glu-C in order to obtain the cyclized core peptides with different leader length (peptide **2** and peptide **3**, respectively). The proteolysis reaction also contained 0.5 mM TCEP to reduce potential disulfide bonds in the core peptide. In contrast, when trying to remove the leader sequence of unmodified ProcA2.8 and linear unmodified XY3-3 by Glu-C, oxidative disulfide bond formation was performed in order to protect the Glu and Asp residues in the core peptide sequence from Glu-C cleavage. The linear peptide (150 µM) was treated with 5 mM each of reduced and oxidized glutathione in 50 mM HEPES buffer for 1 h at room temperature, before Glu-C was added. Purification of the digested sequence was carried out by RP-HPLC on an analytical Phenomenex C18 column, using the gradient described in the General Methods.

#### ***N*-Ethylmaleimide (NEM) or iodoacetamide (IAA) alkylation assay to test for cyclization**

An aliquot of the protease-digested peptide or full-length peptide solution was diluted into twice the volume of NEM alkylation buffer containing 500 mM HEPES, 3 mM NEM, 0.3 mM TCEP, pH 6.5, or IAA alkylation buffer containing 500 mM HEPES, 3 mM IAA, 0.3 mM TCEP, pH 8.0. The reaction was incubated at 37 °C for 30 min (for NEM alkylation) or at room temperature for 1-2 h (for IAA alkylation) in the dark, and analysed by MALDI-TOF MS. Species containing uncyclized free Cys residues that were alkylated were identified by a mass increase of 125 Da (for NEM) or 57 Da (for IAA) for each adduct.

#### **ELISA assay characterizing the disruption of UEV-p6 interaction by lanthipeptides**

To each well in a clear, nickel coated 96-well plate (Thermo Scientific), purified His<sub>6</sub>-UEV protein diluted in PBS (Thermo Scientific; 1 tablet dissolved in 100 mL deionized water; at a final concentration of 1 µM of His<sub>6</sub>-UEV; 100 µL per well), was added and incubated for 1 h at room temperature with rocking. The wells were then washed with PBS containing 0.05% TWEEN 20 (Sigma; 3 × 200 µL, 5 min incubation per wash). The wells were blocked for 1 h with 2% milk solution in deionized water (150 µL per well) at room temperature with rocking. Without washing the wells,

619 descending concentrations of lanthipeptide (serially diluted in 100% DMSO, 10  $\mu$ L) with 0.5  $\mu$ M of  
620 purified GST-tagged p6 protein in PBS (90  $\mu$ L) were added and the plate was incubated for 1 h at room  
621 temperature with rocking (10% final concentration of DMSO). The wells were then washed as before,  
622 and mouse anti-GST Ab-1 antibody (Neomarkers; 1:1000 made up in 2% milk in deionized water; 100  
623  $\mu$ L per well) was added and incubated for 1 h at room temperature with rocking. The wells were washed,  
624 then sheep anti-mouse IgG, peroxidase-linked whole antibody (GE Healthcare; 1:3000 made up in 2%  
625 milk in deionized water; 100  $\mu$ L per well) was added and the plate was incubated for 1 h at room  
626 temperature with rocking. The wells were washed and 1-Step™ Ultra TMB ELISA solution (Thermo  
627 Scientific; 100  $\mu$ L) was added to each well and the plate was incubated for 15 min at room temperature  
628 with rocking. To quench the reaction, 20 mM sulphuric acid (100  $\mu$ L) was added to each well and the  
629 absorbance at 450 nm measured using a Tecan Infinite® M200 PRO plate reader. Data was processed  
630 with Prism 6 (GraphPad); EC<sub>50</sub> and IC<sub>50</sub> values were determined using a non-linear regression model  
631 from three replicates.

632

### 633 **Fluorescein labelling of XY3-3**

634 For fluorescein labelling of XY3-3 using fluorescein-5-maleimide (FM), the pRSFDuet plasmid  
635 with *procA* in MCS1 and *procM* in MCS2 was altered such that a Cys residue was inserted on the N-  
636 terminus of the leader peptide after a thrombin cleavage site (LVPRGC, thrombin cleaves after the Arg  
637 residue (bolded) in the underlined recognition sequence). The peptide sequence was PCR-amplified  
638 using primers 2.8ThrombinCys\_EcoRI\_FP and Lib\_NotI\_RP (Supplementary Table 2), and cloned  
639 between the EcoRI and NotI restriction sites in MCS1 of the pRSFDuet vector containing *procM* in  
640 MCS2. The peptide was overexpressed in *E. coli* and purified as described previously. The additional  
641 Cys residue did not interrupt the modification of the core peptide, as two non-overlapping rings still  
642 formed (Supplementary Fig. 14). The peptide was then subjected to thrombin cleavage. Thrombin was  
643 used to remove the N-terminal His<sub>6</sub>-tag after pro-longed incubation time, and the resulting untagged  
644 Cys-XY3-3 peptide was purified by HPLC on a C4 column. The fluorescein labelling reaction contained  
645 20  $\mu$ M purified peptide, 50  $\mu$ M TCEP, 0.5 mM fluorescein-5-maleimide and 50 mM HEPES, pH 6.9.

The reaction was incubated at room temperature for 3-4 h in the dark and monitored by MALDI-TOF MS (Supplementary Fig. 14). After C4 SPE desalting, the labelled peptide was further purified by HPLC on a C4 column as described under General Methods.

### **Fluorescent peptide pull-down**

His-tagged p6 and UEV (5  $\mu$ M protein in 45  $\mu$ L buffer A) were immobilised into separate wells of a Ni<sup>2+</sup>-coated 96-well plate. Fluorescein-tagged XY3-3 (5  $\mu$ L of a 100  $\mu$ M solution) was added to each well and incubated with shaking at room temperature for 10 minutes. The wells were washed with buffer a (3  $\times$  200  $\mu$ L). The fluorescence of each well was measured using a BMG Labtech ClarioStar plate reader.

### **FITC labelling of peptide 3 and PTAP**

HPLC-purified peptide **3** was dissolved in PBS buffer (pH=8.0) to a final concentration of 50  $\mu$ M. FITC (Thermo Fisher) was dissolved in DMSO (10 mg/mL) and added slowly to the peptide solution to a final concentration of 250  $\mu$ M. The reaction was incubated for 18 h at 25  $^{\circ}$ C. After centrifugation at 16,100  $\times$  g for 10 min, the supernatant was purified on an Agilent 1260 Infinity HPLC system with a Phenomenex Luna C18 column (250 mm  $\times$  4.6 mm, 10  $\mu$ m particle size, 100  $\text{\AA}$  pore size). Solvent A was 0.1% TFA in H<sub>2</sub>O and solvent B was 80% MeCN/H<sub>2</sub>O containing 0.086% TFA. An elution gradient from 0% solvent B to 100% solvent B over 45 min at 1 mL/min was used and FITC labeled XY3-3 core peptide eluted at 66-67% solvent B. PTAP (GGPEPTAPPEE) peptide containing two additional N-terminal Gly residues (synthesized by United Biosynthesis) was labelled with FITC and purified as described above. The FITC labelled PTAP peptide eluted at 48-49% solvent B.

### **Microscale Thermophoresis (MST)**

The purified recombinant His<sub>6</sub>-UEV protein was labelled with the red NT-647-Maleimide dye using a protein-labelling kit (NanoTemper Technologies) according to the manufacturer's instructions. The labelled protein was then purified using a gravity flow column into MST assay buffer (20 mM

HEPES, pH 7.4, 100 mM NaCl, 5% glycerol, 0.01% Tween-20). The inhibitor peptide (XY3-3, peptide **3**) was dissolved in 100% DMSO and then diluted 5-fold into MST buffer. The inhibitor was serially diluted in MST assay buffer supplemented with 20% DMSO (5  $\mu$ L) and then mixed with the labelled protein 1:1 (10  $\mu$ L final volume, 0.01  $\mu$ M final concentration of protein, final DMSO concentration 10%). The samples were incubated at room temperature for 5 min and then loaded into MST premium coated capillaries (NanoTemper Technologies). The samples were analysed on a Monolith NT.115 MST Instrument (NanoTemper Technologies, Excitation Power 20%, MST Power 40%). Data points were excluded if the initial fluorescence was outside of the 10% error range as recommended by the manufacturer. Data was processed with MO.Affinity.Analysis (NanoTemper) and Prism 6 (GraphPad). Binding constants were determined using a non-linear regression model from three replicates.

#### **Fluorescence polarization of UEV binding to peptide 3 and PTAP peptide**

Purified recombinant His<sub>6</sub>-UEV protein was diluted into binding buffer (50 mM HEPES, pH 7.5, 300 mM NaCl, 20% glycerol, 0.5 mM TCEP) and mixed with 150 nM FITC-labeled peptide **3** or FITC-PTAP. Binding assays were incubated at 25 °C for 4 h and analyzed using a Synergy H4 Hybrid plate reader with  $\lambda_{\text{ex}} = 485$  nm and  $\lambda_{\text{em}} = 538$  nm using in triplicate. Then each experiment was repeated two more times for a total of three independent experiments. Dissociation constants were calculated from the 50% saturation point using a dose-response curve fit in Prism 6 (GraphPad) with three replicates. A blank was set up with protein with the FITC fluorophore only.

#### **Fluorescence polarization competition binding assay**

Recombinant His<sub>6</sub>-UEV protein 15  $\mu$ M/4  $\mu$ M was mixed with 1.5  $\mu$ M FITC-PTAP/400 nM FITC-XY3-3 in binding buffer and incubated at 25 °C for 4 h to reach equilibration. Unlabelled peptide **3**/PTAP peptide and thrombin cleaved GST-p6 were serially diluted in binding buffer and added to the pre-incubated mix of UEV with FITC-labelled PTAP/peptide **3**. Binding assays were analysed on the same instruments and analysed with the same software as described above with three replicates.

## Generation of XY3-3-Tat

For Tat labelling of peptide **3**, the Gly residue at the –5 position of XY3-3 was mutated to Cys using primers XY33G-5C\_FP and XY33G-5C\_RP (Supplementary Table 2, Supplementary Figure 12) in a pRSFDuet plasmid with the gene encoding XY3-3 in MCS1 and *procM* in MCS2. The peptide was overexpressed in *E. coli* and purified as described previously. The resulting peptide contained two non-overlapping rings, as desired, suggesting the mutation did not interrupt the modification process (Supplementary Fig. 17). The peptide was then digested by endoproteinase GluC and purified as described previously. The Tat-peptide (CGRKKRRQRRRPPQ) was synthesised using Fmoc solid-phase peptide synthesis on a Glutamine Wang Resin (Matrix Innovation). The cysteine side chain was activated by reacting with aldrithiol-4 (2 equivalents, 10 mL DMF, 16 h, Sigma). The Glu-C-cleaved peptide was dissolved in 3:1 DMF:H<sub>2</sub>O (2 mL) with aldrithiol-activated Tat (3.5 equivalents) and stirred under argon for 18 h. The resulting peptide (named XY3-3-Tat) was precipitated in ice-cold diethyl ether and dried under argon.

## Mammalian cell culture and transfection

HEK293T cells were maintained in DMEM Glutamax supplemented with antibiotics and 10% FBS (Life Technologies). HeLa cells were maintained in DMEM supplemented with 10% FBS. Plasmid DNA transfections were performed using Lipofectamine 2000 (Invitrogen).

## Antibodies and Western Immunoblotting

To visualise expression of transfected proteins, the cells were harvested in cold PBS (ThermoFisher) and then centrifuged (1000 rpm, 4 min, r.t.). The cells were resuspended in RIPA Buffer with protease inhibitors (ThermoFisher) and incubated on ice for 5 min. The cell lysates were then sonicated in an ice water bath (12 cycles of 30 s on, 30 s off) before being centrifuged (10,000 rpm, 20 min, 4 °C) and then the protein concentration in the supernatant was quantified by Bradford Assay. The proteins were separated on a 12% SDS-PAGE gel (v/v) under denaturing conditions (160 V, 50-70 min), then transferred to a nitrocellulose membrane (GE Healthcare, 250 mA, 2 h). The membrane was blocked

with 5% (w/v) non-fat powdered milk with 0.1% Tween-20 (v/v) in PBS and then subjected to immunoblot analysis. The primary antibodies, either anti-eGFP (MA1-952, 1:1000, ThermoFisher), or anti-EGFR (ab252894, 1:1000, Abcam), were diluted in 5% (w/v) non-fat powdered milk with 0.1% Tween-20 (v/v) in PBS and incubated with the membrane overnight at 4 °C. Horseradish peroxidase conjugated anti-mouse (NA931, 1:100,000, GE Healthcare) or anti-rabbit (7074, 1:50,000, Cell Signalling) antibodies were used as the secondary antibodies for anti-eGFP and anti-EGFR respectively. Monoclonal anti- $\beta$ -actin-peroxidase antibody (A3854, 1:100,000, Sigma) was used as a loading control. Bound immunocomplexes were detected using ECL prime Western blot detection reagent (RPN2232, GE Healthcare) and analyzed using a ChemiDoc Imaging System (Bio-Rad) and Image Lab 4.0 Software (Bio-Rad).

#### **Virus-like particle assay**

HEK293T cells were seeded at 25,000 cells/well in a 24-well plate and incubated overnight such that the cells reached 60-70% confluency prior to transfection. The cells were transfected with Lipofectamine 2000 and pMET7-GAG-EGFP (Addgene) in a 4:1 ratio in Opti-MEM (Life Technologies). After 6 h, the cells were either mock-treated with control or Tat-peptide, or treated with increasing concentrations of XY3-3-Tat. After 18 h, the supernatants were collected and the cells washed and lysed prior to immunoblotting analysis. The supernatants were centrifuged (1000 rpm, 4 min), and then filtered through a 20% (w/v) sucrose cushion (12,000 rpm, 90 min, 4 °C). The filtered VLPs (final 15  $\mu$ L) were resuspended with Laemmli Buffer (5  $\mu$ L) and then subjected to SDS-PAGE and immunoblotting analysis.

#### **Cytotoxicity assay**

The cytotoxicity was investigated using the CytoTox96<sup>®</sup> Non-Radioactive Cytotoxicity Assay Kit (Promega) used according to the manufacturer's guidelines and the absorbance measured using a BMG CLARIOstar Plate reader.



#### **EGFR down-regulation assay**

HeLa cells were seeded at 150,000 cells/well in a 6-well plate and incubated overnight such that the cells reached 60-70% confluency. Cells were either treated with Tat peptide (0.5  $\mu$ M) or XY3-3-Tat (0.5  $\mu$ M) in DMEM Glutamax supplemented with 10% FBS and incubated for 18 h. The cells were then starved in serum-free DMEM Glutamax. After 1 h, the cells were either mock-treated or treated with EGF (150 ng/mL, Sigma) for 90 min. The cells were washed with cold PBS and immediately lysed and subjected to SDS-PAGE and immunoblotting analysis.

## References

1. Marsault, E. & Peterson, M.L. Macrocycles are great cycles: applications, opportunities, and challenges of synthetic macrocycles in drug discovery. *J. Med. Chem.* **54**, 1961-2004 (2011).
2. Driggers, E.M., Hale, S.P., Lee, J. & Terrett, N.K. The exploration of macrocycles for drug discovery--an underexploited structural class. *Nat. Rev. Drug Discov.* **7**, 608-624 (2008).
3. Cardote, T.A. & Ciulli, A. Cyclic and macrocyclic peptides as chemical tools to recognise protein surfaces and probe protein-protein interactions. *ChemMedChem* **11**, 787-794 (2015).
4. Gao, M., Cheng, K. & Yin, H. Targeting protein-protein interfaces using macrocyclic peptides. *Biopolymers* **104**, 310-316 (2015).
5. Lennard, K.R. & Tavassoli, A. Peptides come round: using SICLOPPS libraries for early stage drug discovery. *Chemistry* **20**, 10608-10614 (2014).
6. Heinis, C., Rutherford, T., Freund, S. & Winter, G. Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol.* **5**, 502-507 (2009).
7. Miranda, E. et al. A cyclic peptide inhibitor of HIF-1 heterodimerization that inhibits hypoxia signaling in cancer cells. *J. Am. Chem. Soc.* **135**, 10418-10425 (2013).
8. Passioura, T., Katoh, T., Goto, Y. & Suga, H. Selection-based discovery of druglike macrocyclic peptides. *Annu. Rev. Biochem.* **83**, 727-752 (2014).
9. Heinis, C. & Winter, G. Encoded libraries of chemically modified peptides. *Curr. Opin. Chem. Biol.* **26**, 89-98 (2015).
10. Birts, C.N. et al. A cyclic peptide inhibitor of C-terminal binding protein dimerization links metabolism with mitotic fidelity in breast cancer cells. *Chem. Sci.* **4**, 3046-3057 (2013).
11. Tavassoli, A. SICLOPPS cyclic peptide libraries in drug discovery. *Curr. Opin. Chem. Biol.* **38**, 30-35 (2017).
12. Arnison, P.G. et al. Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30**, 108-160 (2013).
13. Ortega, M.A. & van der Donk, W.A. New insights into the biosynthetic logic of ribosomally synthesized and post-translationally modified peptide natural products. *Cell Chem. Biol.* **23**, 31-44 (2016).
14. Kuipers, O.P. et al. Protein engineering of lantibiotics. *Antonie van Leeuwenhoek* **69**, 161-169 (1996).
15. Cotter, P.D. et al. Complete alanine scanning of the two-component lantibiotic lactacin 3147: generating a blueprint for rational drug design. *Mol. Microbiol.* **62**, 735-747 (2006).
16. Pavlova, O., Mukhopadhyay, J., Sineva, E., Ebright, R.H. & Severinov, K. Systematic structure-activity analysis of microcin J25. *J. Biol. Chem.* **283**, 25589-25595 (2008).
17. Islam, M.R. et al. Evaluation of essential and variable residues of nukacin ISK-1 by NNK scanning. *Mol. Microbiol.* **72**, 1438-1447 (2009).
18. Pan, S.J. & Link, A.J. Sequence diversity in the lasso peptide framework: discovery of functional microcin J25 variants with multiple amino acid substitutions. *J. Am. Chem. Soc.* **133**, 5016-5023 (2011).
19. Young, T.S., Dorrestein, P.C. & Walsh, C.T. Codon randomization for rapid exploration of chemical space in thiopeptide antibiotic variants. *Chem. Biol.* **19**, 1600-1610 (2012).
20. Zhang, F. & Kelly, W.L. In vivo production of thiopeptide variants. *Methods Enzymol.* **516**, 3-24 (2012).
21. Boakes, S. et al. Generation of an actagardine A variant library through saturation mutagenesis. *Appl. Microbiol. Biotechnol.* **15**, 1509-1517 (2012).

- 803 22. Deane, C.D., Melby, J.O., Molohon, K.J., Susarrey, A.R. & Mitchell, D.A. Engineering unnatural variants of  
804 plantazolicin through codon reprogramming. *ACS Chem. Biol.* **8**, 1998-2008 (2013).
- 805 23. Weiz, A.R. et al. Harnessing the evolvability of tricyclic microviridins to dissect protease-inhibitor interactions.  
806 *Angew. Chem. Int. Ed.* **53**, 3735-3738 (2014).
- 807 24. Houssen, W.E. et al. An efficient method for the in vitro production of azol(in)e-based cyclic peptides. *Angew. Chem.*  
808 *Int. Ed.*, 10.1002/anie.201408082 (2014).
- 809 25. Ruffner, D.E., Schmidt, E.W. & Heemstra, J.R. Assessing the combinatorial potential of the RiPP cyanobactin tru  
810 pathway. *ACS Synth. Biol.* **4**, 482-492 (2015).
- 811 26. Repka, L.M., Chekan, J.R., Nair, S.K. & van der Donk, W.A. Mechanistic understanding of lanthipeptide biosynthetic  
812 enzymes. *Chem. Rev.* **117**, 5457-5520 (2017).
- 813 27. Li, B. et al. Catalytic promiscuity in the biosynthesis of cyclic peptide secondary metabolites in planktonic marine  
814 cyanobacteria. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10430-10435 (2010).
- 815 28. Tang, W. & van der Donk, W.A. Structural characterization of four prochlorosins: a novel class of lantipeptides  
816 produced by planktonic marine cyanobacteria. *Biochemistry* **51**, 4271-4279 (2012).
- 817 29. Yu, Y., Mukherjee, S. & van der Donk, W.A. Product formation by the promiscuous lanthipeptide synthetase ProcM  
818 is under kinetic control. *J. Am. Chem. Soc.* **137**, 5140-5148 (2015).
- 819 30. Giordanetto, F. & Kihlberg, J. Macrocyclic drugs and clinical candidates: what can medicinal chemists learn from  
820 their properties? *J. Med. Chem.* **57**, 278-295 (2014).
- 821 31. Foster, A.D. et al. Methods for the creation of cyclic peptide libraries for use in lead discovery. *J. Biomol. Screen.* **20**,  
822 563-576 (2015).
- 823 32. Quartararo, J.S. et al. A bicyclic peptide scaffold promotes phosphotyrosine mimicry and cellular uptake. *Bioorg.*  
824 *Med. Chem.* **22**, 6387-6391 (2014).
- 825 33. Tavassoli, A. et al. Inhibition of HIV budding by a genetically selected cyclic peptide targeting the Gag-TSG101  
826 interaction. *ACS Chem. Biol.* **3**, 757-764 (2008).
- 827 34. Garrus, J.E. et al. Tsg101 and the vacuolar protein sorting pathway are essential for HIV-1 budding. *Cell* **107**, 55-65  
828 (2001).
- 829 35. Martin-Serrano, J., Zang, T. & Bieniasz, P.D. HIV-1 and Ebola virus encode small peptide motifs that recruit Tsg101 to  
830 sites of particle assembly to facilitate egress. *Nat. Med.* **7**, 1313-1319 (2001).
- 831 36. Demirov, D.G., Ono, A., Orenstein, J.M. & Freed, E.O. Overexpression of the N-terminal domain of TSG101 inhibits  
832 HIV-1 budding by blocking late domain function. *Proc. Natl. Acad. Sci. USA* **99**, 955-960 (2002).
- 833 37. Shi, Y., Yang, X., Garg, N. & van der Donk, W.A. Production of lantipeptides in *Escherichia coli*. *J. Am. Chem. Soc.* **133**,  
834 2338-2341 (2011).
- 835 38. VerPlank, L. et al. Tsg101, a homologue of ubiquitin-conjugating (E2) enzymes, binds the L domain in HIV type 1  
836 Pr55(Gag). *Proc. Natl. Acad. Sci. USA* **98**, 7724-7729 (2001).
- 837 39. Male, A.L. et al. Targeting *Bacillus anthracis* toxicity with a genetically selected inhibitor of the PA/CMG2 protein-  
838 protein interaction. *Sci. Rep.* **7**, 3104 (2017).
- 839 40. Im, Y.J. et al. Crystallographic and functional analysis of the ESCRT-I /HIV-1 Gag PTAP interaction. *Structure* **18**, 1536-  
840 1547 (2010).
- 841 41. Pornillos, O., Alam, S.L., Davis, D.R. & Sundquist, W.I. Structure of the Tsg101 UEV domain in complex with the PTAP  
842 motif of the HIV-1 p6 protein. *Nat. Struct. Biol.* **9**, 812-817 (2002).
- 843 42. Pornillos, O. et al. Structure and functional interactions of the Tsg101 UEV domain. *EMBO J.* **21**, 2397-2406 (2002).

844 43. Morris, C.R., Stanton, M.J., Manthey, K.C., Oh, K.B. & Wagner, K.U. A knockout of the Tsg101 gene leads to  
845 decreased expression of ErbB receptor tyrosine kinases and induction of autophagy prior to cell death. *PLoS One* **7**,  
846 e34308 (2012).

847 44. Ruland, J. et al. p53 accumulation, defective cell proliferation, and early embryonic lethality in mice lacking tsg101.  
848 *Proc. Natl. Acad. Sci. USA* **98**, 1859-1864 (2001).

849 45. Pornillos, O. et al. HIV Gag mimics the Tsg101-recruiting activity of the human Hrs protein. *J. Cell Biol.* **162**, 425-434  
850 (2003).

851 46. Lu, Q., Hope, L.W., Brasch, M., Reinhard, C. & Cohen, S.N. TSG101 interaction with HRS mediates endosomal  
852 trafficking and receptor down-regulation. *Proc. Natl. Acad. Sci. USA* **100**, 7626-7631 (2003).

853 47. Jencks, W.P. On the attribution and additivity of binding energies. *Proc. Natl. Acad. Sci. USA* **78**, 4046-4050 (1981).

854 48. Knappe, T.A. et al. Introducing lasso peptides as molecular scaffolds for drug design: engineering of an integrin  
855 antagonist. *Angew. Chem. Int. Ed.* **50**, 8714-8717 (2011).

856 49. Conibear, A.C. et al. Approaches to the stabilization of bioactive epitopes by grafting and peptide cyclization.  
857 *Biopolymers* **106**, 89-100 (2016).

858 50. Rink, R. et al. To protect peptide pharmaceuticals against peptidases. *J. Pharmacol. Toxicol. Methods* **61**, 210-218  
859 (2010).

860 51. Newcombe, R.G. Two-sided confidence intervals for the single proportion: comparison of seven methods. *Stat.*  
861 *Med.* **17**, 857-872 (1998).

862

863