Published in final edited form as:

J Nat Prod. 2019 October 25; 82(10): 2744–2753. doi:10.1021/acs.jnatprod.9b00352.

PepSAVI-MS Reveals a Proline-rich Antimicrobial Peptide in *Amaranthus tricolor*

Tessa B. Moyer^{†,||}, Lilian R. Heil^{†,||}, Christine L. Kirkpatrick[†], Dennis Goldfarb[‡], William A. Lefever[§], Nicole C. Parsley[†], Andrew J. Wommack[§], Leslie M. Hicks^{*,†}

[†]Department of Chemistry, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina United States

[‡]Lineberger Comprehensive Cancer Center, University of North Carolina at Chapel Hill, Chapel Hill, North Carolina United States

[§]Department of Chemistry, High Point University, High Point, North Carolina United States

Abstract

Traditional medicinal plants are a rich source of antimicrobials; however, the bioactive peptide constituents of most ethnobotanical species remain largely unexplored. Herein, PepSAVI-MS, a mass spectrometry-based peptidomics pipeline, was implemented for antimicrobial peptide (AMP) discovery in the medicinal plant *Amaranthus tricolor*. This investigation revealed a novel 1.7 kDa AMP with strong activity against *Escherichia coli* ATCC 25922, deemed Atr-AMP1. Initial efforts to determine the sequence of Atr-AMP1 utilized chemical derivatization and enzymatic digestion to provide information about specific residues and post-translational modifications. EThcD (electron-transfer/higher-energy collision dissociation) produced extensive backbone fragmentation and facilitated de novo sequencing, the results of which were consistent with orthogonal characterization experiments. Additionally, multistage HCD (higher-energy collisional dissociation) facilitated discrimination between isobaric leucine and isoleucine. These results revealed a positively charged proline-rich peptide present in a heterogeneous population of multiple peptidofoms, possessing several post-translational modifications including a disulfide bond, methionine oxidation, and proline hydroxylation. Additional bioactivity screening of a simplified fraction containing Atr-AMP1 revealed activity against *Staphylococcus aureus* LAC, demonstrating activity against both a Gram-negative and a Gram-positive bacterial species unlike many known short chain proline-rich antimicrobial peptides.

Graphical abstract

*Corresponding Author Phone: 919-843-6903. lmhicks@unc.edu.

^{||}T.B.M. and L.R.H. contributed equally.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.jnatprod.9b00352.

Supporting mass spectra, Atr-AMP1 oxidation states, bioassays, and database searching results (PDF)

The authors declare no competing financial interest.

The threat of difficult-to-treat infections has increased interest in identifying antimicrobials naturally produced by plants, especially in elucidating the chemical basis of traditional medicinal remedies.^{1,2} Native to Central and South America, the genus *Amaranthus* contains ~60 species and is known worldwide for its medicinal properties, particularly the treatment of urinary infections, diarrhea, jaundice, fungal infections, and skin conditions.^{3–10} Recent research has confirmed activities including anticancer, antibacterial, hepatoprotective, and antimalarial of *Amaranthus* spp. extracts.^{11–16} While the majority of studies investigating *Amaranthus* implement organic extractions and focus on small molecule constituents, antimicrobial peptides (AMPs) have been experimentally validated at the protein level in *A. retroflexus*,¹⁷ *A. caudatus*,¹⁸ and *A. hypochondriacus*.¹⁹ These AMPs belong to the hevein family and exhibit activity against several fungal^{17–19} and Gram-positive bacterial¹⁸ species. AMPs with high sequence homology to Ac-AMP2 isolated from *A. caudatus* have been predicted using PCR gene amplification in several other *Amaranthus* species including *A. tricolor*.²⁰ However, AMP discovery of this nature is reliant on similarity to known peptides and requires further experimental validation of activity. Members of the *Amaranthus* genus are thus promising targets for lead compound antimicrobial peptide discovery, with the potential to selectively and efficiently treat microbial infections.⁴

A. tricolor, commonly known as red spinach or Joseph's coat, has a rich ethnobotanical history. The whole plant has been consumed for centuries in Indian culture to promote general health, and the roots have been used to control bleeding.²¹ Recently, leaves of *A. tricolor* were revealed to be antinociceptive, antidiabetic, hepatoprotective, gastroprotective, neuroprotective, and anti-inflammatory.^{22–27} *A. tricolor* extract activity has been attributed to the whole extract (i.e., a specific active compound was not identified) or to small molecules ranging from phenolic compounds to betalains,^{28–32} but no reports have yet identified AMPs as contributors to its activity.

Herein, the PepSAVI-MS pipeline^{33–36} was implemented for the identification of a new antimicrobial peptide from *A. tricolor* (Figure 1). PepSAVI-MS employs crude fractionation of extracts to generate peptide libraries which are screened for biological activity and profiled with mass spectrometry. Statistical modeling of these data sets is utilized to identify putative bioactive peptides for further molecular and biological characterization. PepSAVI-MS is well suited for bioactive peptide discovery in nonmodel plant species, including many used in traditional medicine, as it is not reliant on prior knowledge of the genome, transcriptome, or specific biosynthetic pathways. Following this pipeline, a peptide library generated from the aerial tissue of *A. tricolor* was screened for activity against the Gram-

negative bacterium *Escherichia coli* ATCC 25922. PepSAVI-MS was employed to prioritize top candidates likely contributing to the observed bioactivity, and from these top contributors, a novel antimicrobial peptide, Atr-AMP1, was identified. Molecular characterization of Atr-AMP1 was determined de novo via complementary biochemical assays, tandem mass spectrometry (MS/MS) including electron-transfer/higher-energy collision dissociation (ET_hCD)³⁷ and higher-energy collisional dissociation (HCD), and MS³ techniques (Figure S1, Supporting Information (SI)). Results reveal that Atr-AMP1 is present in a heterogeneous population of multiple peptidoforms, possessing several post-translational modifications, including a single disulfide bond, methionine oxidation, and hydroxylation of multiple proline residues. A simplified peptide library fraction containing Atr-AMP1 possessed activity against methicillin-resistant *Staphylococcus aureus* LAC, demonstrating that Atr-AMP1 is active against both a Gram-negative and a clinically relevant Gram-positive pathogen.

RESULTS AND DISCUSSION

Peptide Library Bioactivity.

The *A. tricolor* aerial tissue crude extract was subjected to reversed-phase chromatography (Figure 2A) to generate a peptide library. This library was screened for bioactivity against laboratory strain *Escherichia coli* ATCC 25922, and revealed a robust bioactive region between fractions 31 and 35, reaching 80% activity compared to the antibiotic control (Figure 2B), which was prioritized for further investigation.

PepSAVI-MS.

Profiling of *A. tricolor* peptide library fractions 19–46 via LC-MS/MS revealed 5,868 unique features with masses, charge states, and retention times (1,000–15,000 Da, +2–9, and 15–45 min, respectively) in the range of typical AMPs, highlighting the complex nature of the peptide library. The resulting list was filtered so that the maximum abundance of each feature was above 100 and detected in fractions 29–37, with <5% maximum abundance outside of the defined bioactive region (fractions 29–37). This filtered list was modeled against the bioactivity profile using an elastic net penalized linear regression to identify the top 20 peptidyl features most likely contributing to bioactivity (Figure 2B). A 1703.81 Da peptide appeared three times in the top 10 candidates (due to detection of three separate charge states: *m/z* 426.97 (+4), 568.94 (+3), and 852.91 (+2)) and its abundance profile closely mirrored the bioactivity across fractions (Figure 2B). This putative antimicrobial peptide, denoted Atr-AMP1, was prioritized for molecular characterization.

Sequence Characterization.

Oxidation States.—Further inspection of the MS data revealed four peptides whose masses differ from Atr-AMP1 in increments of 15.99 Da (SI Figure S2a). This mass difference may be due to the presence of multiple peptidoforms differing by the addition of an oxygen atom during oxidation or hydroxylation of an amino acid, which has been observed in several known AMPs³⁸ and increase the oxidation state of the peptide. The presence of two oxidation states below Atr-AMP1 suggests that two oxidized or hydroxylated residues are present in the active 1703.81 Da peptide. Additionally, detection

of two oxidation states higher than Atr-AMP1 suggests that there are additional residues that can be modified by oxidation or hydroxylation, creating a mixed population of unresolved isobaric peptidoforms that could complicate sequencing attempts. These oxidation states of Atr-AMP1 are distributed across fractions 27–39 (SI Figure S2b), most of which are inactive, and only Atr-AMP1 appears in the ranked top 20 list from modeling (Figure 2b), suggesting that the active peptidoform includes at least two modified residues. A –64.03 Da neutral loss was observed following collision-induced dissociation (CID) of the intact precursor (SI Figure S3a) as well as CID of the reduced and alkylated peptide (SI Figure S3b). This characteristic mass shift is indicative of an oxidized methionine residue³⁹ and suggests that methionine oxidation accounts for at least one of the oxidative PTMs observed in the active peptide while modifications such as oxidation or hydroxylation of other residues may account for others. Oxidation or hydroxylation of amino acids in AMPs has been shown to impact bioactivity,^{40,41} and in this case these modifications appear necessary for AMP activity.

S–S Identification.—AMPs frequently contain disulfide bonds^{38,42} which can improve structural stability,⁴³ may be necessary for activity,⁴⁴ and often decrease MS/MS fragmentation efficiency.⁴⁵ To determine the presence of any disulfide bonds and improve fragmentation, a fraction containing Atr-AMP1 was first reduced with dithiothreitol (DTT) and then alkylated with iodoacetamide (IAM). MS analysis revealed Atr-AMP1 with a mass shift of +116.04 Da (SI Figure S4), corresponding to one intramolecular disulfide bond between two cysteine residues.³⁹ CID fragmentation of the reduced and alkylated Atr-AMP1 produced richer spectra than the native peptide but did not provide sufficient sequence information for complete de novo sequencing (SI Figure S3). Additional derivatization, digestion, and MS/MS strategies were necessary to elucidate the sequence.

Charged Residue Identification.—A combination of proteolysis and derivatization was used to ascertain the presence of positively charged residues (i.e., lysine/arginine). Tryptic digestion of reduced and alkylated *A. tricolor* fractions resulted in the disappearance of alkylated Atr-AMP1 (data not shown), indicating the presence of at least one lysine or arginine residue. The complex nature of the sample made it impossible to identify the resulting tryptic peptides without knowledge of the sequence. For additional sequence information, the number of primary amine moieties (i.e., unmodified lysine residues and free *N*-terminus) in a peptide can be determined via *N,N*-dimethyl labeling.^{46,47} After reduction and alkylation, formaldehyde was used to modify the peptide so that +28.03 Da is added to each primary amine unit. After *N,N*-dimethyl labeling, an 1847.91 Da feature which was not present in the sample prior to derivatization was observed (SI Figure S5), indicating a +28.02 Da mass shift as a result of a single dimethylation. No other mass shifts corresponding to additional dimethylation events were observed, indicating the presence of only one primary amine moiety. The single primary amine group is likely the free *N*-terminus, suggesting an absence of lysine residues. This combined with the trypsin digestion data indicates the presence of at least one arginine residue in Atr-AMP1.

For further sequence information, the number of acidic residues can be determined via methylation. Methanolic HCl is used to modify the *C*-terminus, aspartic acid, and glutamic

acid, such that a methyl group (+14.01 Da) is added to each hydroxycarbonyl moiety.⁴⁸ Methylation of intact Atr-AMP1 resulted in a mass shift of +13.98 Da, and this 1717.80 Da feature was not observed in the sample before derivatization. This mass shift corresponds to methylation of the C-terminus (SI Figure S6), indicating that the peptide does not contain acidic residues. These derivatization experiments provide valuable insights into the primary sequence composition and facilitate the evaluation of de novo sequencing results.

De Novo Sequencing Via MS/MS.—As noted previously, MS/MS experiments using CID fragmentation of intact or derivatized peptides did not provide sufficient spectra for de novo sequencing. To address this, reduced and alkylated Atr-AMP1 was targeted for fragmentation by EThcD (Figure 3), which produced an extensive fragmentation spectrum. EThcD improves fragmentation along the peptide backbone by first activating the precursor ion with electron transfer dissociation (ETD) and then fragmenting it with HCD. This hybrid fragmentation technique can yield full coverage of difficult-to-sequence peptides.³⁷

Analysis of EThcD spectra using PEAKS Studio resulted in a sequence assignment of M(OX)JCP(HY)SYRRCPRVPP (calculated mass 1819.8644 Da), where all cysteine units are modified with IAM, Met1 is oxidized, J represents isobaric residues isoleucine (Ile)/leucine (Leu), and Pro4 is hydroxylated (Figure 3, SI Figure S7). Other possible assignments vary only in hydroxyproline localization (hyP13, hyP14) suggesting that multiple peptidofoms are present. Hydroxylation was never detected on Pro10, suggesting that proline hydroxylation occurs preferentially at certain positions on this peptide. The sequence assignment is consistent with prior experimental results because it contains two cysteine moieties which form a disulfide bond, an oxidized methionine, additional residues that could be oxidized or hydroxylated (i.e., P, Y), at least one arginine, no lysines, and no acidic residues. The experimental mass (1819.8653 Da) and the calculated mass (1819.8644 Da) of Atr-AMP1 differ by 0.49 ppm. The high mass accuracy, depth of coverage, and consistency with previous experiments provide strong support for this sequence. High-throughput transcriptome sequencing supported the primary sequence of Atr-AMP1.

Subsequent enzymatic digests provided additional validation of the assigned sequence and post-translational modifications. A combination of results from trypsin (Figure 4) and chymotrypsin (Figure 5) digestion yielded peptides confirming 100% coverage of the sequence. These data also confirmed that Atr-AMP1 encompasses at least three peptidofoms that vary in the localization of a single hydroxyproline unit at positions 4 (Figure 4B), 13 (Figure 5B), and 14 (Figure 5C), as suggested by the de novo sequencing results. The hydroxyproline peptidofoms were not chromatographically resolved, making it impossible to determine which modification confers activity.

Discrimination of Ile/Leu.—An MS³ method⁴⁹ was implemented to assign the unknown residue at position 2 because MS/MS fragmentation alone cannot be used to distinguish between the isobaric residues leucine and isoleucine. This approach takes advantage of the differing HCD fragmentation patterns of Ile and Leu immonium ions. Briefly, the peptide is selected for MS/MS fragmentation with HCD, producing an 86.10 Da immonium ion that is selected for MS³ fragmentation. The Ile immonium ion is known to produce a highly abundant *m/z* 69.07 fragment, whereas the Leu immonium ion yields the *m/z* 69.07 product

peak in low abundance relative to the 86.10 Da precursor ion. In order to achieve maximum signal at the MS³ level, the highly abundant peptide corresponding to Atr-AMP1–16 Da, 1687.81 Da, was targeted for MS³ analysis. An EThcD spectrum of Atr-AMP1–16 Da was sequenced (data not shown) to confirm that it contained the same primary sequence as Atr-AMP1 varying only in the loss of hydroxyproline. MS³ analysis of Atr-AMP1–16 yielded a low abundance peak at *m/z* 69.05 (SI Figure S8), indicating that the unknown residue is Leu2.

Additional Structural Characterization.

To validate the observed activity and determine the biological relevance of methionine oxidation, both the oxidized and unoxidized form of each hydroxyproline peptidofragment was synthesized and screened individually and in combination against *E. coli* ATCC 25922 (data not shown). No activity was observed, suggesting a discrepancy between the naturally derived and synthetic peptides. It is possible that a number of stereochemical modifications could be necessary for activity, including the presence of a *cis*-proline⁵⁰ or a D-amino acid.⁵¹

Isobaric stereochemical modifications are difficult to detect using traditional MS approaches, but may be identified chromatographically by comparing the retention times of inactive synthetic and active native peptides.^{50,52,53} During strong cation exchange (SCX) fractionation, it was observed that native Atr-AMP1 has a wide elution profile across SCX fractions (SI Figure S9a) which closely mirrors the activity of SCX fractions against *E. coli* ATCC 25922 (SI Figure S9b). When synthetic Atr-AMP1 (hyp4) was subjected to the same SCX separation conditions, it had a longer retention time than native Atr-AMP1 (SI Figure S9a). Stereochemical modifications of the active peptide could change the secondary structure of Atr-AMP1, affecting the strength of its interaction with the anionic stationary phase, thus altering its retention time. For example, the presence of a *cis*-proline has been shown to decrease SCX retention time and confer activity,^{50,53} additionally the presence of 3-hydroxyproline instead of 4-hydroxyproline has been shown to impact reversed-phase retention and here may alter SCX retention and influence activity.⁵⁴ The extremely low error (less than 1 ppm) between the calculated and experimental mass of Atr-AMP1 and the thorough validation of the sequence via proteolytic digestion and chemical derivatization suggests that one or more unknown isobaric stereochemical modifications such as *cis/trans* isomerization may be responsible for the lack of activity and retention time shift observed in the synthetic peptide.

Additional Bioactivity/MOA.

In order to gain insight into the structure of Atr-AMP1, a homology search was conducted with the Antimicrobial Peptide Database (<http://aps.unmc.edu/AP/main.php>).⁵⁵ This search revealed >40% sequence homology with a class of proline-rich AMPs (PrAMPs) known as Metalnikowins, but provided little insight into the structure and presence of stereochemical modifications. The high proline content and positive charge of Atr-AMP1 suggests that it can be classified as a short-chain PrAMP which are known to exhibit diverse activity⁵⁶ and have been identified in plants,⁵⁷ mammals,⁵⁸ and arthropods.⁵⁹ Several PrAMPs are known to be post-translationally modified, but the most common PTM observed in short chain

PrAMPs, glycosylation of a threonine residue,⁶⁰ is not observed in Atr-AMP1. Disulfide bonds have been identified in longer PrAMPs such as penaeidins⁶¹ and arasins,⁶² but this is the first reported short-chain PrAMP with a disulfide bond. Atr-AMP1 is the first short-chain PrAMP isolated from plants but several botanical long-chain PrAMPs have been identified.^{57,63–65} Future studies should explore the activity of Atr-AMP1 against fungal species because botanical long-chain PrAMPs are active against fungi.

Notably, many known short-chain PrAMPs specifically target Gram-negative bacteria via nonlytic mechanisms of action such as inhibition of protein synthesis in the ribosome.^{56,66} PrAMPs with this mechanism enter the cell via transporter proteins such as SbmA, YgdD, MdtM/YjiO.^{67–69} Mutations which inhibit the function of one or more of these transport proteins have been shown to confer resistance against multiple PrAMPs. The inner-membrane protein SbmA appears to be particularly important for PrAMP uptake and homologous proteins exist among Gram-negative species accounting for the specificity of many PrAMPs for Gram-negative bacteria.⁷⁰ Additionally, the function of the membrane protein YgdD is not known, but it has been shown to be necessary for activity of several PrAMPs against *E. coli*, suggesting that it may function as a peptide transporter in both Gram-positive and Gram-negative bacteria.⁶⁸

To decrease sample complexity in further activity screens, active reversed-phase *A. tricolor* peptide library fractions were pooled together and fractionated via SCX chromatography. The SCX fraction most active against *E. coli* ATCC 25922 also demonstrated activity against methicillin-resistant *S. aureus* LAC, but not rifampin- and streptomycin-resistant *Klebsiella pneumoniae* VK148 (Figure 6a). It is notable that Atr-AMP1 does not possess broad spectrum activity against Gram-negative bacteria and instead demonstrates activity against Gram-positive *S. aureus* which does not produce the SbmA transporter. To probe the impact of transporter proteins on Atr-AMP1, the activity of the simplified fraction against single gene-knockout strains of *E. coli* IK-12 lacking SbmA (*sbmA*, JW0368) or YgdD (*ygdD*, JW2778) was compared to activity against the parent strain (BW25113) (Figure 6b). The simplified fraction containing Atr-AMP1 did not demonstrate a decrease in activity against the mutant *E. coli* strains. These preliminary data suggest that Atr-AMP1 accesses intracellular targets via other transporters or may have a lytic mechanism of action which is unusual for short chain PrAMPs.

In conclusion, we identified and characterized an antimicrobial peptide from *A. tricolor* via PepSAVI-MS. Once Atr-AMP1 was identified as a contributor to activity against *E. coli* ATCC 25922, orthogonal approaches were used to determine the full sequence de novo and localize PTMs. First, chemical derivatization and proteolysis were used to provide information on the presence of specific residues. Subsequent EThcD and multistage HCD fragmentation resulted in full sequence identification consistent with biochemical experiments. Proteolytic digestion confirmed the sequence and helped reveal a mixed population of hydroxyproline peptidoforms. Chromatographic resolution of native and synthetic peptides suggest that additional stereochemical modifications may be present and necessary for activity. Bioactivity against *S. aureus* and the *E. coli* transporter mutants suggest an alternative mechanism of action for Atr-AMP1 to be explored in future studies. This work highlights the utility of PepSAVI-MS to discover AMPs from botanical species

without relying on sequencing data or a priori knowledge of biosynthetic gene clusters, as well as the power of complementary chemical derivatization, mass spectrometry, and chromatographic techniques to enable molecular characterization.

EXPERIMENTAL SECTION

General Experimental Procedure.

Off-line HPLC separations were performed with a Shimadzu Prominence HPLC equipped with a UV-vis detector (220 nm) (Shimadzu, Kyoto, Japan). LC-MS/MS data of the whole peptide library, chemical derivatization experiments, and enzymatic digestions were analyzed using a nano-LC-ESI-MS/MS platform composed of a nanoAcquity UPLC (Waters, Milford, MA) coupled to a TripleTOF5600 (AB Sciex, Framingham, MA). LC-MS/MS data for de novo sequencing was collected using a nanoAcquity UPLC (Waters, Milford, MA) coupled to an Orbitrap Fusion Lumos with EThcD capability (Thermo Fisher Scientific, San Jose, CA). Peptide synthesis was accomplished using a semiautomated flow chemistry instrument built in-house. The mass spectrometry proteomics data have been deposited to the ProteomeXchange Consortium (<http://proteomecentral.proteomexchange.org>) via the PRIDE partner repository⁷¹ with the data set identifier PXD014547.

Biological Material.

Amaranthus tricolor seeds purchased from Strictly Medicinal Seeds (Williams, OR) were planted in nutrient rich soil and grown under controlled temperature (17.5–20.3 °C) and light cycle (14-h light) conditions. Plants (~100) were grown to mature rosette stage (~8 weeks) and harvested in November 2018 with immediate flash freezing and stored at –80 °C until subsequent extraction. A specimen was deposited in the Herbarium of the University of North Carolina at Chapel Hill (NCU, accession number 666014). *Escherichia coli* ATCC 25922 (RS003) was obtained through the American Type Culture Collection. *E. coli* BW25113 and its transporter mutants *E. coli* JW0368 (*sbmA*) and *E. coli* JW2778 (*ygdD*) were obtained from the Keio collection⁷² through the Coli Genetic Stock Center (<http://cgsc2.biology.yale.edu>). *S. aureus* LAC (RS001)⁷³ was obtained from the Richardson laboratory at the University of North Carolina at Chapel Hill and *K. pneumoniae* VK148 (RS0025)⁷⁴ from the Miller laboratory at University of North Carolina at Chapel Hill.

Creation of Peptide Libraries.

The *A. tricolor* peptide library was created as previously described³³ with modifications noted herein. Aerial tissue was selected for this study because it is abundant and can be rapidly harvested. However, plants are known to differentially express antimicrobial peptides across tissue types.^{75–77} Exploration of peptide libraries generated from other tissue types, such as reproductive organs, could result in the identification of additional antimicrobial peptides. Briefly, 200 g of plant material was extracted in a HOAc solution, size exclusion steps were performed to remove large proteins (>30 kDa) and small molecules (<1 kDa), and the crude extract was concentrated to a final volume of 2 mL. Neutral and negatively charged molecules were removed using SCX chromatography, with the fractionation method previously described.³³ SCX eluate from 12 to 47 min was recombined, desalted using a

Sep-Pak C₁₈ cartridge (Waters, 1 g), and dried via vacuum centrifugation. The extract was resuspended in 1 mL HPLC grade water with 5% MeCN and 0.1% formic acid for peptide library creation. Fractionation was performed in duplicate with 0.490 mL injections on a Jupiter C₁₈ column (150 × 4.6 mm, 5 μm, 300 Å, Phenomenex). A 42 min reversed-phase two-step gradient (1 mL/min, 40 °C) was used starting from 100% mobile phase A (95/5/0.1 water/MeCN/TFA) with the following steps: 0–4 min: 0 %B (100/0.1 MeCN/TFA); 4–25 min: 0–45 %B; 25–36 min: 45–100 %B; 36–40 min: 100 %B; 40–42 min: 100–0 %B. Fractions were collected every 0.5 min, dried, and resuspended in 0.1 mL of water to create the final peptide library.

Bioassays.

Bioassays of peptide library fractions were performed in triplicate in a 96-well plate format. Bacterial cultures were inoculated in 5 mL Mueller Hinton Broth (MHB) and grown at 37° C for 16 h with shaking before dilution in 5 mL MHB to an optical density at 600 nm (OD₆₀₀) of 0.25 (all *E. coli* strains and *K. pneumoniae*) or 0.35 (*S. aureus*). After an additional hour of incubation, cultures were added to a 96-well plate containing 10 μL of fraction or control and 10 μL of 2× MHB to a final OD₆₀₀ of 0.1 and a volume of 50 μL in MHB. Ampicillin (0.1 mg/mL) or erythromycin (0.1 mg/mL) was used in *E. coli* or *S. aureus*/*K. pneumoniae* assays, respectively, as the positive control and water as the negative control. All assays were incubated for 4 h (37 °C, 250 rpm) and their OD₆₀₀ was measured before adding 1 μL of 50 mM resazurin to each well. For *E. coli* BW25113, *E. coli* JW0368, and *E. coli* JW2778 the OD₆₀₀ after 4 h was used to measure bioactivity. For all other assays, after one additional hour of incubation with shaking, a fluorescence read of 544 nm (ex) and 590 nm (em) was collected to measure relative fluorescence units (RFU) for each well. Percent activity was calculated as previously described³³ substituting OD₆₀₀ reading for fluorescence for *E. coli* BW25113, *E. coli* JW0368, and *E. coli* JW2778 assays. Synthetic peptides were cleaned up using a C₁₈ Sep Pak (Waters, 1 g), resuspended at a concentration of approximately 1 mM. Each peptide was assayed in triplicate against *E. coli* ATCC 25922 under the same conditions as peptide library fractions. Hydroxyproline containing peptidoforms were also assayed against *E. coli* ATCC 25922 under the same conditions such that the final test concentration of each peptidoform was approximately 0.07 mM.

Peptide Library MS Data Acquisition.

The *A. tricolor* peptide library was analyzed using a nano-LC-ESI-MS/MS platform as previously described with the following specifications: 0.1% formic acid in all mobile phases and a trapping mobile phase composition of 1% MeCN/0.1% formic acid.³³ The MS was operated in positive-ion, high-sensitivity mode with the MS survey spectrum using a mass range of *m/z* 350–1600 in 250 ms and information-dependent acquisition (IDA) of MS/MS data using an eight second dynamic exclusion window. The first 20 features above an intensity threshold of 150 counts and having a charge state of +2 to +5 were fragmented using rolling collision energy (CE) (±5%). Peptide abundance was quantified using Progenesis QI for Proteomics software (Nonlinear Dynamics, v.2.0) as previously described³³ to create a list of mass spectrometric features for statistical modeling. Default peak picking settings were used with the exception that a minimum peak width of 0.05 min was required. Both reduced and alkylated and trypsin digested fractions were analyzed using

this same method, while chymotrypsin digested fractions were analyzed with an adjusted mass range of m/z 250–1600 in order to detect the highly charged chymotryptic peptides of Atr-AMP1.

Data Reduction and Statistical Modeling.

Data reduction and statistical modeling were performed using the PepSAVI-MS statistical analysis package (<https://cran.r-project.org/package=PepSAVImS>).³³ The bioactive region of the *A. tricolor* peptide library was defined as fractions 29–37. Mass spectrometric features with identical charge states, masses within 0.05 Da, and retention times within 5 min of each other were binned. Background ions were eliminated by selecting for retention time (15–45 min), mass (1,000–15,000 Da), and charge-state (+2–9, inclusive). The resulting list was filtered so that the maximum abundance of each feature was above 100 and detected in fractions 28–38, with <5% maximum abundance outside of the defined bioactive region (29–37). 260 features meeting these filter criteria were modeled using the elastic net estimator with a quadratic penalty parameter specification of 0.01. After modeling, 46 features were identified as possible contributors to bioactivity and the top 20 highest ranked candidates were selected for further investigation.

Reduction, Alkylation, and Digestion.

An active *A. tricolor* fraction with at least 50% activity against *E. coli* ATCC 25922 was reduced with 10mM DTT (Millipore-Sigma) at 45 °C, 850rpm, 30 min, and alkylated with 100 mM IAM (Millipore-Sigma) at 25 °C, 850 rpm, 15 min. For trypsin digestion, a reduced and alkylated sample was digested with an enzyme/protein ratio of 1:50 (w/w) with the endoproteinase trypsin (Millipore-Sigma) and incubated at 37 °C overnight. For chymotrypsin digestion, a reduced and alkylated sample was digested with an enzyme/protein ratio of 1:100 (w/w) with the endoproteinase chymotrypsin (Millipore-Sigma) and incubated at room temperature for 4 h. Samples were desalted with C₁₈ ZipTips (Millipore-Sigma) prior to LC-MS/MS analysis.

Chemical Derivatization.

Methyl esterification of acidic residues was performed as previously described⁴⁸ by concentrating 20 μ L of an active *A. tricolor* fraction with at least 50% activity against *E. coli* ATCC 25922 to dryness, resuspending in 50 μ L methanolic HCl (Millipore-Sigma), and incubating for 2 h at room temperature. *N,N*-dimethyl labeling of the lysine residue was performed as previously described⁴⁶ with 20 μ L of reduced and alkylated fraction reconstituted in 0.1 mL of 100 mM tetraethylammonium bicarbonate (Millipore-Sigma) by adding 4 μ L of 4% formaldehyde (Millipore-Sigma) and 4 μ L of 0.6 M sodium cyanoborohydride (NaBH₃CN, Millipore-Sigma) and incubating for 1 h at room temperature with mixing. The reaction was quenched sequentially with 16 μ L of 1% ammonia solution (Millipore-Sigma) and 8 μ L of 5% formic acid (Millipore-Sigma). All derivatized samples were desalted with C₁₈ ZipTips (Millipore-Sigma) prior to LC-MS/MS analysis.

ETcD Analysis.

For de novo sequencing, approximately 1 μ g of material from a reduced and alkylated fraction containing top contributors to activity was subjected to ETcD on an Orbitrap Fusion Lumos (Thermo Scientific). Peptides were trapped (Pepmap 100, 3 μ m particle size, 100 Å pore size, 2 cm, Thermo Scientific) and separated on a 25 cm EASYspray analytical column (75 μ m I.D., 2.0 μ m C₁₈ particle size, 100 Å pore size, Thermo Scientific) at 300 nL/min and 45 °C). A 60 min gradient from 2 to 35% mobile phase B was applied, where mobile phase A was 0.1% formic acid in water and mobile phase B was 0.1% formic acid in MeCN. Targeted ETcD scans were collected for the +4 charge state of Atr-AMP1 (m/z 451.972) in the Orbitrap using an isolation window of 0.7 Da at a resolving power of 60,000. The product ion scan range was m/z 50–1850 with an AGC target of 1.0e6. The peptide was fragmented with ETD using HCD as secondary activation. ETD conditions were automatically calculated based on m/z and charge state information with rolling collision energy for secondary HCD activation. PEAKS software (BSI, Waterloo, Canada) was used for automatic de novo assignments of ETcD MS/MS data with a parent mass tolerance of 15.0 ppm, and fragment mass tolerance of 0.02 Da. Carbamidomethylation was selected as a fixed modification, and methionine oxidation and proline hydroxylation were allowed as variable modifications. The enzyme was specified as “None”.

To distinguish between the leucine and isoleucine moieties, MS³ data was used as described previously.⁴⁹ Briefly, the target peptide was selected for HCD fragmentation at 40 CE. The 86.09 Da immonium ion was then selected for MS³ fragmentation with HCD at 20 CE, and all levels of MS acquisition were collected in the Orbitrap.

Peptide Spectrum Matching.

To generate fully annotated spectra of intact and digested Atr-AMP1, the peptide sequence determined de novo was appended to a publicly available *A. tricolor* transcriptomic database, and peptide spectrum matching was performed with this complex background. Briefly, the *A. tricolor* transcriptome was obtained from the 1000 Plants Initiative (ID#XSSD, <https://sites.google.com/a/ualberta.ca/onekp/>).⁷⁸ Transcripts assembled with SOAPdenovo⁷⁹ were downloaded and duplicate protein sequences were removed. Because the protein database containing 21,758 unique sequences from 7,237,825 residues did not include the sequence of Atr-AMP1, the determined sequence was appended to the transcriptome for Mascot searching (v.2.5.0; Matrix Science, <http://matrixscience.com/>). Raw data files were centroided and converted to Mascot Generic Format (.mgf) using either MSconvert (ProteoWizard) or Progenesis QI for Proteomics (Nonlinear Dynamics, v.2.0) for the intact and chymotrypsin digested data sets or the tryptic data set, respectively. In all searches, a fixed modification of carbamidomethylation on cysteine residues, and variable modifications of oxidation at the methionine and proline units were specified. For the intact peptide analyzed with ETcD, “none” was selected as the enzyme, and peptide/fragment mass tolerances of 10 ppm/0.05 Da were used. Owing to the complex nature of ETcD fragmentation, a custom instrument type was defined to search the a, b, c, x, y, and z ion series as well as their multiply charged and charge-reduced counterparts (a full list of ion series searched is included in SI Figure S9). For proteolyzed fractions fragmented with CID,

searches were performed with the appropriate specificity for each enzyme, allowing up to 2 missed cleavages, and with peptide/fragment mass tolerances of 10 ppm/0.1 Da.

Fmoc-Based Solid-Phase Peptide Synthesis Using Flow Chemistry.

Synthesis was performed as previously described⁸⁰ using 2-chlorotriyl chloride resin (200 mg, 0.68 mmol/g) with each Fmoc amino acid (1 mmol) dissolved in 2.5 mL of 0.38 M hexafluorophosphate benzotriazole tetramethyl uranium (HBTU) in dimethylformamide (DMF) prior to activation with 450 mL of diisopropylethylamine (or 250 mL with His, Cys, or Trp).

Following coupling and deprotection of the final amino acid, the resin-bound peptide was cleaved using 20 mL of trifluoroacetic acid (TFA)/ethanedithiol/trisopropylsilane/H₂O (94:2.5:2.5:1.0). The crude cleaved peptide was filtered and the solution volume was reduced to 5 mL under N₂ atm. The crude peptide was precipitated with cold Et₂O at 4 °C, centrifugation (3000 rpm, 4 °C, 10 min) was performed to collect the crude peptide precipitate, and the pellet was dried in vacuo. The peptide, dissolved in 0.1% aqueous TFA and MeCN (ca. 4:1), was further purified by preparative HPLC (10–40% B over 15 min, 10 mL/min) to obtain reduced peptide as a white powder following lyophilization. Oxidative folding was performed by dissolving 0.1 mg/mL reduced peptide in 0.1% acetic acid with 5% DMSO (v/v). The solution was adjusted to pH 8.0 with NH₄OH and incubated for 12 h at 23 °C. The peptide was directly purified by preparative HPLC (10–40% B over 15 min, 10 mL/min) and dried in vacuo. Methionine oxidation was achieved by incubating the synthetic peptides with 40 μ M H₂O₂ for 5 days at room temperature. Excess H₂O₂ was removed with C₁₈ ZipTips (Millipore-Sigma), and the purity was assessed via LC-MS/MS analysis.

Chromatographic Comparison of Synthetic and Native Peptide.

A. tricolor crude extract and synthetic peptide (hyP4, 5 μ g) were fractionated via SCX as previously described.³³ One-minute fractions were collected across the gradient, desalted via vacuum centrifugation, and profiled using LC-MS/MS analysis on the TripleTOF5600 platform as described above. MS data were processed using the label-free quantification method described above to quantify the abundance of Atr-AMP1 across fractions (SI Figure S9).

Simplification of Active Fractions.

Pooled active *A. tricolor* peptide library fractions (0.20 mL) were injected on an SCX column as previously described³³ collecting 1 min fractions across the gradient. Fractions were desalted with a Sep-Pak C₁₈ cartridge (Waters, 500 mg), concentrated to dryness, and resuspended in 0.20 mL LC-MS water for bioassay.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGMENTS

The authors would like to thank Carol Ann McCormick (NCU Herbarium, Chapel Hill, NC) for curation of *A. tricolor* specimens. The UNC High-Throughput Sequencing Facility assisted with sequencing. This work was supported by NIH-NIGMS under award number 1R01GM125814 to L.M.H. T.B.M. acknowledges support from the NSF Graduate Research Fellowship program (DGE-1650016) and L.R.H received support from a Summer Undergraduate Research Fellowship (UNC-CH).

REFERENCES

- (1). Tang S-S; Prodhan ZH; Biswas SK; Le C-F; Sekaran SD *Phytochemistry* 2018, 154, 94–105. [PubMed: 30031244]
- (2). Ríos JL; Recio MC J. *Ethnopharmacol* 2005, 100, 80–84. [PubMed: 15964727]
- (3). Brenner DM; Baltensperger DD; Kulakow PA; Lehmann JW; Myers RL; Slabbert MM; Sleugh BB *In Plant Breeding Reviews*; John Wiley & Sons, Inc.: Oxford, UK, 2010; pp 227–285.
- (4). Peter K; Gandhi P *Egypt. J. Basic Appl Sci* 2017, 4, 196–205.
- (5). Mowobi GG; Abubakar S; Osuji C; Etim VN; Ogechi N; Egeya JJ *Am. J. Phytochem. Clin. Ther.* 2016, 4, 73–90.
- (6). Bibi S; Sultana J; Sultana H; Malik RN J. *Ethnopharmacol.* 2014, 155, 352–361. [PubMed: 24882732]
- (7). Tugume P; Kakudidi EK; Buyinza M; Namaalwa J; Kamatenesi M; Mucunguzi P; Kalema JJ *Ethnobiol. Ethnomed.* 2016, 12, 5.
- (8). Agra M; de F; Silva KN; Basílio IJLD; de Freitas PF; Barbosa-Filho JM *Rev. Bras. Farmacogn.* 2008, 18, 472–508.
- (9). Baral M; Datta A; Chakraborty S; Chakraborty P; College BCR; Bengal W *Int. J. Appl. Biol. Pharm. Technol* 2011, 2, 41–47.
- (10). Hema ES; Sivadasan M; Anil KN *Ethnobotany* 2003, 18, 122–126.
- (11). Sreelatha S; Dinesh E; Uma C *Asian Pacific J. Cancer Prev.* 2012, 13, 2775–2780.
- (12). Islam M; Ali E; Saeed MA; Jamshaid M; Khan MTJ *Pak. J. Pharm.* 2003, 20, 2–3.
- (13). Maiyo ZC; Ngure RM; Matasyoh JC; Chepkorir R *Afr. J. Biotechnol.* 2010, 9, 3178–3182.
- (14). Lucero López V. R.; Razzeto GS; Giménez MS; Escudero NL *Plant Foods Hum. Nutr.* 2011, 66, 157–162. [PubMed: 21547507]
- (15). Zeashan H; Amresh G; Singh S; Rao CV *Pharm. Biol* 2010, 48, 1157–1163. [PubMed: 20860438]
- (16). Susantiningih T; Ridwan R; Prijanti AR; Sadikin M; Freisleben H-J *Med. J. Indones.* 2012, 21, 66.
- (17). Lipkin A; Anisimova V; Nikonorova A; Babakov A; Krause E; Bienert M; Grishin E; Egorov T *Phytochemistry* 2005, 66, 2426–2431. [PubMed: 16126239]
- (18). Broekaert WF; Marien W; Terras FRG; De Bolle MFC; Proost P; Van Damme, J.; Dillen L; Claeys M; Rees SB; Vanderleyden J; Cammue BPA *Biochemistry* 1992, 31, 4308–4314. [PubMed: 1567877]
- (19). Rivillas-Acevedo LA; Soriano-García MJ *Agric. Food Chem.* 2007, 55, 10156–10161.
- (20). Pribylova R; Kralik P; Pisarikova B; Pavlik I *Biologia (Warsaw, Pol.)* 2008, 63, 217–220.
- (21). Chopra, R. N.; Nayar, S. L.; Chopra, I. C.; Asolkar, L. V; Kakkar, K. K.; Chakre, O. J.; Varma, B. S.; (India), C. of S. & I. R Council of Scientific & Industrial Research: New Delhi, 1956.
- (22). Biswas B; Rogers K; McLaughlin F; Daniels D; Yadav A *Int. J. Microbiol.* 2013, 2013, 1–7.
- (23). Bihani GV; Bodhankar SL; Kadam PP; Zambare GN *Der Pharm. Lett.* 2013, 5, 48–55.
- (24). Devaraj VC *Zhongxiyi Jiehe Xuebao* 2011, 9, 1031–1038. [PubMed: 21906529]
- (25). Al-Dosari MS *Am. J. Chin. Med.* 2010, 38, 1051–1064. [PubMed: 21061460]
- (26). Rahmatullah M; Hosain M; Rahman S; Akter M; Rahman F; Rehana F; Munmun M; Kalpana M. a. *Afr. J. Tradit., Complementary Altern. Med.* 2013, 10, 408–411.
- (27). Amornrit W; Santiyanont R *GMR Genet. Mol. Res.* 2016, 15, 15.

- (28). Maldonado-Cervantes E; Jeong HJ; León-Galván F; Barrera-Pacheco A; De León-Rodríguez A; González de Mejía E.; de Lumen BO; Barba de la Rosa A. P. Peptides 2010, 31, 1635–1642. [PubMed: 20599579]
- (29). Silva-Sánchez C; de la Rosa APB; León-Galván MF; de Lumen BO; de León-Rodríguez A; de Mejía EG J. Agric. Food Chem. 2008, 56, 1233–1240. [PubMed: 18211015]
- (30). Tiengo A; Faria M; Netto FM J. Food Sci. 2009, 74, H121–H126. [PubMed: 19646044]
- (31). Barba de la Rosa A. P.; Barba Montoya A.; Martínez-Cuevas P; Hernández-Ledesma B; León-Galván MF; De León-Rodríguez A; González C Nitric Oxide 2010, 23, 106–111. [PubMed: 20435155]
- (32). Barrio DA; Añón MC Eur. J. Nutr. 2010, 49, 73–82. [PubMed: 19701660]
- (33). Kirkpatrick CL; Broberg CA; McCool EN; Lee WJ; Chao A; McConnell EW; Pritchard DA; Hebert M; Fleeman R; Adams J; Jamil A; Madera L; Strömstedt AA; Göransson U; Liu Y; Hoskin DW; Shaw LN; Hicks LM Anal. Chem. 2017, 89, 1194–1201. [PubMed: 27991763]
- (34). Parsley NC; Kirkpatrick CL; Crittenden CM; Rad JG; Hoskin DW; Brodbelt JS; Hicks LM Phytochemistry 2018, 152, 61–70. [PubMed: 29734037]
- (35). Kirkpatrick CL; Parsley NC; Bartges TE; Cooke ME; Evans WS; Heil LR; Smith TJ; Hicks LM J. Am. Soc. Mass Spectrom. 2018, 29, 859–865. [PubMed: 29404970]
- (36). Kirkpatrick CL; Parsley NC; Bartges TE; Wing CE; Kommineni S; Kristich CJ; Salzman NH; Patrie SM; Hicks LM Microb. Biotechnol. 2018, 11, 943–951. [PubMed: 30014612]
- (37). Frese CK; Altelaar AFM; van den Toorn H; Nolting D; Griep-Raming J; Heck AJR; Mohammed S Anal. Chem. 2012, 84, 9668–9673. [PubMed: 23106539]
- (38). Wang, G. Humana Press: New York, NY, 2015; pp 43–66.
- (39). Lagerwerf FM; van de Weert M; Heerma W; Haverkamp J Rapid Commun. Mass Spectrom. 1996, 10, 1905–1910. [PubMed: 9004526]
- (40). Yin-Shan Yang G. M.; Philippe Roch A. A. C. B. C. J. F. S.; Aumelas A Biochemistry 2000, 39, 14436–14447.
- (41). Nguyen GKT; Lim WH; Nguyen PQT; Tam JP J. Biol. Chem. 2012, 287, 17598–17607.
- (42). Schmitt P; Rosa RD; Destoumieux-Garzón D Biochim. Biophys. Acta, Biomembr. 2016, 1858, 958–970.
- (43). Cooper A; Eyles SJ; Radford SE; Dobson CM J. Mol. Biol. 1992, 225, 939–943. [PubMed: 1613799]
- (44). Haag AF; Kersch B; Dall'Angelo S; Sani M; Longhi R; Balaban M; Wilson HM; Mergaert P; Zanda M; Ferguson GP J. Biol. Chem. 2012, 287, 10791–10798.
- (45). Gorman JJ; Wallis TP; Pitt JJ Mass Spectrom. Rev. 2002, 21, 183–216. [PubMed: 12476442]
- (46). Boersema PJ; Raijmakers R; Lemeer S; Mohammed S; Heck AJR Nat. Protoc. 2009, 4, 484–494. [PubMed: 19300442]
- (47). Hsu JL; Chen SH Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences; The Royal Society, 10 2016.
- (48). Mingming M; Kutz-Naber KK; Li L Anal. Chem. 2007, 79 (2), 673–681. [PubMed: 17222036]
- (49). Xiao Y; Vecchi MM; Wen D Anal. Chem. 2016, 88, 10757–10766.
- (50). Lüders T; Birkemo GA; Nissen-meyer J; Andersen Ø; Nes IF; Lu T; Birkemo GA; Nissen-meyer J; Andersen Ø; Nes IF Antimicrob. Agents Chemother. 2005, 49, 2399–2406. [PubMed: 15917539]
- (51). Manabe T; Kawasaki K Sci. Rep. 2017, 7, 43384.
- (52). Kroeff EP; Pietrzyk DJ Anal. Chem. 1978, 50, 1353–1358.
- (53). Gesquiere JC; Diesis E; Cung MT; Tartar AJ. Chromatogr. A 1989, 478, 121–129.
- (54). van Huizen NA; Burgers PC; Saintmont F; Brocorens P; Gerbaux P; Stingl C; Dekker LJM; Izermans JNM; Luider TM J. Proteome Res. 2019, acs.jproteome.8b00930.
- (55). Wang G; Li X; Wang Z Nucleic Acids Res. 2016, 44, D1087–1093.
- (56). Mishra AK; Choi J; Moon E; Baek KH Molecules. Multidisciplinary Digital Publishing Institute (MDPI), 4 2018; p 815.
- (57). Cao H; Ke T; Liu R; Yu J; Dong C; Cheng M; Huang J; Liu S PLoS One 2015, 10, e0137414.

- (58). Gennaro R; Zanetti M; Benincasa M; Podda E; Miani M *Curr. Pharm. Des.* 2002, 8, 763–778. [PubMed: 11945170]
- (59). Casteels P; Ampe C; Jacobs F; Vaeck M; Tempst P *EMBO J.* 1989, 8, 2387–2391. [PubMed: 2676519]
- (60). Otvos L *Cell. Mol. Life Sci* 2002, 59, 1138–1150. [PubMed: 12222961]
- (61). Destoumieux D; Munoz M; Bulet P; Bachère E *Cell. Mol. Life Sci.* 2000, 57, 1260–1271. [PubMed: 11028917]
- (62). Stensvåg K; Haug T; Sperstad SV; Rekdal Ø; Indrevoll B; Styrvold OB *Dev. Comp. Immunol.* 2008, 32, 275–285. [PubMed: 17658600]
- (63). Wong KH; Tan WL; Serra A; Xiao T; Sze SK; Yang D; Tam JP *Front. Plant Sci.* 2016, 7, 7. [PubMed: 26870046]
- (64). Huang J; Wong KH; Tay SV; Serra A; Sze SK; Tam JP *J. Nat. Prod.* 2019, 82, 194–204. [PubMed: 30758201]
- (65). Astafieva AA; Enyenihi AA; Rogozhin EA; Kozlov SA; Grishin EV; Odintsova TI; Zubarev RA; Egorov TA *Plant Sci.* 2015, 238, 323–329. [PubMed: 26259198]
- (66). Polikanov YS; Aleksashin NA; Beckert B; Wilson DN *Front. Mol. Biosci.* 2018, 5, 48. [PubMed: 29868608]
- (67). Krizsan A; Knappe D; Hoffmann R *Antimicrob. Agents Chemother.* 2015, 59, 5992–5998. [PubMed: 26169420]
- (68). Paulsen VS; Mardirossian M; Blencke H-M; Benincasa M; Runti G; Nepa M; Haug T; Stensvåg K; Scocchi M *Microbiology* 2016, 162, 601–609. [PubMed: 26860543]
- (69). Mattiuzzo M; Bandiera A; Gennaro R; Benincasa M; Pacor S; Antcheva N; Scocchi M *Mol. Microbiol.* 2007, 66, 151–163. [PubMed: 17725560]
- (70). Graf M; Mardirossian M; Nguyen F; Seefeldt AC; Guichard G; Scocchi M; Innis CA; Wilson DN *Nat. Prod. Rep.* 2017, 34, 702–711. [PubMed: 28537612]
- (71). Vizcaíno JA; Csordas A; del-Toro N; Dianas JA; Griss J; Lavidas I; Mayer G; Perez-Riverol Y; Reisinger F; Ternent T; Xu Q-W; Wang R; Hermjakob H *Nucleic Acids Res.* 2016, 44, D447–D456. [PubMed: 26527722]
- (72). Baba T; Ara T; Hasegawa M; Takai Y; Okumura Y; Baba M; Datsenko KA; Tomita M; Wanner BL; Mori H *Mol. Syst. Biol.* 2006, 2, 2006.0008.
- (73). Kennedy AD; Otto M; Braughton KR; Whitney AR; Chen L; Mathema B; Mediavilla JR; Byrne KA; Parkins LD; Tenover FC; Kreiswirth BN; Musser JM; DeLeo FR *Proc. Natl. Acad. Sci. U. S. A.* 2008, 105, 1327–1332. [PubMed: 18216255]
- (74). Lawlor MS; Hsu J; Rick PD; Miller VL *Mol. Microbiol.* 2005, 58, 1054–1073. [PubMed: 16262790]
- (75). Tavares LS; Santos M; de O; Viccini LF; Moreira JS; Miller RNG; Franco OL *Peptides* 2008, 29, 1842–1851. [PubMed: 18602431]
- (76). Meneguetti BT; Machado L; dos S; Oshiro KGN; Nogueira ML; Carvalho CME; Franco OL *Front. Microbiol.* 2017, 7, 7.
- (77). Maróti G; Kereszt A; Kondorosi É; Mergaert P *Res. Microbiol.* 2011, 162, 363–374. [PubMed: 21320593]
- (78). Matasci N; Hung LH; Yan Z; Carpenter EJ; Wickett NJ; Mirarab S; Nguyen N; Warnow T; Ayyampalayam S; Barker M; Burleigh JG; Gitzendanner MA; Wafula E; Der JP; DePamphilis CW; Roure B; Philippe H; Ruhfel BR; Miles NW; Graham SW; Mathews S; Surek B; Melkonian M; Soltis DE; Soltis PS; Rothfels C; Pokorny L; Shaw JA; DeGironimo L; Stevenson DW; Villarreal JC; Chen T; Kutchan TM; Rolf M; Baucom RS; Deyholos MK; Samudrala R; Tian Z; Wu X; Sun X; Zhang Y; Wang J; Leebens-Mack J; Wong GKS *GigaScience; BioMed Central*, 12 2014; p 17.
- (79). Lam T-W; Tang J; Wu G; Li Y; He G; Wang J; Gu S; Xie Y; Huang W; Li S; Zhou X; Wong GK-S; Patterson J; Luo R; Xu X; Liu S *Bioinformatics* 2014, 30, 1660–1666. [PubMed: 24532719]
- (80). Simon MD; Heider PL; Adamo A; Vinogradov AA; Mong SK; Li X; Berger T; Policarpo RL; Zhang C; Zou Y; Liao X; Spokoiny AM; Jensen KF; Pentelute BL *ChemBioChem* 2014, 15, 713–720. [PubMed: 24616230]

Figure 1.

Overview of the workflow used for the identification, isolation, and characterization of a novel antimicrobial peptide from *A. tricolor*. (A) Aerial tissue was extracted in an acetic acid solution with size exclusion steps, subjected to SCX cleanup to remove small molecules, and (B) fractionated via reversed-phase HPLC, generating a peptide library. (C) The peptide library was analyzed via LC-MS/MS, screened for activity against *E. coli* ATCC 25922, and (D) the PepSAVI-MS statistical analysis package was used to identify a likely contributor to activity, Atr-AMP1, whose sequence was (E) evaluated through derivatization experiments and (F) determined de novo.

Figure 2.

(A) Reversed-phase HPLC chromatogram (220 nm) of *A. tricolor* crude extract. Fractions were collected every 0.5 min across the gradient to create a peptide library. The region corresponding to active fractions is highlighted in blue. (B) The peptide library demonstrated strong activity against *E. coli* ATCC 25922 across fractions 31–35. The elution of the top 20 likely contributors to bioactivity identified after statistical modeling is overlaid. Features corresponding to charge states of Atr-AMP1 are denoted with an asterisk.

Figure 3.

Atr-AMP1 spectral annotation. The +4 charge state of reduced and alkylated Atr-AMP1, m/z 455.97, was fragmented using EThcD (ETD 44.59, HCD 40.00) and sequenced de novo using PEAKS Studio. The highest confidence sequence is annotated above. All cysteine units are alkylated with IAM, oxidized residues are noted with “OX”, and hydroxylated residues are noted with “HY”.

Figure 4.

CID fragmentation of the *N*-terminal tryptic peptides produced by Atr-AMP1 confirmed the identities of the first seven residues and revealed two distinct populations of peptidofoms. The first population (A) contains only an oxidized methionine. The second population (B), contains an oxidized methionine and a hydroxyproline at position four. The partial *y*-ion series that differentiate each hydroxylation peptidofom are highlighted such that each colored box marks the mass difference between diagnostic ions. All cysteines are alkylated with IAM, oxidized residues are noted with “OX”, and hydroxylated residues are noted with “HY”.

Figure 5.

CID fragmentation of the C-terminal chymotryptic peptides produced by Atr-AMP1 confirmed the identities of the final eight residues and revealed three different populations of post-translationally modified peptides. The first population (A) contained no post-translational modifications. The second two populations (B, C) were peptidofoms varying only in the position of the hydroxyproline moiety. The partial y-ion series that differentiate each hydroxylation peptidofom are highlighted such that each colored box marks the mass

difference between diagnostic ions. All cysteine residues are alkylated with IAM and hydroxylated residues are noted with “HY”.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

Figure 6.

A simplified fraction containing Atr-AMP1 was screened against additional bacterial species. (A) The fraction was active against *E. coli* ATCC 25922 and methicillin-resistant *S. aureus* LAC but inactive against rifampin- and streptomycin-resistant *K. pneumoniae* VK148. (B) The fraction was also screened against single gene *E. coli* knockouts of genes coding for proteins involved in transporting PrAMPs into the bacterial cell. No increase in activity was

observed between the wild type (BW25113) and transporter knockouts (JW2778 and JW0368).

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript