Letter

New aspercryptins, lipopeptide natural products, revealed by HDAC inhibition in *Aspergillus nidulans*

Matthew T Henke, Alexandra A. Soukup, Anthony Whitney Goering, Ryan A. McClure, Regan J. Thomson, Nancy P. Keller, and Neil L. Kelleher

ACS Chem. Biol., **Just Accepted Manuscript** • DOI: 10.1021/acschembio.6b00398 • Publication Date (Web): 16 Jun 2016

ACS Publications

ACS Chemical Biology is published by the American Chemical Society. 1155 Sixteenth Street N.W., Washington, DC 20036

New aspercryptins, lipopeptide natural products, revealed by HDAC inhibition in *Aspergillus nidulans*

Matthew T. Henke[a], Alexandra A. Soukup[b], Anthony W. Goering[a], Ryan A. McClure[c], Regan J. Thomson[c], Nancy P. Keller[b,d,e] and Neil L. Kelleher[a,c,f*]

[a]Department of Molecular Biosciences, [c]Department of Chemistry and [f]Feinberg School of Medicine, Northwestern University, Evanston, IL 60208, USA
[b]Department of Genetics, [d]Department of Bacteriology, and [e]Department of Medical Microbiology and Immunology University of Wisconsin, Madison, Wisconsin, 53706, USA

ABSTRACT: Unlocking the biochemical stores of fungi is key for developing future pharmaceuticals. Through reduced expression of a critical histone deacetylase in *Aspergillus nidulans*, increases of up to 100-fold were observed in the levels of 15 new aspercryptins, recently described lipopeptides with two non-canonical amino acids derived from octanoic and dodecanoic acids. In addition to two NMR-verified structures, MS/MS networking helped uncover an additional 13 aspercryptins. The aspercryptins break the conventional structural orientation of lipopeptides and appear 'backward' when compared to known compounds of this class. We have also confirmed the 14-gene aspercryptin biosynthetic gene cluster, which encodes two fatty acid synthases and several enzymes to convert saturated octanoic and dodecanoic acid to α-amino acids.

INTRODUCTION

For decades the search for new natural products by screening for bioactivity has been hampered by rediscovering the same compounds.[1] To avoid the rediscovery of natural products, some laboratories have decoupled discovery from screening for bioactivity by measuring accurate mass as the primary, high-throughput screen for the expression of new natural products.[2] This new approach to molecular discovery has been applied to bacteria[3,4] but less so in fungi, which offer an enormous biosynthetic potential.[5] The genome of the mold *Aspergillus nidulans* contains more than 50 gene clusters annotated to be involved in the biosynthesis of natural products. Yet just over 20 of these biosynthetic gene clusters (BGCs) have associated natural products.[6]

Several strategies have been developed to better harness the biosynthetic repertoire of fungi.[7] One such strategy involves the inhibition of histone deacetylase activity (HDACi) and has shown some promise.[8-10] HDACi increases global histone acetylation levels and can increase transcription of otherwise repressed natural product BGCs. Our previous work used quantitative mass spectrometry-based analyses to compare the extracellular metabolomes of *A. nidulans* before and after both chemical and genetic HDACi.[11] We found that HDACi both up-regulated and down-regulated the expression of many compounds. Among the up-regulated compounds were several not known to be produced by *A. nidulans*, such as the fellutamides.[11]

Continuing those efforts here, we have solved the structures of two new lipopeptides, aspercryptin A1 and A2, that are up-regulated by up to 90-fold using the HDACi-based strategy. The aspercryptins are six-amino-acid peptides containing two non-canonical α-amino acids derived from saturated C_8- and C_{12}-fatty acids, and a C-terminal alcohol and related to two previous aspercryptins published by Chiang et al. in 2016.[12] Unlike the overwhelming majority of lipopeptides in the literature, the aspercryptins appear 'backward' and have their lipid tail at the C-terminus. We also map the BGC responsible for the aspercryptins by genetic disruption and Northern blotting. It features a NRPS backbone gene encoding a terminal reductase, atypical activation domains for unusual "fatty amino acid" substrates, and enzymes putatively involved in fatty amino acid biogenesis. The aspercryptins join a small group of lipopeptide natural products where incorporation of the lipid moiety occurs at sites other than the N-terminus.

RESULTS AND DISCUSSION

Comparative metabolomics of HDAC-deficient *A. nidulans* reveals the aspercryptins. In previous work, we generated a mutant of *A. nidulans* with constitutively lower levels of the HDAC RpdA (AN4493).[11] Using differential metabolomics on this strain (*rpdA*^{KD}, RAAS58.4) compared with wildtype fungus (Figure 1, left)), we detected many signals that were up-regulated in the *rpdA*^{KD} mutant (Figure 1, middle panels). Dereplication of these metabolites based on their accurate masses allowed us to cull the list to only those that have not yet been characterized (Figure 1).

Several dozen of the putatively new metabolites displayed a high degree of similarity in their MS/MS fragmentation spectra, suggesting structural and therefore biosynthetic relatedness (Figure 2). For the members of this compound family, all were up-regulated over a wide range from 2- to 130-fold in *rpdA*^{KD} vs. wildtype (data for one compound are shown in Figure 1, third panel). The structures of two of these compounds, aspercryptin A1 (*m/z* 758.5386 $[M+H]^+$, $C_{37}H_{71}N_7O_9$, -0.02 ppm) and aspercryptin A2 (*m/z* 742.5440 $[M+H]^+$, $C_{37}H_{71}N_7O_8$, 0.45 ppm), (Figure 1, right panel) were determined completely by MS with stable isotope feeding and extensive NMR (Supporting Information Table S1).

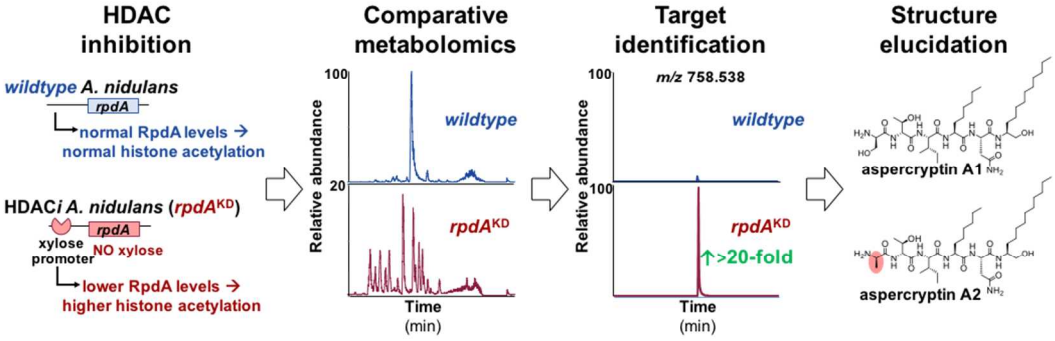

Figure 1. Workflow to compare the metabolomes of wildtype *Aspergillus nidulans* and a strain where the HDAC RpdA is constitutively repressed (*rpdA*^{KD}) leading to hyperacetylation in bulk chromatin. Knockdown of RpdA leads to many newly observed metabolites (second panel), from which new metabolites can be viewed selectively (third panel) and targeted for structure determination when they display unique mass signatures. Aspercryptins A1 and A2 are lipopeptides made by *A. nidulans* (fourth panel at far right). Full NMR data can be found in Supporting Information. Stable isotope incorporation experiments verify much of these structures (Supporting Information Figure S2). Annotated MS/MS spectra for both aspercryptin A1 and A2 (Supporting Information Figures S3 and S4) are included for reference.

The aspercryptins are 'backward' lipopeptides. The aspercryptins are linear lipopeptides built from six amino acids and appear to be the first example of peptide natural products with two lipid groups. They also seem to be the only known lipopeptides with a lipid tail at the C-terminus; thus they appear 'backward' to other lipopeptides. This is despite the overwhelming literature precedent for lipopeptides having an N-terminal lipid group.

Aspercryptin A1 and A2 differ from each other only at the N-terminal residue - serine for aspercryptin A1 and alanine for aspercryptin A2. The most striking feature of the aspercryptins is that of the six amino acids, two are highly unusual and non-proteogenic: 2-amino-octanoic acid and 2-amino-dodecanol. Stereochemical assignment of the aspercryptins was determined for each residue by derivatization with Marfey's reagent and comparison with standards using HPLC-MS (Supporting Information Figure S1). For aspercryptin A1 the first two residues (serine and threonine) are epimerized to D-serine and D-*allo*-threonine while the remaining 4 residues are the L-isomers. The stereocenters of aspercryptin A2 are the same as those in aspercryptin A1; however, there is also an epimer of aspercryptin A2 such that the alanine is the L-isomer, thus we name this analogue *epi*-aspercryptin A2 (Supporting Information Figure S1). Several of these monomers (threonine, isoleucine and serine) were also confirmed by metabolic feeding of stable isotope analogues (Supporting Information Figure S2). Furthermore, stable-isotope labeling shows that when fed d_3-serine, labeled aspercryptin A1 only shows incorporation of 2 deuterons (Supporting Information Figure S2d). This further supports the epimerization of this residue by loss of the deuteron on the alpha carbon.

MS/MS networking reveals a large family of aspercryptins. We observed that the MS/MS spectra of many other metabolites up-regulated by *rpdA*^{KD} had striking similarity to those of aspercryptin A1 and A2. We then turned to MS/MS networking to visualize the relatedness of these metabolites in the dataset.

The ability for MS/MS networking to cluster biosynthetically related compounds is now established.[13] When performing network analysis of the wildtype and *rpdA*^{KD} extracts we saw several examples of biosynthetically related natural products clustering together as expected; for example, the emericellamides group tightly together based on MS/MS spectral similarity (Figure 2a, circled).

One cluster of metabolites consistently upregulated in the *rpdA*^{KD} extracts (Figure 2a, boxed) contained aspercryptins A1 and A2 (Figure 2b). Using the MS/MS fragmentation patterns of aspercryptin A1 and A2 (Supporting Information Figures S3 and S4) and the recently published aspercryptins B1 and B3[12] as anchor points, we detected and have proposed putative structures for an additional 13 aspercryptins (Figure 2b, yellow circles, Table 1). These new aspercryptins fall into sub-families and we were readily able to quantify their abundance increases in the *rpdA*^{KD} mutant (Supporting Information Figure S5).

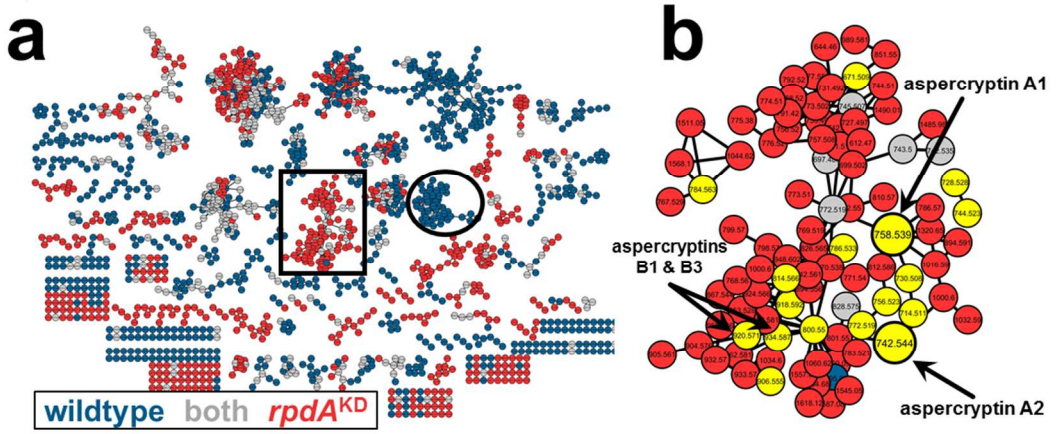

Figure 2. MS/MS networking to identify and characterize a large cluster of aspercryptins. (a) MS/MS networking clusters natural products into molecular families. Here the metabolites made by wildtype *A. nidulans* (blue) are compared to those made by the *rpdA*^{KD} mutant (red). Some metabolites are expressed equally (grey). Some molecular families are seen in only one biological state. The aspercryptins are boxed and the emericellamides are circled. (b) Aspercryptin A1 and A2 (larger yellow circles) fall into an MS/MS cluster of metabolites mostly expressed by *rpdA*^{KD} mutant. Comparing the MS/MS spectra of 5 NMR-elucidated aspercryptins allows for characterization of an additional 13 aspercryptins in this molecular family for a total of 18 structurally characterized aspercryptins, which includes aspercryptins B1 and B3 (yellow).

Table 1. The formulae and expression ratios for the 18 aspercryptins from Figure 2, including the five aspercryptins elucidated by NMR (**bold**) and the 13 additional aspercryptins whose putative structures are supported by comparing MS/MS spectra. [a]NMR structures described in this work; [b]NMR structures described by Chiang et al;[12] [c]ratio is an unresolved combination of epimers

Name	*m/z* [M+H]+	molecular formula	[*rpdA*^{KD}] [wildtype]	Name	*m/z* [M+H]+	molecular formula	[*rpdA*^{KD}] [wildtype]
aspercryptin A1[a]	758.539	$C_{37}H_{71}N_7O_9$	20	**aspercryptin B1**[b]	934.586	$C_{47}H_{79}N_7O_{12}$	60
aspercryptin A2[a]	742.544	$C_{37}H_{71}N_7O_8$	90[c]	aspercryptin B2	918.591	$C_{47}H_{79}N_7O_{11}$	130
***epi*-aspercryptin A2**[a]	742.544	$C_{37}H_{71}N_7O_8$	90[c]	**aspercryptin B3**[b]	920.571	$C_{46}H_{77}N_7O_{12}$	50
aspercryptin A3	744.523	$C_{36}H_{69}N_7O_9$	2	aspercryptin B4	906.555	$C_{45}H_{75}N_7O_{12}$	120
aspercryptin A4	730.508	$C_{35}H_{67}N_7O_9$	2	aspercryptin C1	800.549	$C_{39}H_{73}N_7O_{10}$	20
aspercryptin A5	728.528	$C_{36}H_{69}N_7O_8$	20	aspercryptin C2	784.554	$C_{39}H_{73}N_7O_9$	60
aspercryptin A6	714.513	$C_{35}H_{67}N_7O_8$	20	aspercryptin C3	786.534	$C_{38}H_{71}N_7O_{10}$	3
aspercryptin A7	671.507	$C_{34}H_{66}N_6O_7$	20	aspercryptin C4	772.518	$C_{37}H_{69}N_7O_{10}$	2
aspercryptin D1	814.566	$C_{40}H_{75}N_7O_{10}$	5	aspercryptin C6	756.523	$C_{37}H_{69}N_7O_9$	20

The putative structures for 8 of the additional 13 are shown in Figure 3. A subset was chosen for clarity in the main text (for a complete version of this chart see Supporting Information Figure S6). The variations in the structures of the aspercryptins are standard for what is seen from NRPS pathways: incorporation of alanine instead of serine, valine for isoleucine and a C_{10} fatty amino alcohol instead of the C_{12} version. Taking the structure of aspercryptin A1 as the 'canonical' sequence of the aspercryptins,

we have made a hierarchical nomenclature system, where the letter represents the status of the N-terminus: A for a free amine, B for cichorine capped, C for acetyl and D for propionyl; and where the number represents 'variants' of the base amino acid sequence: 1 for canonical, 2 for serine to alanine, etc. Isolation and full NMR structure determination for each of the aspercryptins are beyond the scope of this work. While the *de novo* structure elucidation of natural products from analysis of MS and MS/MS data alone is not possible, such analyses can readily be used for the detection and putative characterization of highly similar structural analogues. Though the stereochemistry of the aspercryptins proposed by MS/MS analysis is likely the same as those that have been experimentally determined, we have refrained from assuming that this is indeed the case (Figure 3).

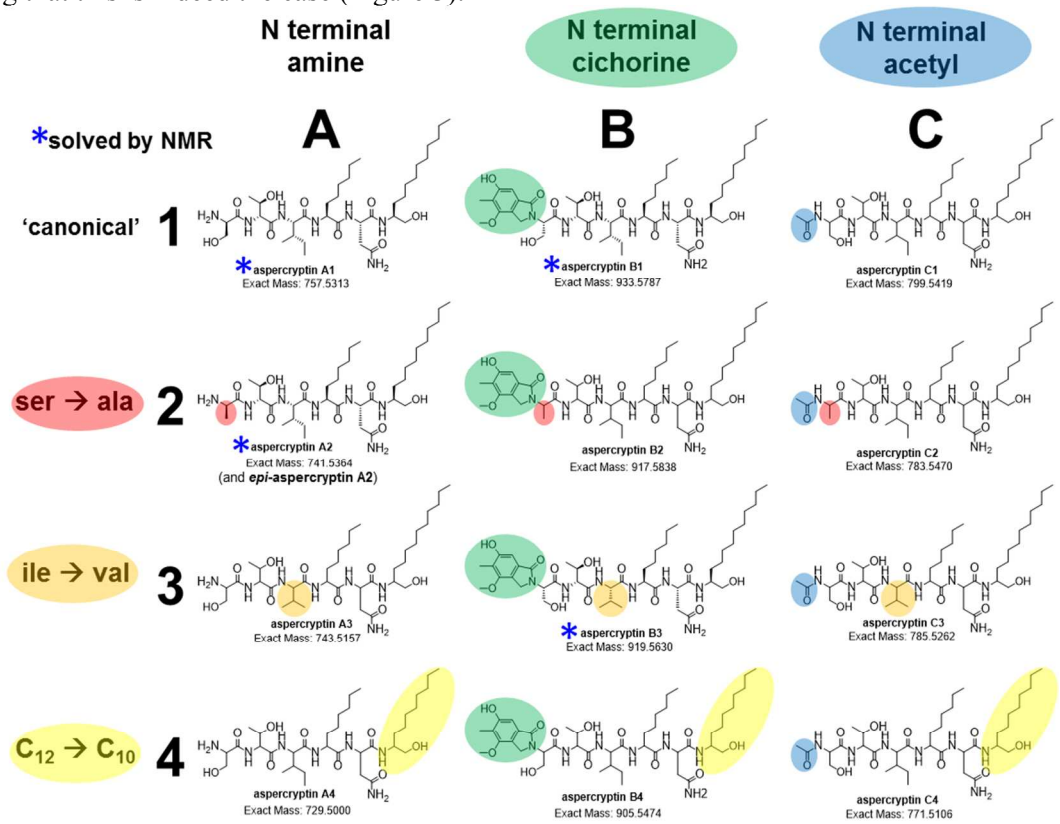

Figure 3. Putative structures for 8 of the 13 aspercryptins based on analysis of MS spectral differences. A complete version of this chart with the 18 structurally characterized aspercryptins is shown in the Supporting Information. Heavily annotated MS/MS spectra of aspercryptins solved by NMR (blue asterisk) were used as the basis for comparison to other metabolites (Supporting Information Figures S3 and S4). Conservatively, stereochemical assignment has not been made for the MS-based structures.

Confirmation of the aspercryptin biosynthetic gene cluster. Linking natural products to their gene clusters is key to annotating the biosynthetic potential of phylogenetically diverse fungi. The genome of *A. nidulans* contains a 14-gene BGC[12] (AN7884 to AN7872, Figure 4a and Supporting Information Table S3) that was recently shown to produce aspercryptins B1 and B3.[12] This BGC contains a six-module NRPS (AN7884) and two fatty acid synthase (FAS) subunits (AN7880 and AN7873). To experimentally link all of the aspercryptins to this BGC, we deleted the NRPS gene, AN7884 and observed no detectable levels of any of the aspercryptins (Figure 4b). We also generated an overexpression mutant of the nearby transcription factor, AN7872, which led to increased transcript levels for many of the genes predicted to be in the cluster. This mirrored the overexpression of these genes in the *rpdA*^{KD} strain (Supporting Information Figure S7) and confirmed the predicted cluster boundaries. Further solidifying its role as the transcription factor for the cluster, we also observed a >2-fold increase in the levels of aspercryptin A1 upon overexpression of AN7872 (Figure 4c). Additionally deletion of the FAS gene AN7880 abolished production of the aspercryptins. Levels of the aspercryptins could be rescued partially by supplementing the media with octanoic and dodecanoic acids, which are likely the direct products of the FAS enzymes (Figure 4c).

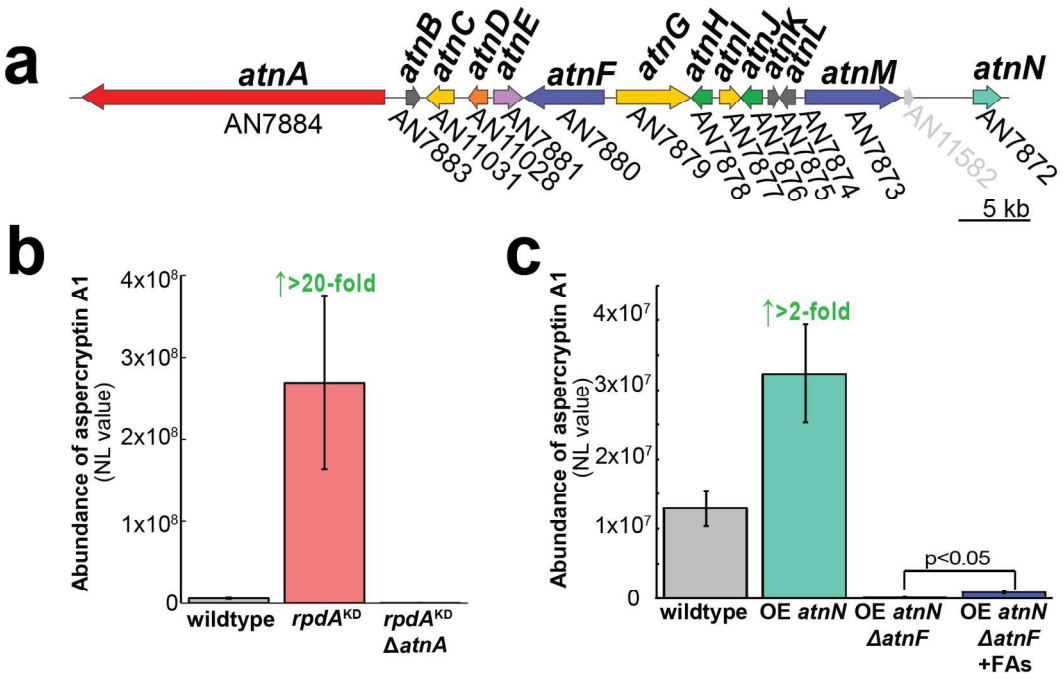

Figure 4. Evidence for the association of the aspercryptins to their BGC. (a) The BGC responsible for the biosynthesis of the aspercryptins contains a **non-ribosomal peptide synthetase** encoded by *atnA* (AN7884), which contains 6 adenylation domains and a terminal reductase domain; the two **fatty acid synthase** subunits *atnF* (AN7880) and *atnM* (AN7873); a **transcription factor** *atnN* (AN7872); two **aminotransferases** *atnH* (AN7878) and *atnJ* (AN7876); a **cytochrome P450** *atnE* (AN7881); an **oxidoreductase** *atnD* (AN11028); and 3 for **transport** *atnC* (AN11031) and *atnF* (AN7879) and **resistance** *atnI* (AN7877). (b) Deletion of the NRPS gene *atnA* in the background of *rpdA*^{KD} showed it is necessary for biosynthesis of the aspercryptins (only aspercryptin A1 is shown here for clarity). (c) Overexpression of the transcription factor *atnN* (OE *atnN*) led to a doubling of aspercryptin A1 levels. Subsequent deletion of the FAS gene *atnF* (OE *atnN*, Δ*atnF*) abolished levels of aspercryptins, which could rescued by supplementing media with the fatty acids octanoic and dodecanoic acids (OE *atnN*, Δ*atnF* + FAs).

The NRPS, the transcription factor, and the intervening genes have been named *atnA* (NRPS) through *atnN* (transcription factor). Upon exploring the phylogenetic distribution of the *atn* BGC in the *Aspergillus* genome repository (http://www.aspergillusgenome.org), we found the *atn* BGC present in six other Aspergilli and to be most conserved in *A. versicolor* (NRPS genes 87% identical, Supporting Information Figure S8). The BGC borders were confirmed by transcript mapping using Northern blots of all genes in the cluster (Supporting Information Figure S7).

Proposed biosynthesis for the aspercryptins. Based on annotations of the BGC and the above deletion experiments, we propose the following for aspercryptin biosynthesis (Supporting Information Figure S9). The first step is the generation of the fatty acid precursors, octanoic and dodecanoic acids, by the FAS subunits AtnF and AtnM (AN7880 and AN7873). *A. nidulans* has 3 other pairs of FAS genes. One, *fasA* and *fasB*, is required for fatty acids involved in primary metabolism; deletion of either *fasA* or *fasB* is lethal.[14] The other 2 pairs of FAS genes make precursor fatty acids for natural products – *pkiB* and *pkiC* for the polyketides made by the PKS *pkiA* (AN3386)[15] and *stcJ* and *stcK* for sterigmatocystin biosynthesis.[14]

The fatty acid precursors are perhaps the most interesting aspect of the system and are likely transformed into the corresponding α-amino fatty acids in three steps. First they are hydroxylated by the cytochrome P450 AtnE (AN7881), then oxidized to the corresponding α-keto acids by the NAD(P)-dependent oxidoreductase AtnD (AN11028), and finally converted to the α-amino fatty acids by the PLP-dependent aminotransferases AtnH or AtnJ (AN7878 and AN7876). Similar pathways to convert a fatty acid or polyketide precursor to the corresponding α-amino acid have been proposed for the cyclosporins and apicidins/HC-toxin, and experimentally supported to varying degrees.[16-18] The only experimental evidence to support this transformation pathway from these systems comes from the deletion of a branched chain aminotransferase which eliminates HC-toxin biosynthesis.[19] A recent publication by Chiang et al. also proposes this pathway for production of the fatty amino acids, and showed that deletion

of these genes abolished levels of aspercryptin B1 (except in the case of the aminotransferases, AtnH and AtnJ, which could reasonably compensate for each other).[12]

Unlike the other FAS genes in the *A. nidulans* genome, the aspercryptin FAS elements are 'interrupted' by the aminotransferase genes that we suggest are involved in generation of the α-amino fatty acids (Supporting Information Figure S10). Perhaps this chromosomal organization evolved to decrease the likelihood that the gene cassette necessary for the fatty amino acid biosynthesis will be separated by chromosomal reorganization. If they were not disrupted by the aminotransferases, the expression of the FAS genes could conceivably result in unproductive synthesis of free medium-chain fatty acids that could disrupt membrane stability. In fact, addition of dodecanoic acid above 10 μM for rescue experiments resulted in little to no growth.

Once made, we propose that the α-amino fatty acids, 2-amino-octanoic and 2-amino-dodecanoic acids, are recognized, activated and covalently tethered to the NRPS AtnA by its fourth and sixth adenylation domains (Supporting Information Figure S9). For typical lipopeptides, lipid moieties are added to the N-terminus by an initial terminal condensation domain of the NRPS.[20] Incorporation of the lipid group as an α-amino fatty acid by an adenylation domain has been proposed for members of the apicidin/HC-toxin family and demonstrated in the biosynthesis of cyclosporin, where the NRPS has been shown to adenylate and covalently bind to the polyketide-derived α-amino acid (4R)-4-[(E)-2-butenyl]-4-methyl-L-threonine (bmt).[21,22] Experiments are on-going to directly establish the activity of adenylation domains 4 and 6 of AtnA in recognizing and activating 2-amino-octanoic and 2-amino-dodecanoic acids, respectively. To date, only a handful of α-amino fatty acids have been observed in cyclic lipopeptides (Supporting Information Figure S11). Intriguingly all of these are derived from fungi, with the exception of the piperazimycins.[23] Perhaps this strategy for lipid incorporation into natural products is a fungal innovation that spread into Streptomycetes.

In general, using the NRPSpredictor2 (modelled from bacterial NRPS) did not confidently predict the amino acid substrates recognized by each of the six adenylation domains of AtnA (Supporting Information Table S4). While Chiang et al. propose cichorine-serine as the monomer activated by the first adenylation domain of AtnA for the biosynthesis of aspercryptin B1,[12] we believe it to be serine, and that the N-terminal amine of mature aspercryptin A1 subsequently reacts during the biosynthesis of cichorine[24] to form aspercryptin B1. Additionally, despite AtnA having only one epimerase domain in the threonine module, the first two amino acids of aspercryptin A1 are D-serine and D-*allo*-threonine. This suggests that serine is either loaded directly as D-serine, or that the epimerase domain in the threonine module epimerizes both L-serine and L-threonine. Iterative epimerase domains are not without precedent.[25] Because we observed that the alanine residue of aspercryptin A2 exists as a mixture of L- and D- enantiomers, we predict that the epimerase domain acts iteratively. Chronologically, L-serine (for aspercryptin A1) or L-alanine (for aspercryptin A2) is activated by the first adenylation domain of AtnA, and while serine is fully epimerized to the D-enantiomer by the epimerase domain, alanine is only partially converted to the D-enantiomer. The complete conversion of serine and threonine and partial conversion of alanine suggests that the sidechain hydroxyl of serine and threonine may aid in governing substrate recognition by the epimerase domain.

The final step in the biosynthesis of the aspercryptins is the reduction of the C-terminus to an alcohol. Terminal reductase domains generally use the energy from NAD(P)H to install a reactive aldehyde that serves as a warhead[26] or as an intermediate that rearranges to yield a mature natural product.[25] For the aspercryptins, the removal of a charge-bearing site from dominantly hydrophobic portion of the compound is a potential reason for the installation of this alcohol at the C-termini of aspercryptins. In fact, the putatively membrane-associated nature of the aspercryptins may be relevant to their function and is consistent with our ability to isolate these compounds mainly from the cell mass and little from the extracellular medium. Many microbes use lipopeptides as a means to adhere to or move across surfaces, and to establish biofilms.[27] Because we could not assign an anti-bacterial activity, the aspercryptins may serve a similar structural/motility function for *A. nidulans*. We also note that *atnH* and *atnJ* are induced by ethanol (the other *atn* genes were not examined), which possibly reflects a role for this metabolite during hypoxic stress.[28]

CONCLUSION

Here we uncovered over a dozen family members of the recently discovered lipopeptides, the aspercryptins, and mapped the complete gene cluster responsible for their biosynthesis. Instrumental in the discovery of the aspercryptins was the marriage of HDAC inhibition with MS-based metabolite screening. By inhibiting HDAC function, we were able to tease the production of a new family of natural

products. By taking advantage of MS/MS networking, we could characterize the structural variation of the aspercryptins. Our successful use of HDACi in the discovery of the aspercryptins demonstrates its potential to survey fungal extracts for new chemical matter, even in species that have been so intensively studied, such as *A. nidulans*. Discovery of new natural product scaffolds from the microbial world at rates far higher than in past decades now represents a promising path for reinvigorating the pipeline of compounds flowing into pharmaceutical screening platforms.

METHODS

Fungal Transformations. The generation of the *rpdA*^{KD} strain was previously described.[11] All other strains in this study are described in Table S5. For construction of overexpression strains, 1 kb of *AN7884* or *AN7872* 5' flanking regions and 3' coding regions were amplified and fused to *A. parasiticus pyrG* (amplified from pJW24) and a 401 bp fragment of the *alcA* promoter using double joint PCR.[29,30] The resulting overexpression construct was transformed into RJMP1.1 to create strains TAAS393.2 and TAAS394.2. For deletion constructs, 1 kb of flanking regions were amplified and fused to *A. fumigatus riboB* using double joint PCR. The resulting knockout construct(s) were transformed into RJMP1.1 and/or TAAS393.2 to create strains TAAS176.3 TAAS395.1, and TAAS217.1. Transformants were examined for targeted replacement of the native loci by PCR and Southern blotting, and expression levels were confirmed by northern analysis.

Growth and extraction of fungal strains for LC-MS/MS analysis. All strains were grown with initial inoculations of 10^6 spores/mL in 250 mL GMM in 1L unbaffled flasks, grown in the dark at 37˚C for 4 days at 200 rpm. For the initial screening of the *rpdA*^{KD}, overexpression and deletant mutants, extractions and analysis were performed as previously described.[11] For induction of overexpression strains, lactose minimal medium + 30 mM cyclopentanone was used. For rescue experiments, media was supplemented with 10 μM octanoic and doceanoic acids in acetone. However, for stable isotope incorporation experiments, 5 mL of GMM in 13 mL culture tubes were inoculated with 10^6 spores/mL and grown as described above, with the additional step of spiking with 1 mM sterile-filtered stable-isotope amino acids after 48 hours growth. After a total of 4 days of growth, whole cultures were extracted with an equal volume of ethyl acetate with sonication and vortexing for several minutes, followed by overnight incubation at 4˚C. Organic layers were then dried down and analyzed as previously described.[11]

Isolation and structural determination of aspercryptins A1 and A2. For isolation of aspercryptin from *rpdA*^{KD} mutant, large-scale growths (500 mL growths in 2L flask, total of 4 L) were performed as described above. Mycelium was separated from the spent media with coffee filters. The spent media was extracted with dichloromethane to generate an emulsion layer. The emulsion was dried down. The cellular material was washed with excess methanol. Methanol extracts and emulsions from DCM extraction were pooled in methanol and dried onto excess silica. Silica was washed with excess ethyl acetate and aspercryptins were eluted from silica with methanol. Methanol fraction was pHed to ~8.5 and run over SAX resin (Dowex® 1X2 chloride form), flowthrough contained aspercryptins. SAX Resin washed with MeOH (pH 5) and pooled with flowthrough. This was then fractionated over preparative RPLC to afford ~4 mg of aspercryptin A1 and A2 in a ~2:1 mixture that could not be chromatographically resolved.

All NMR experiments were performed in DMSO-d_6 on an Agilent 600 MHz DD2 with HCN cryoprobe, except for ^{13}C-NMR, which was acquired with AVANCE III 500 MHz with direct cryoprobe.

Stereochemistry of amino acids was determined following a standard Marfey's Test protocol.[31] Briefly, 1 mg of a ~2:1 mixture of aspercryptins A1 and A2 was hydrolyzed in 6N HCl overnight at 110ºC. The hydrolyzed mixture was dried *in vacuo*, to which 100 μL 1% Marfey's reagent (FDAA) in acetone and 20 μL 1M $NaHCO_3$ were added. Amino acid standards were derivatized as described above; 2.5 μmoles were used of each DL-alanine, D-serine, L-serine, DL-threonine, DL-*allo*-threonine, DL-isoleucine, DL-aspartic acid (for asparagine), synthetic DL-2-amino-octanoic acid and synthetic DL-2-amino-dodecanol.

MS/MS networking to identify aspercryptin molecular family. All MS/MS spectra were pre-processed by removing the 25% lowest intensity peaks, applying a non-linear transformation to the peak intensities by taking their square root, and normalizing peak intensities to the sum of the intensities of all remaining peaks in the spectrum. Spectra with less than 10 remaining peaks, along with those where the base peak constituted more than 75% of the total scan intensity were removed from the analysis. MS/MS spectra were compared using the cosine similarity method. When comparing two spectra, if more than six matching peaks were found, the remaining unmatched peaks were aligned by shifting their *m/z* by the difference in precursor mass. The resulting output is a cosine score between 0 and 1 that describes the similarity between two spectra, where 1 represents a perfect match. MS/MS spectra from the same precursor were determined by a cosine similarity of >0.7 and a precursor match within 0.01 *m/z*. In cases where multiple spectra from the same precursor were observed, the spectrum with the higher intensity was used and the lower intensity spectrum was discarded.

Using the cosine comparisons, a network containing 1590 nodes and 3396 edges was constructed and visualized in Cytoscape, where each node is a representative spectrum from a unique precursor, and each edge is a cosine comparison with a value of 0.4 or greater.

Synthesis of lipid amino acid monomers. DL-2-amino-octanoic acid. ±-2-bromo-octanoic acid (1 g, 4.5 mmol) was dissolved in 10 mL 1:1 H_2O:acetone. NaN_3 (0.5 g, 7.7 mmol) was added. The solution was vigorously stirred overnight at room temperature (reaction mostly complete after 1 hour) to yield ±-2-azido-octanoic acid. The 2-azido-octanoic acid was then hydrogenated in dry THF with 20% Pd/C under 1 atm H_2 with constant stirring at room temperature overnight. Mixture was filtered through celite to remove Pd/C.[32] A small aliquot was purified by RP-LC to yield 36 µg of pure DL-2-amino-octanoic acid (*m/z* 160.1333 $[M+H]^+$, $C_8H_{17}NOH^+$, 0.6 ppm). White waxy solid.^{1}H- and ^{13}C-NMR spectra can be found in Supporting Information. **DL-2-amino-dodecanol.** ±-2-bromo-dodecanoic acid (1.25 g, 4.5 mmol) was dissolved in 10 mL 1:1 H_2O:acetone. NaN_3 (0.5 g, 7.7 mmol) was added. The solution was vigorously stirred overnight at room temperature (reaction mostly complete after 1 hour) to yield ±-2-azido-dodecanoic acid. The 2-azido-dodecanoic acid was then reduced in dry THF with $LiAlH_4$ (0.35 g, 9.34 mmol) with constant stirring on an ice bath for 2 hours.[33] The reaction was quenched with 5% KHSO4 and extracted with ethyl acetate. A small aliquot was purified by RP-LC to yield 70 µg of DL-2-amino-dodecanol (*m/z* 202.2166 $[M+H]^+$, $C_{12}H_{27}NOH^+$, 0.2 ppm). Off-white waxy solid. ^{1}H- and ^{13}C-NMR spectra can be found in Supporting Information.

Bioactivity assays. Fifteen µg of purified aspercryptins A1 and A2 (~2:1) were tested in a standard disk diffusion assay against *M. luteus*, *E. coli*, *P. aeruginosa*, *S. epidermidis*, *K. pneumoniae*, *B. subtilis*, *A. nidulans* and *P. citrinum*.

ASSOCIATED CONTENT

Supporting Information. Figures are largely of structural validation of the aspercryptins, including MS/MS spectra. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: n-kelleher@northwestern.edu

Author Contributions

MTH grew strains, analyzed extracts by LC-MS, determined the structures of aspercryptins A1 and A2 by NMR, the other 13 aspercryptins by MS/MS and prepared the manuscript. AAS generated all mutants used in this study. AWG performed MS/MS networking. RAM performed reduction reactions and Marfey's derivatization. NPK annotated the distribution of the *atn* gene cluster. AAS, NLK and NPK heavily revised several versions of the manuscript.

Notes

ACKNOWLEDGEMENTS

We would like to thank the National Institutes of Health R01 AT009143 (N. Kelleher), T32 GM07133 and NRSA AI55397 (A. Soukup), and NIH 5PO1GM084077 (N. Keller). We also thank Y. Zhang of IMSERC at Northwestern University for NMR assistance (supported by the NIH 1S10OD012016-01 and 1S10RR019071-01A1).

REFERENCES

(1) Baltz, R. H. (2006) Marcel Faber Roundtable: is our antibiotic pipeline unproductive because of starvation, constipation or lack of inspiration? *J. Ind. Microbiol. Biotechnol*. *33*, 507-513.

(2) Hoffmann, T.; Krug, D.; Huttel, S.; Muller, R. (2014) Improving natural products identification through targeted LC-MS/MS in an untargeted secondary metabolomics workflow *Anal. Chem*. *86*, 10780-10788.

(3) Sidebottom, A. M.; Johnson, A. R.; Karty, J. A.; Trader, D. J.; Carlson, E. E. (2013) Integrated metabolomics approach facilitates discovery of an unpredicted natural product suite from *Streptomyces coelicolor* M145 *ACS Chem. Biol. 8*, 2009-2016.

(4) Derewacz, D. K.; Covington, B. C.; McLean, J. A.; Bachmann, B. O. (2015) Mapping Microbial Response Metabolomes for Induced Natural Product Discovery *ACS Chem. Biol. 10*, 1998-2006.

(5) Khaldi, N.; Seifuddin, F. T.; Turner, G.; Haft, D.; Nierman, W. C.; Wolfe, K. H.; Fedorova, N. D. (2010) SMURF: Genomic mapping of fungal secondary metabolite clusters *Fungal Genet. Biol. 47*, 736-741.

(6) Yaegashi, J.; Oakley, B. R.; Wang, C. C. (2014) Recent advances in genome mining of secondary metabolite biosynthetic gene clusters and the development of heterologous expression systems in *Aspergillus nidulans J. Ind. Microbiol. Biotechnol*. *41*, 433-442.

(7) Brakhage, A. A.; Schroeckh, V. (2011) Fungal secondary metabolites - strategies to activate silent gene clusters *Fungal Genet. Biol. 48*, 15-22.

(8) Fisch, K. M.; Gillaspy, A. F.; Gipson, M.; Henrikson, J. C.; Hoover, A. R.; Jackson, L.; Najar, F. Z.; Wagele, H.; Cichewicz, R. H. (2009) Chemical induction of silent biosynthetic pathway transcription in *Aspergillus niger J. Ind. Microbiol. Biotechnol*. *36*, 1199-1213.

(9) Henrikson, J. C.; Hoover, A. R.; Joyner, P. M.; Cichewicz, R. H. (2009) A chemical epigenetics approach for engineering the in situ biosynthesis of a cryptic natural product from *Aspergillus niger Org. Biomol. Chem.7*, 435-438.

(10) Mao, X. M.; Xu, W.; Li, D.; Yin, W. B.; Chooi, Y. H.; Li, Y. Q.; Tang, Y.; Hu, Y. (2015) Epigenetic genome mining of an endophytic fungus leads to the pleiotropic biosynthesis of natural products *Angew. Chem. 54*, 7592-7596.
(11) Albright, J. C.; Henke, M. T.; Soukup, A. A.; McClure, R. A.; Thomson, R. J.; Keller, N. P.; Kelleher, N. L. (2015) Large-scale metabolomics reveals a complex response of *Aspergillus nidulans* to epigenetic perturbation *ACS Chem. Biol. 10*, 1535-1541.
(12) Chiang, Y. M.; Ahuja, M.; Oakley, C. E.; Entwistle, R.; Asokan, A.; Zutz, C.; Wang, C. C.; Oakley, B. R. (2016) Development of Genetic Dereplication Strains in *Aspergillus nidulans* Results in the Discovery of Aspercryptin *Angew. Chem. 55*, 1662-1665.
(13) Nguyen, D. D.; Wu, C. H.; Moree, W. J.; Lamsa, A.; Medema, M. H.; Zhao, X.; Gavilan, R. G.; Aparicio, M.; Atencio, L.; Jackson, C.; Ballesteros, J.; Sanchez, J.; Watrous, J. D.; Phelan, V. V.; van de Wiel, C.; Kersten, R. D.; Mehnaz, S.; De Mot, R.; Shank, E. A.; Charusanti, P.; Nagarajan, H.; Duggan, B. M.; Moore, B. S.; Bandeira, N.; Palsson, B. O.; Pogliano, K.; Gutierrez, M.; Dorrestein, P. C. (2013) MS/MS networking guided analysis of molecule and gene cluster families *Proc. Natl. Acad. Sci. U. S. A. 110*, E2611-2620.
(14) Brown, D. W.; Adams, T. H.; Keller, N. P. (1996) Aspergillus has distinct fatty acid synthases for primary and secondary metabolism *Proc. Natl. Acad. Sci. U. S. A. 93*, 14873-14877.
(15) Ahuja, M.; Chiang, Y. M.; Chang, S. L.; Praseuth, M. B.; Entwistle, R.; Sanchez, J. F.; Lo, H. C.; Yeh, H. H.; Oakley, B. R.; Wang, C. C. (2012) Illuminating the diversity of aromatic polyketide synthases in *Aspergillus nidulans J. Am. Chem. Soc. 134*, 8212-8221.
(16) Offenzeller, M.; Santer, G.; Totschnig, K.; Su, Z.; Moser, H.; Traber, R.; Schneider-Scherzer, E. (1996) Biosynthesis of the unusual amino acid (4R)-4-[(E)-2-butenyl]-4-methyl-L-threonine of cyclosporin A: enzymatic analysis of the reaction sequence including identification of the methylation precursor in a polyketide pathway *Biochemistry 35*, 8401-8412.
(17) Bushley, K. E.; Raja, R.; Jaiswal, P.; Cumbie, J. S.; Nonogaki, M.; Boyd, A. E.; Owensby, C. A.; Knaus, B. J.; Elser, J.; Miller, D.; Di, Y.; McPhail, K. L.; Spatafora, J. W. (2013) The genome of *Tolypocladium inflatum*: evolution, organization, and expression of the cyclosporin biosynthetic gene cluster *PLoS Genet. 9*, e1003496.
(18) Niehaus, E. M.; Janevska, S.; von Bargen, K. W.; Sieber, C. M.; Harrer, H.; Humpf, H. U.; Tudzynski, B. (2014) Apicidin F: characterization and genetic manipulation of a new secondary metabolite gene cluster in the rice pathogen *Fusarium fujikuroi PloS One 9*, e103336.
(19) Cheng, Y.-Q.; Ahn, J.-H.; Walton, J. D. (1999) A putative branched-chain-amino-acid transaminase gene required for HC-toxin biosynthesis and pathogenicity in *Cochliobolus carbonum Microbiology 145*, 3539-3546.
(20) Chooi, Y. H.; Tang, Y. (2010) Adding the lipo to lipopeptides: do more with less *Chem. Biol. 17*, 791-793.
(21) Dittmann, J.; Wenger, R. M.; Kleinkauf, H.; Lawen, A. (1994) Mechanism of cyclosporin A biosynthesis. Evidence for synthesis via a single linear undecapeptide precursor *J. Biol. Chem.*, *269*, 2841-2846.
(22) Lawen, A.; Zocher, R. (1990) Cyclosporin synthetase. The most complex peptide synthesizing multienzyme polypeptide so far described *J. Biol. Chem.*, *265*, 11355-11360.
(23) Miller, E. D.; Kauffman, C. A.; Jensen, P. R.; Fenical, W. (2007) Piperazimycins: Cytotoxic Hexadepsipeptides from a Marine-Derived Bacterium of the Genus Streptomyces *J. Org. Chem. 72*, 323-330.
(24) Sanchez, J. F.; Entwistle, R.; Corcoran, D.; Oakley, B. R.; Wang, C. C. (2012) Identification and molecular genetic analysis of the cichorine gene cluster in *Aspergillus nidulans MedChemComm*, *3*, 997-1002.
(25) Evans, B. S.; Ntai, I.; Chen, Y.; Robinson, S. J.; Kelleher, N. L. (2011) Proteomics-based discovery of koranimine, a cyclic imine natural product *J. Am. Chem. Soc. 133*, 7316-7319.
(26) Shigemori, H.; Wakuri, S.; Yazawa, K.; Nakamura, T.; Sasaki, T.; Kobayashi, J. i. (1991) Fellutamides A and B, cytotoxic peptides from a marine fish-possessing fungus *Penicillium fellutanum Tetrahedron*, *47*, 8529-8534.
(27) Raaijmakers, J. M.; De Bruijn, I.; Nybroe, O.; Ongena, M. (2010) Natural functions of lipopeptides from Bacillus and Pseudomonas: more than surfactants and antibiotics *FEMS Microbiol. Rev. 34*, 1037-1062.
(28) Shimizu, M.; Fujii, T.; Masuo, S.; Takaya, N. (2010) Mechanism of de novo branched-chain amino acid synthesis as an alternative electron sink in hypoxic *Aspergillus nidulans* cells *Appl. Environ. Microbiol. 76*, 1507-1515.
(29) Yu, J. H.; Hamari, Z.; Han, K. H.; Seo, J. A.; Reyes-Domínguez, Y.; Scazzocchio, C. (2004) Double-joint PCR: a PCR-based molecular tool for gene manipulations in filamentous fungi *Fungal Genet. Biol*. *41*, 973-981.
(30) Calvo, A. M.; Wilson, R. A.; Bok, J. W.; Keller, N. P. (2002) Relationship between secondary metabolism and fungal development *Microbiol. Mol. Biol. Rev. 66*, 447-459.
(31) Bhushan, R.; Bruckner, H. (2004) Marfey's reagent for chiral amino acid analysis: a review *Amino acids 27*, 231-247.
(32) Corey, E. J.; Link, J. O. (1992) A general, catalytic, and enantioselective synthesis of α-amino acids *J. Am. Chem. Soc. 114*, 1906-1908.
(33) Boyer, J. H. (1951) Reduction of Organic Azides to Primary Amines with Lithium Aluminum Hydride *J. Am. Chem. Soc. 73*, 5865-5866.

HDACi + LC/MS

wildtype

VS

HDACi

MS/MS networking

Backward lipopeptides