The genetic origin of evolidine, the first cyclopeptide discovered in plants, and related orbitides

Received for publication, June 12, 2020, and in revised form, August 11, 2020 Published, Papers in Press, August 19, 2020, DOI 10.1074/jbc.RA120.014781

Mark F. Fisher[1], Colton D. Payne[2], Thaveshini Chetty[1], Darren Crayn[3], Oliver Berkowitz[4], James Whelan[4], K. Johan Rosengren[2], and Joshua S. Mylne[1,*]

From the [1]University of Western Australia, School of Molecular Sciences & The ARC Centre of Excellence in Plant Energy Biology, Crawley, Western Australia, Australia, the [2]University of Queensland, Faculty of Medicine, School of Biomedical Sciences, Brisbane, Queensland, Australia, the [3]Australian Tropical Herbarium, James Cook University, Cairns, Queensland, Australia, and the [4]Department of Animal, Plant and Soil Sciences, School of Life Sciences & ARC Centre of Excellence in Plant Energy Biology, AgriBio, The Centre for AgriBioscience, La Trobe University, Bundoora, Victoria, Australia

Edited by Joseph M. Jez

Cyclic peptides are reported to have antibacterial, antifungal, and other bioactivities. Orbitides are a class of cyclic peptides that are small, head-to-tail cyclized, composed of proteinogenic amino acids and lack disulfide bonds; they are also known in several genera of the plant family Rutaceae. *Melicope xanthoxyloides* is the Australian rain forest tree of the Rutaceae family in which evolidine, the first plant cyclic peptide, was discovered. Evolidine (cyclo-SFLPVNL) has subsequently been all but forgotten in the academic literature, so to redress this we used tandem MS and *de novo* transcriptomics to rediscover evolidine and decipher its biosynthetic origin from a short precursor just 48 residues in length. We also identified another six *M. xanthoxyloides* orbitides using the same techniques. These peptides have atypically diverse C termini consisting of residues not recognized by either of the known proteases plants use to macrocyclize peptides, suggesting new cyclizing enzymes await discovery. We examined the structure of two of the novel orbitides by NMR, finding one had a definable structure, whereas the other did not. Mining RNA-seq and whole genome sequencing data from other species of the Rutaceae family revealed that a large and diverse family of peptides is encoded by similar sequences across the family and demonstrates how powerful *de novo* transcriptomics can be at accelerating the discovery of new peptide families.

Orbitides are a group of cyclic peptides found in plants. They contain between five and 16 amino acid residues, are homodetic cycles, and have no disulfide bonds (1, 2). Orbitides are thought to be ribosomally synthesized and post-translationally modified peptides. Their genetic origins are known only for a few groups: the segetalins of *Vaccaria hispanica* (3), some linusorbs (previously cyclolinopeptides) from flax (*Linum usitatissimum*) (4), two curcacyclines from *Jatropha curcas* (1), a number of orbitides in *Citrus* species (3, 5), the annomuricatins of *Annona muricata* (6) and the PawL-derived peptides (PLPs) from members of the daisy family (Asteraceae) (7–9).

Evolidine is an orbitide from the leaves of *Melicope xanthoxyloides* (F. Muell.) T.G. Hartley (synonym *Euodia xanthoxyloides* F. Muell.), a species of the Rutaceae family, which is found in the rainforest of northern Queensland in Australia, the Lesser Sunda Islands, the Moluccas, Timor Leste, and New Guinea (10–12). Evolidine was first isolated at the University of Sydney, Australia, in 1952 (13), and its amino acid content and cyclic nature were determined in 1955 (14). The amino acid sequence was confirmed in 1961 (15) and the first synthesis of the peptide was carried out in 1965 (16). A number of subsequent studies of the structure of evolidine have been carried out, using both NMR (17, 18) and X-ray crystallography (Fig. 1) (19). In 2005, evolidine was reported to have antibacterial and antifungal activity (20).

Despite being the first cyclic peptide isolated from a plant, evolidine has been overlooked by the natural product literature. It was not included in the most comprehensive review of plant cyclic peptides to date (21). Instead, perhaps because of the economic importance of flaxseed, from which it comes, cyclolinopeptide A is often cited as the first cyclic peptide (or sometimes the first orbitide) discovered in a plant (1, 4, 21–25). However, cyclolinopeptide A was not isolated until 1959 (26) and its chemical structure solved only in 1966 (27).

We were interested in investigating the biosynthesis of evolidine, the first plant cyclopeptide, and so we assembled a *de novo* transcriptome for leaves of *M. xanthoxyloides.* We found mRNA transcripts that encode evolidine and used MS to rediscover evolidine plus confirm six novel orbitides encoded by similar transcripts. The C termini of these core peptides are Phe, Lys, Leu, and Ser, none of which are residues recognized by either of the known plant macrocyclizing enzymes, namely asparaginyl endopeptidase (which recognizes Asp or Asn in the P1 position) or prolyl oligopeptidase (which recognizes Pro or Ala). Two of the novel orbitides were examined by NMR spectroscopy and their behavior varied in solution. Using the precursor sequences from *M. xanthoxyloides*, we analyzed RNA-seq and whole genome sequencing (WGS) data from the Sequence Read Archive (28) (SRA) and found genes and transcripts encoding putative peptides in four genera from the Aurantioideae subfamily of the Rutaceae.

Results

We first confirmed the presence of evolidine in our tissue sample by LC–MS. The peptide was sequenced by tandem MS

This article contains supporting information.

* For correspondence: Joshua S. Mylne, joshua.mylne@uwa.edu.au.

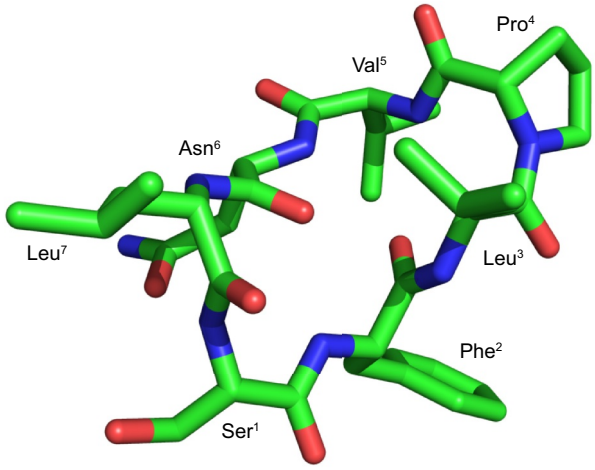

Figure 1. Structure of evolidine. Deposition number 1266027 at the Cambridge Structural Database, obtained by X-ray crystallography (19).

(MS/MS) (Fig. S1). The observed *m/z* and lack of *y* ions showed that the peptide was cyclic. To find the transcript that encodes evolidine, we extracted total RNA from *M. xanthoxyloides* leaves, performed RNA-seq, and assembled a transcriptome using established methods (7). We searched the transcriptome with tBLASTn for sequences encoding evolidine by using five cyclic peptide precursor sequences from *Citrus* species (translated from GenBank® accession numbers EF175924–EF175297 and FC872925); we reasoned that these could share sequence similarity with the putative gene coding for evolidine because both *Melicope* and *Citrus* belong to the Rutaceae family. We found two sequences potentially coding for evolidine along with ~13 similar, putative peptide-encoding sequences. Aligning these sequences showed absolute conservation for the residues flanking evolidine (Glu and Ser), which indicated where the peptide sequence might lie.

Confirmation of further cyclic peptides by tandem MS and Sanger sequencing

Having the sequences for transcripts encoding putative novel cyclic peptides facilitated their identification and sequencing from LC–MS/MS data. We found peaks corresponding to six additional peptides of ~13 predicted by transcriptomic data and sequenced the peptides by MS/MS (Table 1). We named these peptides xanthoxycyclins A to F, and the corresponding genes *Proxanthoxyclins A* to *F* (Fig. 2). The LC–MS data for xanthoxycyclin A are shown in Fig. 3; data for xanthoxycyclins B to F are shown in Fig. S3–S7.

The sequences of the core peptides lack conservation except some at the N- and C-terminal residues. The N-terminal residues were ones with small side chains, namely Gly, Ser, and Ala. The C-terminal residue was one of four possibilities: Phe or Leu, Lys for xanthoxycyclin C, and Ser for xanthoxycyclin F. In the sequence preceding the peptides, Glu is absolutely conserved in the P1 position and the P2 residue is a highly-conserved Ser, although for the two evolidine transcripts a Trp is encoded. There are many other highly conserved residues, especially in the sequence following the core peptide, where the P1′ residue is always Ser and P2′ is most often Ile (Fig. 2). The core peptides described above are typical orbitides in that they contain mainly hydrophobic residues (1). With between six and eight residues, they fall into the low-middle range of sizes for orbitides, which can be as large as 16 amino acids (2).

We confirmed the sequences obtained by transcriptome assembly and LC–MS/MS by designing primers against the ends of each contig to amplify either genomic DNA, or a full-length coding sequence using cDNA produced from poly-T-primed RNA as the template. We were able to confirm one of the *Proevolidine* sequences (Fig. 4), and *Proxanthoxycyclin E* (Fig. S2*A*) from genomic DNA. *Proxanthoxycyclins B–D* were sequenced from cDNA and found to be a 100% match to the contigs assembled by RNA-seq. It was not possible to Sanger sequence *Proxanthoxycyclin A* due to the coding sequence being almost identical to that of *Proxanthoxycyclin B* and the repetitive nature of one of the regions to be primed. We were also unable to PCR amplify the putative second *Proevolidine* gene or clone a sequence for *Proxanthoxycyclin F*.

It is perhaps worth noting that when amplifying *Proxanthoxycyclin E* from genomic DNA template, as well as finding the leaf peptide sequence we were searching for, we also amplified six other sequences using the same primers (Fig. S2*B*). Five of these were complete ORFs with a high degree of sequence similarity to the *Proxanthoxycyclins*; another appeared to be a truncated form. There was no evidence to support these in the leaf MS/MS data, but these genes related to *Proxanthoxycyclin* might not be expressed or could encode xanthoxycyclins that are expressed in other tissues or specific growth stages not represented within the leaf sample we analyzed.

Potential peptide-encoding transcripts and genes in other Rutaceae species

To investigate the prevalence of this kind of peptide-encoding transcript within the Rutaceae, we searched the SRA to find RNA-seq and WGS data for other species in the family. We downloaded and assembled RNA-seq data from *Aegle marmelos* fruit pulp, *Atalantia buxifolia* mature fruit, *Bergera koenigii* leaf, *Clausena excavata* leaf, *Zanthoxylum armatum* young leaf and *Zanthoxylum bungeanum* young leaf, as well as WGS data for *A. buxifolia* and *Clausena lansium*. No peptide-encoding transcripts were detected in either *Zanthoxylum* species, despite their being closely related to *Melicope* and there being three known orbitides in three different species in the genus *Zanthoxylum* (29–31). Putative peptide-encoding transcripts were found in all other RNA-seq datasets examined (Fig. 5), which come from five species in four genera. All of these species are in the Aurantioideae subfamily of the Rutaceae, to which *Citrus* also belongs.

Like those of *M. xanthoxyloides*, the potential propeptides encoded within these five species also show a high sequence similarity outside the core peptide sequence. Within the core peptide, Gly is the most common residue at the N terminus, but the C-terminal residues are even less conserved than the core peptides of *M. xanthoxyloides*. In *A. marmelos*, Phe and Leu are most common, as in *M. xanthoxyloides*, but the other three species have a wide range of C-terminal residues, including Val, Thr, Trp, Pro, and Ser, among others. It must be noted,

Table 1

Summary of *M. xanthoxyloides* cyclic peptides, the genes that encode each and evidence for the peptides from mass spectrometry

The peptides are all backbone cyclic, with the sequences beginning with their proto-N terminus and ending with their proto-C terminus that are joined during processing.

Peptide	Sequence	Gene(s)	Peptide evidence
Evolidine	SFLPVNL	*Proevolidine1, Proevolidine2*	Eastwood *et al.* (14), Fig. S1
Xanthoxycyclin A	GFAILF	*Proxanthoxycyclin A*	Fig. 4
Xanthoxycyclin B	GFAILFF	*Proxanthoxycyclin B*	Fig. S3
Xanthoxycyclin C	GYAPFPFK	*Proxanthoxycyclin C*	Fig. S4
Xanthoxycyclin D	GTVAVQFL	*Proxanthoxycyclin D*	Fig. S5
Xanthoxycyclin E	SSSPPIF	*Proxanthoxycyclin E*	Fig. S6
Xanthoxycyclin F	AFVVPPRIS	*ProxanthoxycyclinF*	Fig. S7

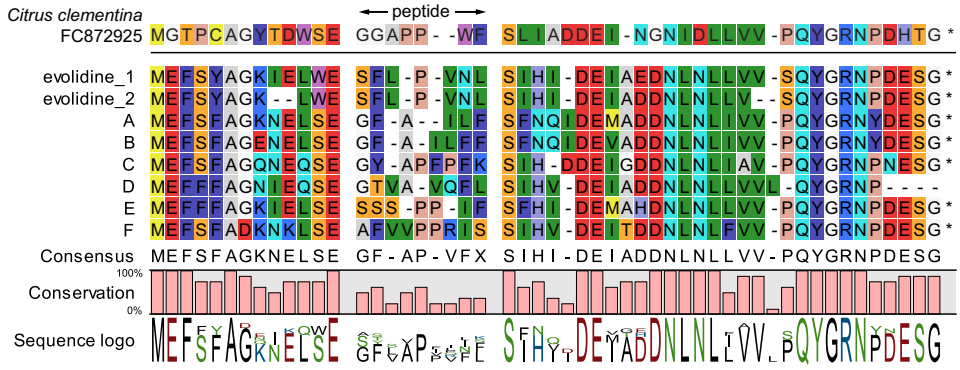

Figure 2. *M. xanthoxyloides* peptide precursors aligned with a known *Citrus* orbitide precursor. The translated coding sequences for seven cyclic peptides in *M. xanthoxyloides* aligned to *C. clementina* GenBank accession FC872925. The letters *A* to *F* stand for *Proxanthoxycyclins A* to *F*.

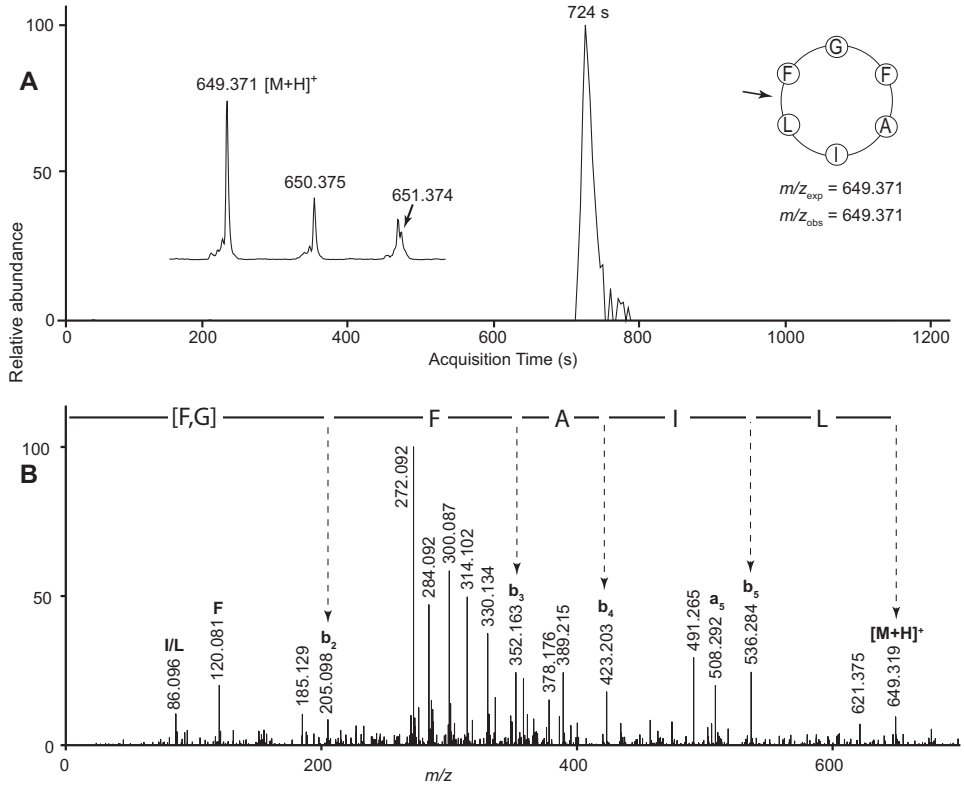

Figure 3. LC–MS/MS data for xanthoxycyclin A. *A,* extracted ion chromatogram at *m/z* 649.371. *Inset left*: mass spectrum at the chromatogram peak. *Inset right*, cyclic peptide sequence with MS/MS ring cleavage point shown by an *arrow*, expected *m/z* (m/z_{exp}) and observed *m/z* (m/z_{obs}). *B,* MS/MS spectrum of protonated molecule at *m/z* 649.371. A strong sequence of *b* ions is derived from initial ring opening between Leu and Phe residues. Immonium ions are labeled with the corresponding amino acid residue.

however, that there is no peptide evidence yet to indicate that these encoded sequences are actually processed into orbitides, so it would be unwise to speculate overly; indeed some sequences appear to encode cysteine residues that, to our knowledge, have never been found in an orbitide. The high degree of conservation, with up to 50% identity and 60% similarity between

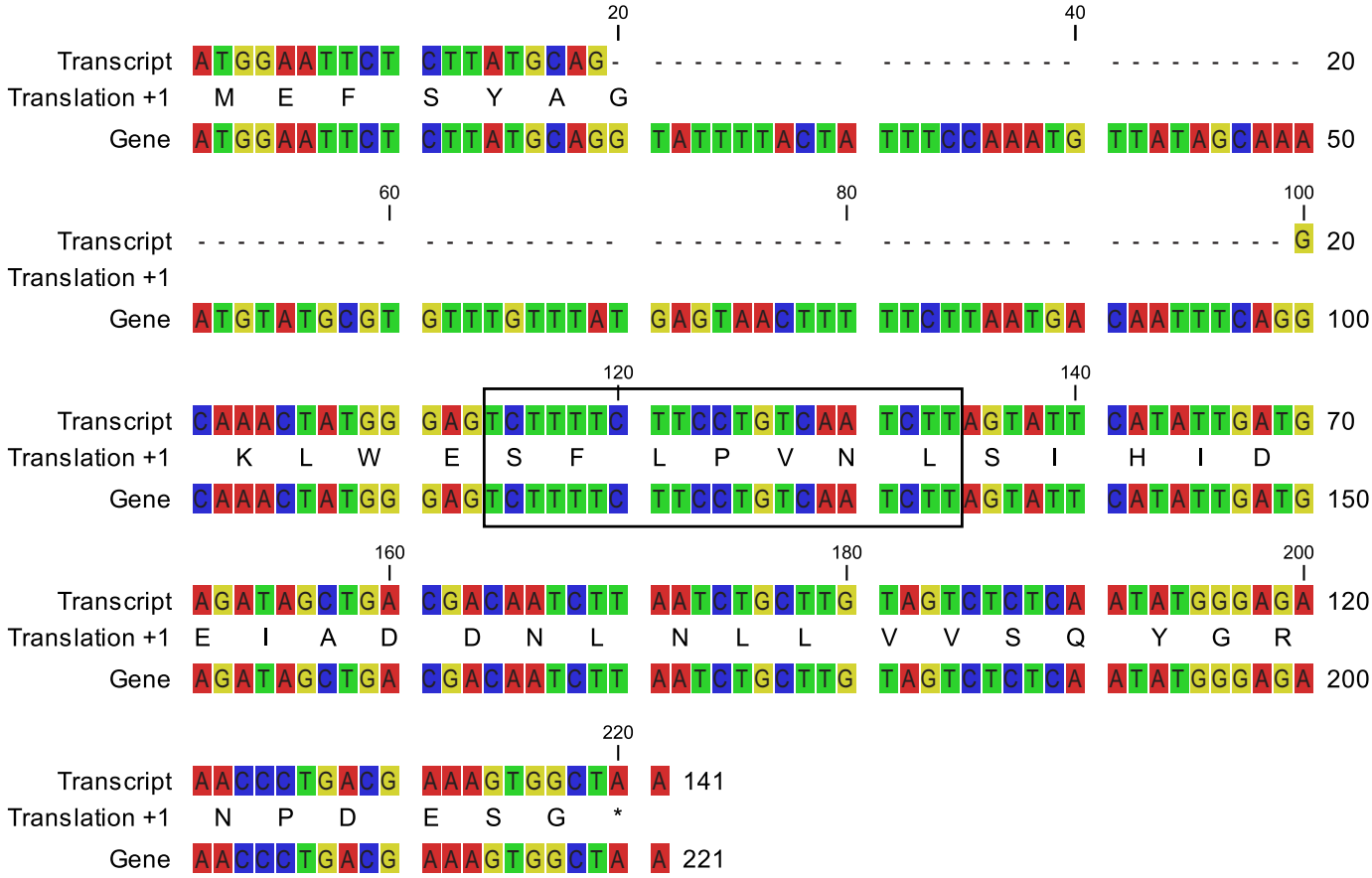

Figure 4. Alignment of *Proevolidine* transcript and gene. The Sanger-sequenced gene data matches perfectly with the RNA-seq transcript, except that the gene contains an 80-bp intron. The translated peptide sequence is shown between the nucleotide sequences; evolidine is shown *within a box*.

the flanking regions of the *M. xanthoxyloides* propeptides and the *Citrus clementina* sequence FC872925, suggests that these regions are important for peptide processing.

WGS data from the species *A. buxifolia* and *C. lansium* contained many potential peptide-encoding sequences (Figs. S8 and S9, respectively), although it was not possible to say how many might be expressed without transcriptomic data. One of the sequences from *C. lansium* appeared to encode the known peptide clausenlanin A (cyclo-GLILLLLLL) previously found in that species (32). Another sequence potentially encoding a peptide from *C. lansium*, clausenlanin B (cyclo-GLVLLLLLL) (32), was found, but not in the data for that species; rather, it was found in WGS data from *A. buxifolia*. The two putative propeptide sequences show a remarkably high degree of conservation, differing only at three residues outside the core peptide sequence (Fig. S10). The strong degree of sequence similarity in the propeptides shown among the species studied here and *Citrus* species investigated elsewhere (3, 5), especially in the sequence following the core peptide, suggests a common point of evolution for all these peptides across the Rutaceae.

NMR spectroscopy of synthetic xanthoxycyclins

Orbitides, due to their small size and restrained nature, contain a limited variety of structural features. The majority of orbitides are five to 10 residues with structural features being limited to ordered loops (2), turns (33, 34), and mini-β-sheets (35). Despite over 100 orbitides being discovered and reported in CyBase (36), fewer than 10 3D structures have been solved using either solution NMR spectroscopy or X-ray crystallography and submitted to the Protein Data Bank (PDB) (36).

To explore the structural features of orbitides, we subjected synthetic xanthoxycyclin D and xanthoxycyclin F to solution NMR spectroscopy. The 1D NMR spectrum of xanthoxycyclin D demonstrated sharp peaks with good signal dispersion in H_2O/D_2O. A set of 2D NMR spectra, including TOCSY (total correlation spectroscopy) (37) and NOESY (nuclear Overhauser effect spectroscopy) (38), were recorded and assigned using a sequential assignment strategy. The NOESY spectrum contained a large number of medium range NOEs (Table 2), consistent with the peptide adopting an ordered structure. Xanthoxycyclin D demonstrated relatively small deviations from random coil Hα shifts (Fig. S11), however, deviations in Cα and Cβ chemical shifts (Fig. S11) allowed for prediction of backbone dihedral angles by TALOS-N (torsion angle likelihood obtained from shift and sequence similarity) (39) for seven of the eight residues.

The dihedral angle restraints and NOE-based distance restraints were used as input for simulated annealing and energy minimization in explicit water, and resulted in a tightly overlapping structural ensemble with a root mean square deviation of 0.20 Å for the backbone atoms and 0.93 Å for all heavy atoms (Fig. 6). Only one residue, Gln-6, appears to have

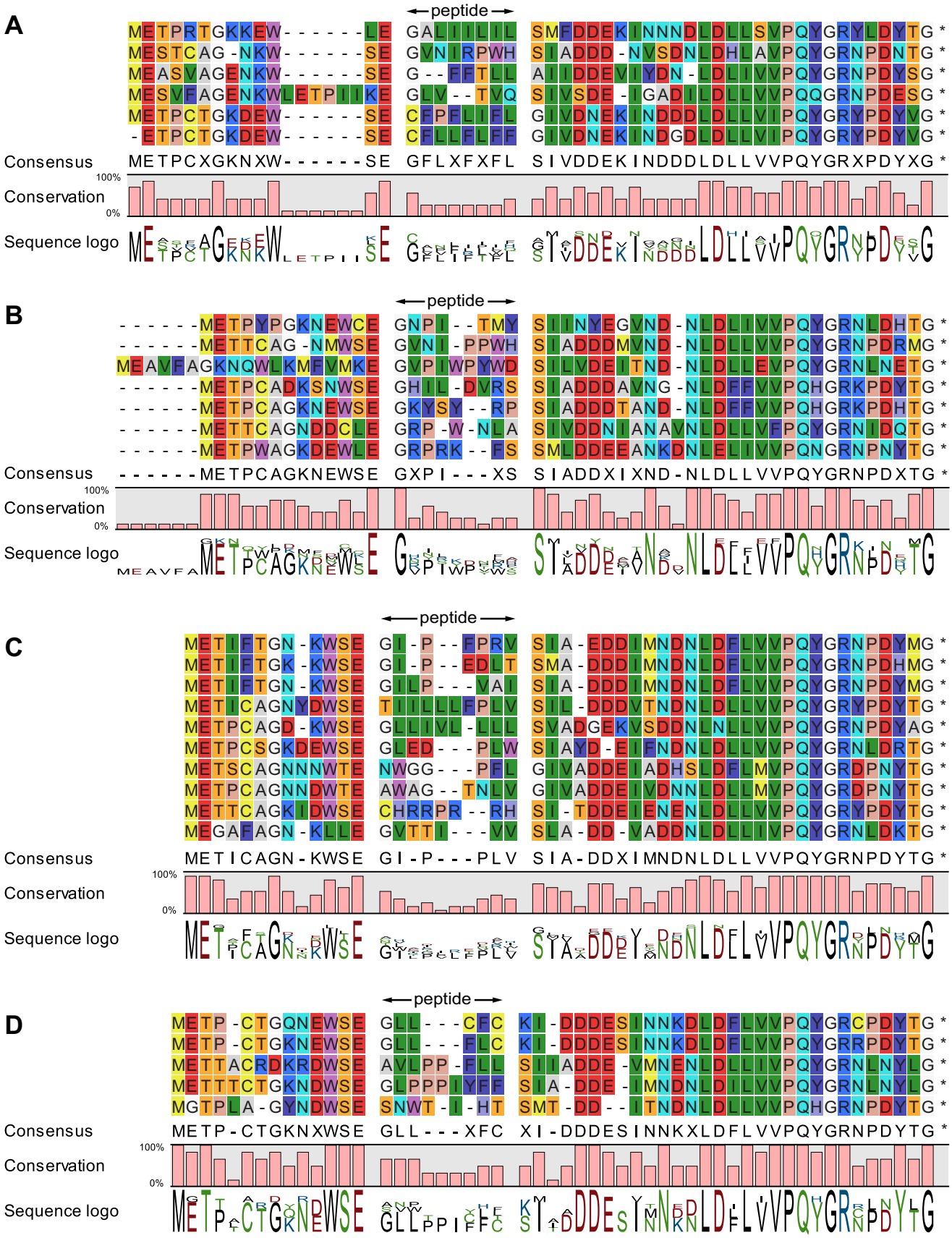

Figure 5. Alignments of putative propeptide sequences from four Rutaceae species. Sequences are from (*A*) *A. marmelos*; (*B*) *A. buxifolia*, (*C*) *B. koenigii*, and (*D*) *C. excavata*. Sequences are shown with background RasMol colors, and the putative peptide is marked and spaced from the surrounding sequence. *Asterisks* represent stop codons. Below each alignment is the consensus sequence, the degree of conservation and a sequence logo.

Table 2
NMR structure statistics of xanthoxycyclin D

Energies (kcal/mol)	
Overall	−225.70 ± 0.18
Bonds	1.88 ± 0.01
Angles	5.43 ± 0.06
Improper	2.16 ± 0.03
Dihedral	34.76 ± 0.03
Van der Waals	−16.82 ± 0.08
Electrostatic	−253.16 ± 0.19
NOE	0.006 ± 0.001
Cdih	0.036 ± 0.002
MolProbity statistics	
Clashes (>0.4 Å/1000 atoms)	0
Poor rotamers	0
Ramachandran outliers (%)	0
Ramachandran favored (%)	100 ± 0
MolProbity score	0.5 ± 0
MolProbity score percentile	100 ± 0
Atomic root mean square deviations (Å)	
Mean global backbone	0.20 ± 0.06
Mean global heavy	0.93 ± 0.24
Experimental restraints	
Distance restraints	
Short range (i-j < 2)	50
Medium range (/i-j/ < 5)	24
Long range (/i-j/ > 5)	0
Hydrogen bond restraints	8 (4 H-bonds)
Total	82
Dihedral angle restraints	
ϕ	7
ψ	7
$\chi 1$	1
Total	15
Violations from experimental restraints	
Total NOE violations exceeding 0.2 Å	0
Total dihedral violations exceeding 2.0°	0

complete conformational freedom of its side chain. The side chain of Phe-7 adopts two conformational states; these states are reflected in the NMR data, which show a large chemical shift separation of the two Hβ protons but averaged $^{3}J_{H\alpha H\beta}$ coupling constants. The side chain of Leu-8 is locked into a singular conformation (gauche-trans); again reflected in the NMR data, which shows a slightly weaker NH-Hβ2 NOE compared with the NH-Hβ3 NOE and a large difference in $^{3}J_{H\alpha H\beta}$-coupling constants, with the coupling between Hα and Hβ2 being significantly weaker than that between Hα and Hβ3 (40).

The structure of xanthoxycyclin D has two notable features. First, a typical type III β-turn between residues Val-5 and Leu-8. The type III turn is defined by a distance between the Cα carbons of Val-5 and Leu-8 (i and $i + 3$) of 4.5 Å, with the ϕ and ψ dihedral angles of Gln-6, −62°, −34° ($i + 1$), and Phe-7, −60°, −31° ($i + 2$), respectively. The turn is stabilized by a hydrogen bond between the NH of Leu-8 and CO of Val-5. Second, an atypical type III β-turn is formed between Thr-2 and Val-5. Although the distance between the Cα carbons of Thr-2 and Val-5, which is 5 Å, is consistent with a turn, and the ϕ and ψ dihedral angles of Val-3, −79°, −31° ($i + 1$), and Ala-4, −79°, −39° ($i + 2$) are similar to the common type III β-turn angles, the hydrogen bond which stabilizes this turn is formed between the HN of Val-5 and the side chain Oγ1 of Thr-2 rather than the backbone carbonyl. Hydrogen bonds are also present between Gly-1 HN–Gln-6 O and Thr-2 HN–Val-5 O. Additionally, the hydroxyl proton Hγ1 of Thr-2 is within hydrogen bond distance of the carbonyl of Val-5 in half of the models of the structural ensemble. A weak resonance at 5.15 ppm was indeed detected and assigned to Thr-2 Hγ1 in the NMR data, which is consistent with it being involved in a hydrogen bond. However, this implies that there are three protons potentially competing to form hydrogen bonds with the same carbonyl oxygen, thus rearrangements allowing this sharing of acceptor would likely be required.

Xanthoxycyclin F was also studied using solution NMR spectroscopy. In contrast to xanthoxycyclin D, the 1D spectrum was of poor quality, and despite the appearance of a single pure compound on HPLC analysis, the peptide adopted multiple conformations in solution, evident from the numerous NMR peaks observed. The 2D TOCSY and NOESY data confirmed numerous conformations at almost every residue, presenting a challenge for the assignment of the data. Despite this, a dominant conformation was able to be assigned and confirmed using a natural abundance ^{13}C HSQC (heteronuclear single quantum coherence spectroscopy). The secondary shifts of the dominant confirmation were determined for all residues of xanthoxycyclin F (Fig. S12). It demonstrated secondary shifts that deviate from random coil for most residues, but the complete lack of any nonsequential NOEs, coupled with the complexity of separating information from overlapping resonances from different conformations, prevented sufficient data from being derived to determine the structure of any of the conformers. A previous study showed that some members of the PLP subfamily of orbitides (PLPs-2, -4, -10, and -12) also demonstrate secondary Hα shifts that deviate from random coil (8), indicating at least some degree of ordered backbone, yet they displayed only irregular nonrigid structures lacking classical secondary structural features.

Biological activities of M. xanthoxyloides orbitides

Poojary and Belagali (20) reported that evolidine inhibits the growth of some bacteria and fungi. We attempted to replicate these results and to extend the study to xanthoxycyclins D and F, using disc diffusion assays like those in the prior work. Our synthetic peptides were generated in house by solid phase peptide synthesis or purchased so that we had enough material for NMR structure determination and to test the antifungal and antibacterial activity of the peptides. LC–MS/MS was used to compare the properties of the natural and synthetic compounds to verify that they were identical. The LC retention times were compared and mirror plots of the MS/MS spectra for each natural/synthetic pair were made. We concluded that the natural and synthetic compounds were identical. This was done for all three synthetic peptides: evolidine (Fig. S13), xanthoxycyclin D (Fig. S14), and xanthoxycyclin F (Fig. S15).

Our results showed no growth inhibition by any of the three peptides against the Gram-negative bacterium *Escherichia coli* K-12 nor against the Gram-positive bacterium *Bacillus subtilis* Marburg No. 168, whereas a positive control disc containing 50 μg of kanamycin showed a clear zone of inhibition in every case (Fig. S16). Similarly, the peptides were not active against either *Candida albicans* or *Aspergillus fumigatus* (Fig. S17), although a positive control disc containing 50 μg of amphotericin B inhibited fungal growth. We attempted to use griseofulvin as a positive control to follow the method of Poojary and Belagali,

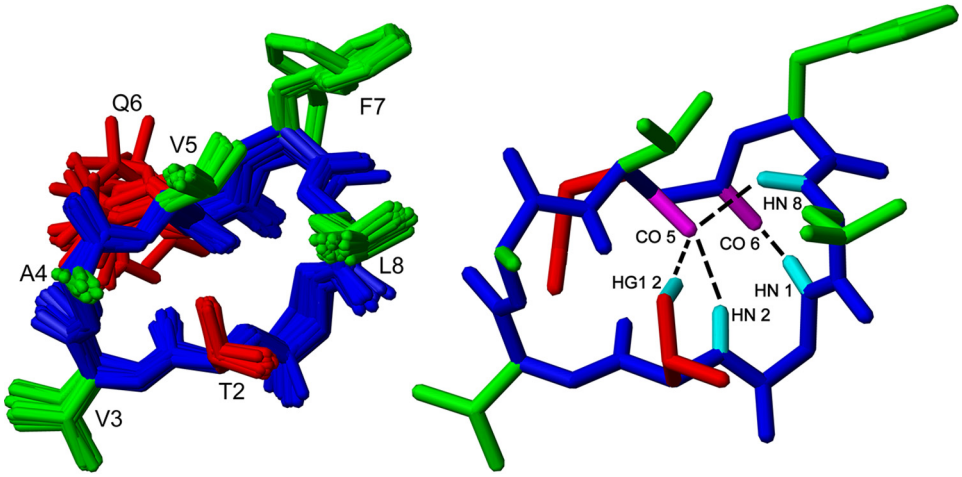

Figure 6. NMR solution structure of xanthoxycyclin D. *Left image* is of the structural ensemble of 20 best structures superimposed, shown in stick format. *Right image* is a single structure of xanthoxycyclin D in stick format illustrating hydrogen bonding with *dashed lines*. Hydrogen bond donors are shown in *cyan*, with hydrogen bond acceptors being shown in *magenta*. Both *left* and *right images* show the peptide backbone in *blue*, hydrophobic side chains in *green* and polar side chains in *red*.

but found it had no effect on either fungal species, even with a quantity four times that used in the earlier study (Fig. S18).

Discussion

Here we have shown that evolidine and six novel orbitides, xanthoxycyclins A–F, in the leaves of *M. xanthoxyloides* are genetically encoded and we have solved the solution structure of xanthoxycyclin D by NMR spectroscopy. Using public RNA-seq and WGS data from the SRA, we have also identified similar genes and transcripts in other plants from the Aurantioideae subfamily of the Rutaceae, which may encode comparable cyclic peptides, perhaps from a single evolutionary origin. Our ability to discover all these sequences by searching with peptide precursors from *Citrus* species points to a common ancestor and orbitides made from precursors like these may be present in many of the ~1,900 species of the Rutaceae (41). A study of other species across this family could reveal more such orbitides that might have interesting pharmacological properties.

Structure of xanthoxycyclin D

The 3D structure was determined using solution NMR spectroscopy. We showed that xanthoxycyclin D adopts a well-defined structure in solution, comprising two distinct turns. Although it is one residue longer than evolidine and with quite a different sequence, the structures share some general features that highlight the nature of the orbitide fold. Evolidine also comprises two ordered β-turns involving residues Phe-2–Val-5 and Asn-6–Phe-2 (19). The conformation is stabilized by hydrogen bonds involving both backbone and side chain atoms. Evolidine-like xanthoxycyclin D contains a bifurcated hydrogen bond in which the Phe-2 carbonyl serves as an acceptor for the amide protons of both Val-5 and Asn-6. Both structures lack substantial hydrophobic interactions, with the small ring size resulting in the majority of side chains projecting away from the core. Xanthoxycyclin F in contrast does not adopt a single well-defined structure in solution, instead adopting multiple conformations, as evident from a large number of additional resonance signals. Thus, not all sequences are able to support a single stable fold.

Bioactivity of evolidine

Our bioactivity results for evolidine were in contrast to the prior work (20), which found that evolidine inhibited growth of *B. subtilis* (strain not specified), and the fungi *C. albicans* and *Aspergillus niger* using a similar assay to ours. The same study was consistent with our finding that *E. coli* is unaffected by evolidine. The prior study did not state what quantities of evolidine were applied to the discs, preventing a direct comparison of the results. This is not the first example of cyclic peptide bioactivity being nonreproducible, although this occurred mostly when synthetic compounds were compared with natural products. This was previously observed in yunnanins A and C (orbitides from *Stellaria yunnanensis*), where the synthetic compounds had activity orders of magnitude lower than the natural products (42). The authors attributed this either to conformational differences between the natural and synthetic products, or traces of another compound in the natural extract. A similar phenomenon has been seen for other, nonorbitide cyclic peptides, such as phakellistatins 1 and 10 (42), phakellistatin 11 (43), and dolastatin 16 (44). This last case is intriguing, because the study compared the synthetic product with natural samples extracted and tested previously (45). The natural samples were repurified using the same method applied to the synthetic compound and their activity was found to be greatly reduced compared with the earlier results (44, 45). This lends weight to the hypothesis that trace impurities may be responsible for bioactivities found only in natural extracts.

In our case, the compounds in our study and that of Poojary and Belagali (20) were both synthetic, but differences in the method of purification could account for the observed differences in activity. Poojary and Belagali (20) purified evolidine using silica gel chromatography followed by recrystallization from EtOAc–*n*-hexane, whereas we performed two rounds of reversed phase-HPLC (RP-HPLC). In addition, the earlier study used griseofulvin as a positive control against *C. albicans*,

although it has long been known to be ineffective against *Candida* species (46). Although our results showed no biological activity for evolidine or the two other xanthoxycyclins tested, previous work on orbitides in the *Citrus* genus has suggested that such peptides might act as potassium ionophores in plant phloem (5).

Biosynthesis of M. xanthoxyloides orbitides

Several studies have investigated the post-translational processing required to mature plant cyclic peptides. Asparaginyl endopeptidase, a protease that recognizes Asp or Asn at the C terminus of the core peptide for cleavage, has also been recruited by evolution multiple times to perform cyclization of different cyclic peptide families via a transpeptidation reaction. These include (i) kalata-type cyclic peptides (cyclotides), a family of plant macrocyclic peptides containing three disulfide bonds (47–49); (ii) the cyclic knottins, a family of mostly trypsin-inhibitory peptides from the squash family (48, 50); and (iii) the single-disulfide PawS-derived peptides of the daisy family (51) and presumably the closely-related PawL-derived peptides, a group of orbitides in the daisy family with an absolutely conserved Asp at the C terminus of the core peptide (7, 8).

Previous work has revealed that a second cyclization mechanism in plants involving a prolyl oligopeptidase named peptide cyclase 1 (PCY1) is responsible for cyclization of the segetalin orbitides in *Vaccaria hispanica* (*Saponaria vaccaria*) (3, 23). This enzyme recognizes Pro, or sometimes Ala, at the C terminus of the core peptide. It is likely that a prolyl oligopeptidase is similarly responsible for cyclizing the annomuricatin orbitides in *A. muricata*, which all have Pro or Ala in an equivalent position (6).

The xanthoxycyclins of *M. xanthoxyloides* differ from both the above groups because their core peptide sequences terminate with Phe, Leu, Ser, or Lys. None of these four residues are recognized by the aforementioned asparaginyl endopeptidase (Asn/Asp) or prolyl oligopeptidase (Pro/Ala). If xanthoxycyclins are matured by proteases performing a similar cleavage-coupled intra-molecular transpeptidation with the C-terminal residue, our confirmation of these orbitide sequences indicates one or more novel cyclization mechanisms are yet to be discovered in this species, perhaps involving a different class of protease. We would consider the next step to be to attempt to purify this cyclizing enzyme from fractioned plant extracts, by following the desired cyclizing activity.

Here we hope we have redressed priority to evolidine as being the first plant cyclopeptide to be discovered and, by revealing the biosynthetic precursor gene for evolidine and its relatives in the Rutaceae, are able to propose that a novel type of cyclizing enzyme exists in plants awaiting discovery.

Experimental procedures

Plant material

Leaves and stems of *M. xanthoxyloides* were collected on 13 June 2019, by S. Venter (reference 15917) from Smithfield Conservation Park, Queensland, Australia (16.80609° S, 145.67583° E) under permit PTU18-001474. A voucher specimen has been deposited in the Australian Tropical Herbarium. After removing the stems, the leaves were ground to a powder under liquid nitrogen and stored at −70 °C.

Synthetic peptides

Evolidine was synthesized using solid phase peptide synthesis on an automated peptide synthesiser (CS Bio Pty. Ltd.) using fluorenylmethyloxycarbonyl (Fmoc) chemistry on 2-chlorotrityl resin (0.25 mmol scale). Loading of the C-terminal Leu residue (1 molar eq.) was achieved by dissolving the Fmoc-protected Leu residue in dichloromethane, adding 4 molar eq of *N*, *N*′-diisopropylethylamine (DIPEA) to the dissolved residue and adding this to the resin. Subsequent amino acids (4 eq.) were activated in 0.5 M 2-(1*H*-benzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate and DIPEA (8 eq.) in *N*, *N*′-dimethylformamide. The coupling reaction was performed twice for amino acids containing branched β-carbons or aromatic rings. The peptide was cleaved from the resin whilst maintaining attachment of sidechain protecting groups using 2% TFA (v/v) in dichloromethane for 3 min. This was repeated 10 times. Acetonitrile was added to the peptide solution and dichloromethane and TFA were removed via rotary-evaporation and the peptide portion lyophilized. The peptide was then cyclized in dimethylformamide; this was done by adding 1-[bis (dimethylamino)methylene]-1*H*-1,2,3-triazolo[4,5-b]pyridinium 3-oxide hexafluorophosphate (1 eq.) to a 1 mM peptide solution, followed by the slow addition of DIPEA (10 eq.). The reaction was monitored by electrospray ionization MS over 3 h. The peptide was then extracted from the dimethylformamide using phase extraction, prior to lyophilization. The remaining side chain protecting groups were cleaved using TFA/water/triisopropylsilane (97:2:1) for 2 h, the peptide was precipitated with cold diethyl ether, filtered, and dissolved in 50% acetonitrile before lyophilization. The crude peptide was purified with RP-HPLC on a C18 preparative column (300 Å, 10 μm, 21.20 mm internal diameter × 250 mm, Phenomenex). Electrospray ionization MS was used to confirm the mass of evolidine, with further purifications being conducted on a semi-preparative C18 column (300 Å, 5 μm, 10 mm internal diameter × 250 mm, Vydac) with purity being assessed with a C18 analytical column (300 Å, 5 μm, 2.1 mm internal diameter × 150 mm, Vydac). Xanthoxycyclin D and xanthoxycyclin F were purchased from GenScript Inc. (Piscataway, NJ). The purity of all peptides was assessed by RP-HPLC with a C18 analytical column (300 Å, 5 μm, 2.1 mm internal diameter × 150 mm, Vydac) and confirmed to be $\geq$95%.

Peptide extraction and purification

Peptides were extracted from plants as described previously (7). Briefly, about 1 g of leaf powder was ground with 50 ml of 50% methanol, 50% dichloromethane (v/v). The extract was dried with anhydrous magnesium sulfate and filtered under a vacuum with two 50-ml washes of solvent. After flash chromatography though silica gel, the samples were dried in a rotary evaporator (Heidolph). These leaf peptide extracts were then purified as previously described (8). Briefly, the crude extract was purified by solid-phase extraction using a 30-mg Strata-X polymeric reversed-phase column (Phenomenex). The dried

extract was dissolved in 2 ml of an aqueous solution of 5% acetonitrile (v/v), 0.1% formic acid (v/v) and dispensed onto the column, after which purified peptides were eluted with 75% acetonitrile (v/v), 0.1% formic acid (v/v). The purified extract was desiccated in a vacuum centrifuge (Labconco) and then prepared for LC–MS analysis by redissolving it in 50 μl of HPLC-grade 5% acetonitrile (v/v), 0.1% formic acid (v/v) (Honeywell).

LC–MS/MS for peptide sequencing

Samples (2 μl) were separated by a gradient elution from 5% solvent B to 95% solvent B over 15 min on a high-capacity nano-LC chip (Agilent Technologies; part number G4240-62010) driven by a 1200 series nano-flow HPLC system (Agilent) at a flow rate of 300 nl/min. Solvent A was 0.1% (v/v) formic acid in water and solvent B was 0.1% (v/v) formic acid in acetonitrile. Both solvents were HPLC-grade (Honeywell). Peptides were detected with either a 6520 or 6550 Q-TOF mass spectrometer (Agilent) via electrospray ionization in positive mode with one MS scan per second. The source voltage was set between 2100 and 2175 V. MS/MS data were gathered by collision-induced dissociation at a rate of two scans per second, with the expected peptide *m/z* values set as "preferred" for fragmentation. Visual inspection of MS/MS data were used to sequence the cyclic peptides, guided by transcriptomic data for the six novel peptides, and the known sequence for evolidine.

RNA-seq and transcriptome assembly

Total RNA was isolated from frozen ground leaf tissue of *M. xanthoxyloides* using the mini hot phenol method (52), which in turn is derived from the protocol of Botella *et al.* (53). Total RNA was treated with DNase and subsequently purified with the NucleoSpin RNA Clean-Up kit (Macherey-Nagel). It was validated by both agarose gel electrophoresis and absorbance measurement on a NanoDrop2000 (Thermo Fisher Scientific). RNA-seq libraries were generated using the TruSeq Stranded Total RNA with Ribo-Zero Plant kit according to the manufacturer's instructions (Illumina) and sequenced on a NextSeq 550 system (Illumina) as single end reads with a length of 75 bp. The data set contained 9.7×10^9 raw reads and 95% of reads had an average quality score at or above Q30. The leaf transcriptome was assembled using CLC Genomics Workbench 12.0.2 (Qiagen Aarhus A/S). The raw reads were trimmed to a quality threshold of Q30 and minimum length 15, and assemblies were performed with word sizes 23, 30, 40, 50, and 64 and minimum contig length 200. Other parameters remained at their default values.

Analysis of data from the sequence read archive

We searched the SRA for RNA-seq and WGS data for other species in the Rutaceae family. Using CLC Genomics Workbench, RNA-seq data from *A. marmelos* fruit pulp (SRA run SRR8190707), *A. buxifolia* mature fruit (SRR6349678), *B. koenigii* leaf (SRR2970920); *C. excavata* leaf (SRR 6438389), *Z. armatum* young leaf (SRR9179743) and *Z. bungeanum* young leaf (SRR9179746), and WGS data for *A. buxifolia* (SRR6188460) and *C. lansium* (SRR5796634) were downloaded. Data were assembled to a quality threshold Q30 and minimum contig length 100; other parameters were set to their default values, including word size, which the software determined automatically. The assembled transcriptomic and genomic data were subjected to tBLASTn searches using the evolidine and xanthoxycyclin propeptide sequences to find highly similar sequences. The maximum expected value for search results was set to 1 and BLOSUM45 was used as the substitution database.

Extraction of genomic DNA

Genomic DNA (gDNA) was extracted from 2 g of frozen ground powder of *M. xanthoxyloides* leaves with the DNEasy Plant Mini Kit (Qiagen) according to the manufacturer's instructions. Purified DNA was quantified on a NanoDrop 2000 (Thermo Fisher Scientific).

First strand cDNA synthesis

Complementary DNA (cDNA) was generated from previously extracted total RNA (900 ng) using the SMARTER RACE 5′/3′ kit (Takara Bio USA) according to the manufacturer's instructions. A poly-T primer was used. The 20-μl reaction was diluted to 110 μl with water.

PCR and cloning of propeptide genes and transcripts

The DNA template was amplified by the PCR using *Pfu* high-fidelity DNA polymerase (Agilent Technologies). Each 50-μl reaction consisted of gDNA (~27 ng) or cDNA (1 μl) and 1 unit of *Pfu* DNA polymerase solution, with $MgCl_2$ to a final concentration of 2 mM, 400 μM mixed dNTPs, 10 mM KCl, 10 mM $(NH_4)_2SO_4$, 20 mM Tris-HCl (pH 8.8), 0.1 mg/ml of BSA, 0.1% Triton X-100, and forward and reverse primers, 0.4 μM (Table S1).

PCR amplification was carried out in a Veriti 96-well thermocycler (Applied Biosystems) programmed as follows: 95 °C for 2 min followed by 35 cycles of 95 °C for 30 s; 55 °C for 30 s; 72 °C for 30 s; and finally 72 °C for 10 min. In some cases it was necessary to repeat the PCR a second time using 2 μl of the initial PCR product to produce sufficient DNA for subsequent reactions. PCR products were purified using the QIAquick PCR Purification Kit (Qiagen) according to the manufacturer's instructions. DNA was eluted in 30 μl of water after allowing incubation on the column membrane for 1 min, and quantified on a NanoDrop 2000. The *Proxanthoxycyclin D* cDNA was Sanger sequenced directly from the PCR product using the same primers as the PCR (Table S1).

The other products were cloned as follows: overhanging adenosine bases were added to PCR products in an "A-tailing" reaction using *Taq* DNA polymerase. Each 20-μl reaction consisted of ~150 ng of PCR product, 1 unit of *Taq* DNA polymerase, dNTPs to a final concentration of 400 μM, 2 mM $MgCl_2$, 20 mM Tris-HCl (pH 8.8), and 50 mM KCl. Reactions were incubated at 37 °C for 20 min.

Either ~12.5 ng (*Proxanthoxycyclin C* cDNA), ~25 ng (*Proxanthoxycyclin B* cDNA), or ~30 ng (*Proevolidine* and *Proxanthoxycyclin E* gDNA) of the A-tailed PCR product were ligated into 50 ng of the pGEM-T Easy vector (Promega) according to the manufacturer's instructions. Each reaction was incubated

overnight at 4 °C. LB agar plates were prepared containing 100 μg/ml of ampicillin and 100 μM of isopropyl β-D-1-thiogalactopyranoside. Plates were spread with 20 μl of 50 mg/ml 5-bromo-4-chloro-3-indolyl-β-D-galactoside, which was allowed to be absorbed for at least 30 min. These plates were inoculated with a culture of TOP10 *E. coli* cells transformed with 2 μl of the ligated vector and incubated overnight at 37 °C. Aliquots of LB (5 ml) were each inoculated with a single colony from a plate and incubated overnight at 37 °C with shaking.

Plasmid DNA was extracted with the GeneJET Plasmid Miniprep Kit (Thermo Fisher Scientific) according to the manufacturer's instructions. DNA was eluted in 50 μl of water and sent for dideoxy sequencing using the M13F primer of the pGEM-T Easy vector (Australian Genome Research Facility, Perth).

NMR spectroscopy

Xanthoxycyclin samples were prepared by dissolving 2 mg of peptide in 550 μl of H_2O/D_2O (90:10) at pH ~3.5. One-dimensional ^1H data, homonuclear ^1H-^1H two-dimensional TOCSY (37), NOESY (38), and DQF-COSY (double quantum filtered-correlated spectroscopy) experiments were recorded at 298 K on a 900 MHz Bruker Avance III spectrometer equipped with a cryoprobe. TOCSY experiments were recorded with 8 scans and 512 increments and two NOESY experiments were recorded with 48 scans and 512 increments with mixing times of 100 and 200 ms. Sweep widths of 10 ppm were sufficient to cover all proton resonances. Additionally, a ^1H-^{13}C HSQC spectrum was recorded at natural abundance using 128 scans and 256 increments with a sweep width of 10 ppm in the F2 direct dimension and 80 ppm in the F1 dimension. All data were processed using Topspin 4.0.3 (Bruker), with the solvent signal referenced to 4.77 ppm at 298 K. The spectra were analyzed and assigned with the program CARA (computer aided resonance assignment) (54) using sequential assignment strategies (55). Observed Hα, Cα, and Cβ shifts were determined and compared with random coil values (56) for all residues to generate secondary shift graphs. Additionally, amide proton temperature dependence was determined by recording TOCSY data at varying temperatures (288, 293, 298, 303, and 308 K). The chemical shift of each HN proton was plotted against temperature to determine the temperature coefficients.

Solution structure calculations

Integrated peak volumes from NOESY data were used to determine inter-proton distance restraints for xanthoxycyclin D. TALOS-N (39) was used to predict the ϕ (C−1-N-Cα-C) and ψ (N-Cα-C-N+1) backbone dihedral angles. Hydrogen bond donors were identified by backbone amide temperature coefficients. Values ≥4.6 ppb/K for the coefficient are indicative of hydrogen bond donation (57); hydrogen bond acceptors were identified from preliminary structure calculations (57). Preliminary structures (50) were calculated using torsion angle simulated annealing in CYANA (combined assignment and dynamics algorithm for NMR applications) (58). The results of these calculations defined the starting coordinates and distance restraints used for refinement in the program CNS (crystallography and NMR system) (59). NOE distances, dihedral restraints from TALOS-N, and hydrogen bond restraints were all used as input for CNS. Simulated annealing using torsion angle dynamics was performed in CNS to generate 50 structures. All 50 structures were then further minimized in water using Cartesian dynamics to generate the final structures of xanthoxycyclin D. MolProbity (60) was used for stereochemical analysis of the structures via comparison to previously published structures in the PDB. Images of the structure of the xanthoxycyclin D were generated by the program MOLMOL (61). The best 20 structures based on good geometry, no violations above 0.2 Å or 2°, and low energy were chosen to represent the solution structure of xanthoxycyclin D.

Antibacterial and antifungal disc diffusion assays

Synthetic peptides were dissolved in DMSO to a concentration of either 20 mg/ml (evolidine) or 50 mg/ml (xanthoxycyclins D and F). Antibacterial assays were carried out by a previously described method (8). Briefly, sterile LB agar plates were prepared and evenly inoculated with cultures of *E. coli* K-12 or *B. subtilis* Marburg No. 165 using a sterile swab. The cultures had been prepared in liquid LB medium and diluted to an optical density at 600 nm (A_{600}) of 0.1. We also prepared sterile 8-mm filter paper discs containing either a positive control (50 μg of kanamycin), negative control (5 μl of sterile water), or one of the three peptides. Six discs contained 200, 100, 50, 25, 10, or 5 μg of evolidine or 200, 100, 50, 25, 12.5 or 5 μg of xanthoxycyclin D or F. The discs were placed onto the six bacterial plates, incubated overnight at 37 °C, and then inspected for any inhibition bacterial growth by the discs.

A similar but simplified disc diffusion assay was carried out against two species of fungi: *A. fumigatus* and *C. albicans*. A glycerol stock of each species was streaked onto a sterile yeast extract, peptone and dextrose (YPD) agar plate. Cultures were incubated at 37 °C for ~48 h before spores were harvested with 3 ml of sterile 0.1% Tween 80. The spore suspensions were then diluted to an A_{600} of 0.05 and spread onto fresh YPD plates with a sterile swab. Control discs of sterile 8-mm filter paper were prepared with 50 μg of amphotericin B (positive control) or 5 μl of DMSO (negative control). Another three discs contained 200 μg of each peptide. A set of discs was placed onto each of the two fungal plates and incubated for ~24 h at 37 °C before inspection for any inhibition of fungal growth.

Data availability

The *M. xanthoxyloides* RNA-seq raw reads were deposited in the SRA under BioProject accession number PRJNA561449. The seven evolidine- and proxanthoxycyclin-encoding genes or transcripts from this study were deposited in GenBank under accession numbers MN655996 and MT473297. The structure of xanthoxycyclin D was submitted to the PDB under the identifier 6WPV and to the Biological Magnetic Resonance Bank (30747) MS data were deposited to the EMBL-EBI MetaboLights database (62) with the identifier MTBLS1963.

Acknowledgments—We thank Dr. Fanie Venter, James Cook University, for collecting the plant material used in this study. We

thank Dr. Nicolas L. Taylor, the School of Molecular Sciences at the University of Western Australia, for providing advice and assistance in mass spectrometry aspects of this project.

Author contributions—M. F. F., O. B., and K. J. R. data curation; M. F. F., C. D. P., and K. J. R. formal analysis; M. F. F., C. D. P., T. C., O. B., K. J. R., and J. S. M. investigation; M. F. F. and C. D. P. visualization; M. F. F. methodology; M. F. F. and C. D. P. writing-original draft; D. M. C. and J. W. resources; J. W. K. J. R., and J. S. M. supervision; K. J. R. and J. S. M. writing-review and editing; J. S. M. conceptualization; J. S. M. funding acquisition; J. S. M. project administration.

Funding and additional information—This work was supported by Australian Research Council Grants DP190102058 (to J. S. M. and K. J. R.) and CE140100008 (to J. W.). M. F. F. was supported by the Australian Research Training Program and a Bruce and Betty Green Postgraduate Research Scholarship.

Conflict of interest—The authors declare that they have no conflicts of interest with the contents of this article.

Abbreviations—The abbreviations used are: PLP, PawL-derived peptide; SRA, sequence read archive; TOCSY, total correlation spectroscopy; PDB, Protein Data Bank; HSQC, heteronuclear single quantum coherence spectroscopy; RP-HPLC, reversed phase-high performance liquid chromatography; Fmoc, fluorenylmethyloxycarbonyl; DIPEA, N,N′-diisopropylethylamine; gDNA, genomic DNA.

References

1. Arnison, P. G., Bibb, M. J., Bierbaum, G., Bowers, A. A., Bugni, T. S., Bulaj, G., Camarero, J. A., Campopiano, D. J., Challis, G. L., Clardy, J., Cotter, P. D., Craik, D. J., Dawson, M., Dittmann, E., Donadio, S., *et al.* (2013) Ribosomally synthesized and post-translationally modified peptide natural products: overview and recommendations for a universal nomenclature. *Nat. Prod. Rep.* **30,** 108–160 CrossRef Medline
2. Fisher, M. F., Payne, C. D., Rosengren, K. J., and Mylne, J. S. (2019) An orbitide from *Ratibida columnifera* seed containing 16 amino acid residues. *J. Nat. Prod.* **82,** 2152–2158 CrossRef Medline
3. Condie, J. A., Nowak, G., Reed, D. W., Balsevich, J. J., Reaney, M. J. T., Arnison, P. G., and Covello, P. S. (2011) The biosynthesis of Caryophyllaceae-like cyclic peptides in *Saponaria vaccaria* L. from DNA-encoded precursors. *Plant J.* **67,** 682–690 CrossRef Medline
4. Okinyo-Owiti, D. P., Young, L., Burnett, P.-G. G., and Reaney, M. J. T. (2014) New flaxseed orbitides: detection, sequencing, and ^{15}N incorporation. *Peptide Sci.* **102,** 168–175 CrossRef Medline
5. Belknap, W. R., McCue, K. F., Harden, L. A., Vensel, W. H., Bausher, M. G., and Stover, E. (2015) A family of small cyclic amphipathic peptides (SCAmpPs) genes in citrus. *BMC Genomics* **16,** 303 CrossRef Medline
6. Fisher, M. F., Zhang, J., Berkowitz, O., Whelan, J., and Mylne, J. S. (2020) Cyclic peptides in seed of *Annona muricata* are ribosomally synthesized. *J. Nat. Prod.* **83,** 1167–1173 CrossRef Medline
7. Jayasena, A. S., Fisher, M. F., Panero, J. L., Secco, D., Bernath-Levin, K., Berkowitz, O., Taylor, N. L., Schilling, E. E., Whelan, J., and Mylne, J. S. (2017) Stepwise evolution of a buried inhibitor peptide over 45 My. *Mol. Biol. Evol.* **34,** 1505–1516 CrossRef Medline
8. Fisher, M. F., Zhang, J., Taylor, N. L., Howard, M. J., Berkowitz, O., Debowski, A. W., Behsaz, B., Whelan, J., Pevzner, P. A., and Mylne, J. S. (2018) A family of small, cyclic peptides buried in preproalbumin since the Eocene epoch. *Plant Direct* **2,** e00042 CrossRef
9. Behsaz, B., Mohimani, H., Gurevich, A., Prjibelski, A., Fisher, M., Vargas, F., Smarr, L., Dorrestein, P. C., Mylne, J. S., and Pevzner, P. A. (2020) *De novo* peptide sequencing reveals many cyclopeptides in the human gut and other environments. *Cell Systems* **10,** 99–108 CrossRef Medline
10. Hughes, G. K., and Neill, K. G. (1949) Alkaloids of the Australian Rutaceae: *Evodia xanthoxyloides*: I. evoxanthine. *Aust. J. Chem.* **2,** 429–436 CrossRef
11. Brophy, J. J., Goldsack, R. J., and Forster, P. I. (2004) Composition of the leaf oils of the Australian species of *Euodia* and *Melicope* (Rutaceae). *J. Essent. Oil Res.* **16,** 286–293 CrossRef
12. Joyce, E. M., Thiele, K. R., Slik, F. J. W., and Crayn, D. M. (2020) Checklist of the vascular flora of the Sunda-Sahul convergence zone. *Biodivers. Data J.* **8,** e51094 CrossRef Medline
13. Hughes, G. K., Neill, K. G., and Ritchie, E. (1952) Alkaloids of the Australian Rutaceae: *Evodia xanthoxyloides* F. Muell: II. isolation of the alkaloids from the leaves. *Aust. J. Sci. Res.* **5,** 401–405 CrossRef
14. Eastwood, S. F. W., Hughes, G. K., Ritchie, E., and Curtis, R. M. (1955) Alkaloids of the Australian Rutaceae: *Evodia xanthoxyloides* F. Muell: V. a note on the constitution of evolidine. *Aust. J. Chem.* **8,** 552–555 CrossRef
15. Law, H. D., Millar, I. T., and Springall, H. D. (1961) 50. The structure of evolidine. *J. Chem. Soc. (Resumed)*, 279–284 CrossRef
16. Studer, R. O., and Lergier, W. (1965) Struktur und Synthese von Evolidin. *Helv. Chim. Acta* **48,** 460–470 CrossRef
17. Kopple, K. D. (1971) Conformations of cyclic peptides V. A proton magnetic resonance study of evolidine, cyclo-Ser-Phe-Leu-Pro-Val-Asn-Leu. *Biopolymers* **10,** 1139–1152 CrossRef Medline
18. Peishoff, C. E., Bean, J. W., and Kopple, K. D. (1991) Conformation of a cyclic heptapeptide in solution: an NMR constrained distance geometry search procedure. *J. Am. Chem. Soc.* **113,** 4416–4421 CrossRef
19. Eggleston, D. S., Baures, P. W., Peishoff, C. E., and Kopple, K. D. (1991) Conformations of cyclic heptapeptides: crystal structure and computational studies of evolidine. *J. Am. Chem. Soc.* **113,** 4410–4416 CrossRef
20. Poojary, B., and Belagali, S. L. (2005) Synthesis, characterization and biological evaluation of cyclic peptides: viscumamide, yunnanin A and evolidine. *Z. Naturforsch. B* **60,** 1313–1320 CrossRef
21. Tan, N.-H., and Zhou, J. (2006) Plant cyclopeptides. *Chem. Rev.* **106,** 840–895 CrossRef Medline
22. Mehmood, R., and Malik, A. (2010) Isolation and characterization of crotosparsamide, a new cyclic nonapeptide from *Croton sparsiflorus*. *Nat. Prod. Commun.* **5,** 1885–1888 Medline
23. Barber, C. J. S., Pujara, P. T., Reed, D. W., Chiwocha, S., Zhang, H., and Covello, P. S. (2013) The two-step biosynthesis of cyclic peptides from linear precursors in a member of the plant family *Caryophyllaceae* involves cyclization by a serine protease-like enzyme. *J. Biol. Chem.* **288,** 12500–12510 CrossRef Medline
24. Chekan, J. R., Estrada, P., Covello, P. S., and Nair, S. K. (2017) Characterization of the macrocyclase involved in the biosynthesis of RiPP cyclic peptides in plants. *Proc. Natl. Acad. Sci. U.S.A.* **114,** 6551–6556 CrossRef Medline
25. Liu, F. J., Zhu, Z. H., Jiang, Y., and Li, H. J. (2018) A pair of cyclopeptide epimers from the seeds of *Celosia argentea*. *Chin. J. Nat. Med.* **16,** 63–69 CrossRef Medline
26. Kaufmann, H. P., and Tobschirbel, A. (1959) Über ein Oligopeptid aus Leinsamen. *Chemische Berichte* **92,** 2805–2809 CrossRef
27. Prox, A., and Weygand, F. (1966) Sequenzanalyse von Peptiden durch Kombination von Gaschromatographie und Massenspektrometrie. in *Proceedings of the Eighth European Peptide Symposium* (Beyerman, H. C. H. C., van de Linde, A., and Maassen van den Brink, W., eds) North-Holland Publishing Company, Noordwijk, The Netherlands
28. Kodama, Y., Shumway, M., and Leinonen, R, International Nucleotide Sequence Database Collaboration, (2012) The Sequence Read Archive: explosive growth of sequencing data. *Nucleic Acids Res.* **40,** D54–D56 CrossRef Medline
29. Ribeiro, T. A. N., da Silva, L. R., de Sousa Júnior, P. T., Castro, R. N., and De Carvalho, M. G. (2012) A new cyclopeptide and other constituents from the leaves of *Zanthoxylum rigidum* Humb. & Bonpl. ex Willd. (Rutaceae). *Helv. Chim. Acta* **95,** 935–939 CrossRef
30. Yao-Kouassi, P. A., Caron, C., Aké-Assi, E., Martinez, A., Le Magrex-Debar, E., Gangloff, S. C., and Zèches-Hanrot, M. (2014) A new

cycloheptapeptide from *Zanthoxylum mezoneurispinosum* Aké Assi (Rutaceae). *Mediterranean J. Chem.* **3,** 1013–1020 CrossRef
31. Beirigo, P. J., dos, S., Torquato, H. F. V., dos Santos, C. H. C., de Carvalho, M. G., Castro, R. N., Paredes-Gamero, E. J., de Sousa, P. T., Jacinto, M. J., and da Silva, V. C. (2016) [1-8-NαC]-Zanriorb A1, a proapoptotic orbitide from leaves of *Zanthoxylum riedelianum*. *J. Nat. Prod.* **79,** 1454–1458 CrossRef Medline
32. Hu, S.-P., Song, W.-W., Zhao, S.-M., and Tan, N.-H. (2017) Clausenlanins A and B, two leucine-rich cyclic nonapeptides from *Clausena lansium*. *Nat. Prod. Bioprospect.* **7,** 307–313 CrossRef Medline
33. Wélé, A., Zhang, Y., Caux, C., Brouard, J.-P., Pousset, J.-L., and Bodo, B. (2004) Annomuricatin C, a novel cyclohexapeptide from the seeds of *Annona muricata*. *C. R. Chim.* **7,** 981–988 CrossRef
34. Pinto, M. E. F., Batista, J. M., Jr., Koehbach, J., Gaur, P., Sharma, A., Nakabashi, M., Cilli, E. M., Giesel, G. M., Verli, H., Gruber, C. W., Blanch, E. W., Tavares, J. F., da Silva, M. S., Garcia, C. R. S., and Bolzani, V. S. (2015) Ribifolin, an orbitide from *Jatropha ribifolia*, and its potential antimalarial activity. *J. Nat. Prod.* **78,** 374–380 CrossRef Medline
35. Baraguey, C., Blond, A., Correia, I., Pousset, J.-L., Bodo, B., and Auvin-Guette, C. (2000) Mahafacyclin A, a cyclic heptapeptide from *Jatropha mahafalensis* exhibiting β-bulge conformation. *Tetrahedron Lett.* **41,** 325–329 CrossRef
36. Wang, C. K., Kaas, Q., Chiche, L., and Craik, D. J. (2008) CyBase: a database of cyclic protein sequences and structures, with applications in protein discovery and engineering. *Nucleic Acids Res.* **36,** D206–D210 CrossRef Medline
37. Braunschweiler, L., and Ernst, R. R. (1983) Coherence transfer by isotropic mixing: application to proton correlation spectroscopy. *J. Magn. Reson.* **53,** 521–528 CrossRef
38. Jeener, J., Meier, B. H., Bachmann, P., and Ernst, R. R. (1979) Investigation of exchange processes by two-dimensional NMR spectroscopy. *J. Chem. Phys.* **71,** 4546–4553 CrossRef
39. Shen, Y., and Bax, A. (2013) Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56,** 227–241 CrossRef Medline
40. Wagner, G. (1990) NMR investigations of protein structure. *Prog. Nuclear Magn. Reson. Spectroscopy* **22,** 101–139 CrossRef
41. Groppo, M., Pirani, J. R., Salatino, M. L. F., Blanco, S. R., and Kallunki, J. A. (2008) Phylogeny of Rutaceae based on two noncoding regions from cpDNA. *Am. J. Bot.* **95,** 985–1005 CrossRef Medline
42. Napolitano, A., Rodriquez, M., Bruno, I., Marzocco, S., Autore, G., Riccio, R., and Gomez-Paloma, L. (2003) Synthesis, structural aspects and cytotoxicity of the natural cyclopeptides yunnanins A, C, and phakellistatins 1,10. *Tetrahedron* **59,** 10203–10211 CrossRef
43. Pettit, G. R., Lippert, J. W., Taylor, S. R., Tan, R., and Williams, M. D. (2001) Synthesis of phakellistatin 11: a Micronesia (Chuuk) marine sponge cyclooctapeptide. *J. Nat. Prod.* **64,** 883–891 CrossRef Medline
44. Pettit, G. R., Smith, T. H., Arce, P. M., Flahive, E. J., Anderson, C. R., Chapuis, J.-C., Xu, J.-P., Groy, T. L., Belcher, P. E., and Macdonald, C. B. (2015) Antineoplastic agents. 599: total synthesis of dolastatin 16. *J. Nat. Prod.* **78,** 476–485 CrossRef Medline
45. Pettit, G. R., Xu, J-P., Hogan, F., Williams, M. D., Doubek, D. L., Schmidt, J. M., Cerny, R. L., and Boyd, M. R. (1997) Isolation and structure of the human cancer cell growth inhibitory cyclodepsipeptide dolastatin 16. *J. Nat. Prod.* **60,** 752–754 CrossRef Medline
46. Blank, H. (1965) Antifungal and other effects of griseofulvin. *Am. J. Med.* **39,** 831–838 CrossRef Medline
47. Gillon, A. D., Saska, I., Jennings, C. V., Guarino, R. F., Craik, D. J., and Anderson, M. A. (2008) Biosynthesis of circular proteins in plants. *Plant J.* **53,** 505–515 CrossRef Medline
48. Mylne, J. S., Chan, L. Y., Chanson, A. H., Daly, N. L., Schaefer, H., Bailey, T. L., Nguyencong, P., Cascales, L., and Craik, D. J. (2012) Cyclic peptides arising by evolutionary parallelism via asparaginyl-endopeptidase-mediated biosynthesis. *Plant Cell* **24,** 2765–2778 CrossRef Medline
49. Saska, I., Gillon, A. D., Hatsugai, N., Dietzgen, R. G., Hara-Nishimura, I., Anderson, M. A., and Craik, D. J. (2007) An asparaginyl endopeptidase mediates *in vivo* protein backbone cyclization. *J. Biol. Chem.* **282,** 29721–29728 CrossRef Medline
50. Mahatmanto, T., Mylne, J. S., Poth, A. G., Swedberg, J. E., Kaas, Q., Schaefer, H., and Craik, D. J. (2015) The evolution of *Momordica* cyclic peptides. *Mol. Biol. Evol.* **32,** 392–405 CrossRef Medline
51. Bernath-Levin, K., Nelson, C., Elliott, A. G., Jayasena, A. S., Millar, A. H., Craik, D. J., and Mylne, J. S. (2015) Peptide macrocyclization by a bifunctional endoprotease. *Chem. Biol.* **22,** 571–582 CrossRef Medline
52. Box, M. S., Coustham, V., Dean, C., and Mylne, J. S. (2011) Protocol: A simple phenol-based method for 96-well extraction of high quality RNA from *Arabidopsis*. *Plant Methods* **7,** 7 CrossRef Medline
53. Botella, J. R., Arteca, J. M., Schlagnhaufer, C. D., Arteca, R. N., and Phillips, A. T. (1992) Identification and characterization of a full-length cDNA encoding for an auxin-induced 1-aminocyclopropane-1-carboxylate synthase from etiolated mung bean hypocotyl segments and expression of its mRNA in response to indole-3-acetic acid. *Plant Mol. Biol.* **20,** 425–436 CrossRef Medline
54. Keller, R. L. J. (2004) *The Computer Aided Resonance Assignment Tutorial*, 1 Ed., Cantina Verlag, Switzerland
55. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley-Interscience, New York
56. Wishart, D. S., Bigam, C. G., Holm, A., Hodges, R. S., and Sykes, B. D. (1995) ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids: I. investigations of nearest-neighbor effects. *J. Biomol. NMR* **5,** 67–81 CrossRef Medline
57. Cierpicki, T., and Otlewski, J. (2001) Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **21,** 249–261 CrossRef
58. Güntert, P. (2004) Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **278,** 353–378 CrossRef Medline
59. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J. S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54,** 905–921 CrossRef Medline
60. Chen, V. B., Arendall, W. B., III, Headd, J. J., Keedy, D. A., Immormino, R. M., Kapral, G. J., Murray, L. W., Richardson, J. S., and Richardson, D. C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66,** 12–21 CrossRef Medline
61. Koradi, R., Billeter, M., and Wüthrich, K. (1996) MOLMOL: a program for display and analysis of macromolecular structures. *J. Mol. Graph.* **14,** 51–55 CrossRef Medline
62. Haug, K., Cochrane, K., Nainala, V. C., Williams, M., Chang, J., Jayaseelan, K. V., and O'Donovan, C. (2019) MetaboLights: a resource evolving in response to the needs of its scientific community. *Nucleic Acids Res.* **48,** D440–D444 Medline