



<b>Title</b>	Versatility of enzymes catalyzing late steps in polyene 67-121C biosynthesis
<b>Authors(s)</b>	Stephens, Niamh, Rawlings, Bernard, Caffrey, Patrick
<b>Publication date</b>	2013-05-22
<b>Publication information</b>	Stephens, Niamh, Bernard Rawlings, and Patrick Caffrey. "Versatility of Enzymes Catalyzing Late Steps in Polyene 67-121C Biosynthesis." Taylor and Francis, May 22, 2013. <a href="https://doi.org/10.1271/bbb.120961">https://doi.org/10.1271/bbb.120961</a> .
<b>Publisher</b>	Taylor and Francis
<b>Item record/more information</b>	<a href="http://hdl.handle.net/10197/8288">http://hdl.handle.net/10197/8288</a>
<b>Publisher's statement</b>	This is an electronic version of an article published in Bioscience Biotechnology and Biochemistry 77(4): 880-883 (2013). Bioscience Biotechnology and Biochemistry is available online at: <a href="http://www.tandfonline.com/doi/abs/10.1271/bbb.120961">www.tandfonline.com/doi/abs/10.1271/bbb.120961</a> .
<b>Publisher's version (DOI)</b>	10.1271/bbb.120961

Downloaded 2026-05-02 00:30:12

The UCD community has made this article openly available. Please share how this access benefits you. Your story matters! (@ucd\_oa)



© Some rights reserved. For more information

Running title: Enzymes Catalyzing Late Steps in 67-121C Biosynthesis

Versatility of Enzymes Catalyzing Late Steps in Polyene 67-121C Biosynthesis

Niamh STEPHENS,<sup>1</sup> Bernard RAWLINGS,<sup>2</sup> and Patrick CAFFREY<sup>1,†</sup>

*<sup>1</sup>School of Biomolecular and Biomedical Science, Centre for Synthesis and Chemical Biology, University College Dublin, Belfield, Dublin 4, Ireland*

*<sup>2</sup>Department of Chemistry, University of Leicester, University Road, Leicester LE1 7RH, United Kingdom*

Received December 14, 2012; Accepted December 30, 2012

<sup>†</sup>To whom correspondence should be addressed. Fax: +353 1-7161188; E-mail: patrick.caffrey@ucd.ie

*Abbreviations:* ESMS, electrospray mass spectrometry; GT, glycosyltransferase; HPLC, high performance liquid chromatography; PKS, polyketide synthase; LC-MS, liquid chromatography-mass spectrometry

*Actinoplanes caeruleus* produces 67-121C, a heptaene macrolide modified with a D-mannosyl-D-mycosaminyl disaccharide. Draft genome sequencing revealed genes encoding mycosaminyltransferase, mycosamine synthase, a cytochrome P450 that modifies the macrolactone core, and the extending mannosyltransferase. Only the mycosamine synthase and P450 were active in the biosynthesis of amphotericins in *Streptomyces nodosus*, the amphotericin producer.

**Key words:** polyene macrolide biosynthesis; extending glycosyltransferase

Polyene macrolides are important antifungal agents that have severe side effects such as nephrotoxicity. Since very few antifungal drugs are available, there is considerable interest in developing non-toxic analog by genetic engineering of polyene-producing bacteria.<sup>1)</sup>

The biosynthetic gene clusters for several polyene macrolides have been characterized.<sup>1, 2)</sup> The pathway begins with assembly of a macrolactone core, catalyzed by a modular polyketide synthase (PKS). In the late stages, a cytochrome P450 oxidizes a methyl side chain to form an exocyclic carboxyl group, and a glycosyltransferase (GT) catalyzes addition of mycosamine, an amino deoxy sugar derived from GDP-D-mannose. In some cases, another P450 subsequently introduces a hydroxyl group or epoxide.<sup>3,4)</sup> Chemical modification studies have shown that adding sugar residues to polyenes can improve their pharmacological properties.<sup>5-7)</sup> Enzymatic methods of extending natural product glycosylation have been developed in recent years.<sup>8)</sup> The *Streptomyces antibioticus* OleD enzyme catalyzes transfer of a glucosyl residue to the 2'-OH of the desosamine sugar of oleandomycin.<sup>9)</sup> An engineered version of the enzyme has increased acceptor tolerance and glycosylates many diverse compounds *in vitro*. Wild-type OleD adds an extra sugar residue to the tetraene nystatin A1 *in vitro*.<sup>10)</sup> Recent studies have discovered naturally occurring polyenes with

additional sugar residues attached to mycosamine. A disaccharide-modified nystatin has been isolated from *Pseudonocardia* species P1, a symbiont of leaf-cutting ants.<sup>11)</sup> The second sugar residue has not been fully identified, but is thought to be a hexose. The gene for the extending GT has been identified and was named *nypY*. Lee and co-workers characterized another nystatin analog, produced by *Pseudonocardia autotrophica*.<sup>12)</sup> This polyene contains an *N*-acetylglucosamine residue  $\beta$ -1,4 linked to the mycosaminyl sugar residue. The new compound is less hemolytic than nystatin, and more than 300 times more water-soluble. The gene for this extending GT was not found in the main biosynthetic gene cluster.

*Actinoplanes caeruleus* synthesizes 67-121C (**1**), an aromatic heptaene that is modified with a mannosyl-mycosaminyl disaccharide. Co-metabolites such as 67-121A (**2**) are produced in smaller amounts; these contain a single mycosaminyl sugar (Fig. 1).<sup>13, 14)</sup> In this study we identified the late genes involved in 67-121 biosynthesis and assessed the activities of the encoded enzymes towards aglycones of amphotericin B (**3**), a medically important polyene synthesized by *Streptomyces nodosus*. In this bacterium, the amphotericin PKS produces 8-deoxy-16-descarboxyl-16-methyl-amphoterolides B (**4**) and A (**8**). The AmphN P450 forms C-16 carboxyl groups, to give 8-deoxyamphoterolides B (**5**) and A (**9**). The AmphDI GT catalyzes mycosaminylation, and finally AmphL, another P450, catalyzes hydroxylation at C-8 to give amphotericins B and A (**3** and **7**).<sup>15)</sup> Mutant strains deficient in the late steps have been generated.<sup>16,17)</sup> These were used as hosts for heterologous expression of *A. caeruleus* polyene modification genes.

*A. caeruleus* DSM43900 was purchased from Leibniz Institute DSMZ, the German collection of micro-organisms and cell cultures (Braunschweig, Germany). Production of **1** was confirmed by growing starter cultures on GYE medium (2% w/v glucose, 1% w/v yeast extract, and 2% w/v CaCO<sub>3</sub>) and production cultures on fructose-dextrin-soya production medium.<sup>18)</sup> Methanol extracts of the mycelia contained a major heptaene with activity against *Saccharomyces cerevisiae*. The sample was analyzed by LC-MS on a XEVO

instrument with Waters Acquity UP-LC (Waters, Manchester, UK). This revealed a heptaene with a molecular mass expected for **1** ( $[M + H]^+ = 1,289.6431$ ). To obtain the required genes, a draft genome sequencing approach was used rather than traditional library construction and screening. For genomic DNA isolation, *A. caeruleus* was grown on 50 mL of tryptic soy broth at 30°C for 4 d with shaking. Cells were harvested by centrifugation at 10,000 x g for 10 min at 4°C, washed twice with purified water, and resuspended in 15 mL of 50 mM Tris-HCl 50 mM EDTA (pH 8.0) containing 0.5% (v/v) Tween 20, 0.5% (v/v) Triton X-100, and 200 µg/mL RNase. Since *A. caeruleus* is resistant to lysozyme, cells were disrupted in a French pressure cell at 5,000 psi. Then 10 mg of proteinase K was added, and the lysate was incubated at 37°C for 30 min to allow degradation of protein and RNA. The suspension was diluted with 2 volumes of QBT buffer (50 mM MOPS pH 7.0, 750 mM NaCl, 0.15% Triton X-100 and 15% v/v isopropanol) and genomic DNA was isolated using a Qiagen 500/G column. The manufacturer's instructions were followed. Genome sequencing was carried out by BaseClear (Leiden, Netherlands) using an Illumina instrument (Illumina Inc., San Diego, California, USA).

The genome sequence analysis yielded 2,194 contigs with an average size of 3,696 bp. The total length of DNA sequenced was 8,109,345 bp. This was adequate for the present study. Efforts to obtain an improved genome sequence are in progress. The sequences were annotated using the Artemis program (downloaded from <http://www.sanger.ac.uk/resources/software/artemis/>). One 10,697bp contig was found to contain homologs of known mycosaminyltransferase, mycosamine synthase, cytochrome P450 and ferredoxin enzymes and proteins. The highest levels of sequence identity were 68% for GT, 79% for mycosamine synthase, 69% for P450, and 54% for ferredoxin. The *A. caeruleus* genes were named *aceDI*, *aceDII*, *aceN*, and *aceM*, respectively. Adjacent genes encoded an editing thioesterase, a *p*-aminobenzoic acid synthase, and a putative PKS loading module. The order of genes was identical to that in the

FR008/candicidin biosynthetic gene cluster.<sup>19)</sup> BLAST searches were carried out with the remaining contigs to identify the gene encoding the GT that adds the second sugar residue. Another 9,988bp sequence contained the most likely candidate, which was named *pegA* (putative polyene extending glycosyltransferase). The protein was homologous (48% sequence identity overall) to known mycosaminyl transferases. The degrees of sequence identity were 39% for the N-terminal acceptor-binding domain and 60% for the C-terminal region, which binds the GDP-sugar donor substrate.<sup>16)</sup> AmphDI GT can utilize GDP-D-mannose in the absence of GDP-D-mycosamine *in vitro* and *in vivo*.<sup>15,20)</sup> The PegA protein showed end-to-end sequence homology (51% identity overall) with the *Pseudonocardia* Sp1 NypY extending glycosyltransferase, which adds a hexose to the mycosamine of a nystatin polyene (Fig. 2).<sup>11)</sup> In this case the sequence similarity was pronounced for the N-terminal domain, which binds the acceptor (47% identity). Bioinformatic analysis confirmed that PegA catalyzes addition of a mannosyl residue to the mycosaminyl sugar during 67-121C biosynthesis. The *pegA* gene was flanked by genes for transposases. This suggests that the gene was acquired by horizontal transmission. The new sequences have been deposited in the GenBank database under accession no.s KC292379 and KC292380.

The functions of the *A. caeruleus* genes were investigated by heterologous expression in *S. nodosus* strains. The *AceDI-DII-N-M* region was amplified by PCR using high-fidelity Phusion DNA polymerase and primers AceF 5' **GCATCTGCAGGTTTCGTGGGATCAAGGAAGG** 3' and AceR2 5' **GATCAAGCTTGATGTCGCTCATACAAACCGAG** 3'. The *Pst* I and *Hind* III sites incorporated into the primers (in bold type) were used to introduce the amplified DNA between the same sites of the pIAGO expression vector,<sup>16)</sup> to give pIAGO-Ace1. This construct was transformed into *S. nodosus amphDI-DII-NM*, which produces 8-deoxy-16-descarboxyl-16-methyl-amphoterolides A and B (4 and 8, Fig. 1).<sup>17)</sup> Polyenes were extracted and analysed by HPLC, as described

previously.<sup>18)</sup> This revealed that the transformant gave 8-deoxyamphoteronolides A and B (**9** and **5**) as about 4.6% of total polyene. The extent of conversion of **8** to **9** ( Fig. 3B) indicated that AceN P450 weakly recognises the amphotericin aglycones. No glycosylated forms were observed, indicating that AceDI GT did not act efficiently on the amphoteronolide acceptors, but we estimate that less than 2% glycosylation would be undetectable by HPLC. To investigate further, the pIAGO-Ace1 plasmid was transformed into *S. nodosus amphDII*, which lacks mycosamine synthase but retains AmphDI GT and produces 8-deoxyamphoteronolides A and B (**5** and **9**) and 6-deoxyhexosylated amphoteronolides.<sup>16)</sup> An analysis of polyenes from the transformant revealed that efficient production of amphotericin B (**3**) was restored to a level of 47% of total heptaene (Fig. 3D). The AceDII enzyme encounters its normal substrate (GDP-3,6-dideoxy-3-keto-D-mannose) in *S. nodosus* and catalyzes formation of GDP-mycosamine, which is then used by AmphDI. Taken together, the complementation studies of pIAGO-Ace1 indicate that the AceDII and AceN enzymes can function in amphotericin biosynthesis in *S. nodosus*. The upstream *aceDI* gene is transcribed, but 67-121 mycosaminyltransferase does not recognize the amphotericin aglycone. We have found that amphoteronolides are not recognised *in vivo* by the GT that functions in the biosynthesis of perimycin, another aromatic heptaene.<sup>16)</sup> In contrast, AmphDI GT tolerates several different aglycone acceptor substrates.<sup>20 - 22)</sup>

The *pegA* gene was amplified by PCR and cloned into pIAGO. The primers used were PEGF4 5' AGTCAGATCTGCTAAGGAAATTCGGCACCGATC 3' and NGTR1 5' GATCAAGCTTAAGGCACGTCAGCCGGACGGAAG 3'. The amplified DNA was digested with *Bgl* II and *Hind* III and cloned between the *Bam*HI and *Hind* III sites of the vector to give pIAGO-*pegA1*. This construct was introduced into *S. nodosus*, and polyenes were extracted and analyzed by HPLC. The main products were amphotericins A (**7**) and B (**3**). No extra polyene species were detected as convincing peaks in the chromatograms, but ESMS analysis of

the crude polyenes revealed additional trace ions with molecular masses appropriate for mannosyl-amphotericin A (**10**) ( $[M + H]^+ = 1,088.5$ ,  $[M - H]^- = 1,086.5$ ) and mannosyl-amphotericin B (**6**) ( $[M + H]^+ = 1,086.5$ ,  $[M - H]^- = 1,084.5$ ). These were present in about 1-2% of the abundance of the amphotericin A and B ions, and were absent from the mass spectra analyzing extracts of control *S. nodosus* containing the empty pIAGO vector. High resolution MS of the most abundant new species revealed a molecular formula appropriate for mannosyl-amphotericin A (**10**) ( $[M + H]^+ = 1,088.5665$  observed, calculated for  $C_{53} H_{86} N O_{22} = 1,088.5641$ ). These results indicate that PegA is the extending glycosyltransferase that adds the second sugar of 67-121C. Expression of the gene in *S. nodosus* does not lead to the production of high levels of disaccharide-modified amphotericins. This might have been due to failure of recognition of amphotericins A and B by PegA, or to efficient export of these polyenes before mannosylation occurred. Further work is required to investigate these possibilities.

Recombinant GTs are used to glycosylate natural products *in vitro*.<sup>20)</sup> This approach yields only small amounts of glycosylated material, but is a sensitive method of assessing substrate tolerance. Our attempts to overproduce PegA were unsuccessful. A recombinant form with a C-terminal hexahistidine tag was overproduced in *Escherichia coli* from a pET28 construct, but the protein was insoluble. pIAGO constructs were designed to express forms with N-terminal or C-terminal tags in *Streptomyces lividans*. These did not give detectable levels of recombinant proteins, even after Ni-NTA resins were used to enrich the target protein by affinity purification.

This study provides new information on the substrate specificities of glycosyltransferases and cytochrome P450 enzymes that act on aromatic polyenes. Despite the high degree of sequence homology between AmphN and AceN (69% identity), the latter enzyme only weakly recognized amphoteronolide substrates. To date it has not been possible to overproduce an active recombinant cytochrome

P450 that forms an exocyclic carboxyl group in a polyene macrolide. The mechanism of action of these enzymes can be investigated by expressing engineered versions in streptomycete hosts that produce macrolactone substrates. It might be possible to engineer the biosynthesis of amphotericin analogs in which the methyl branch at C-16 is oxidized to a hydroxyl group rather than a carboxylic acid. The *A. caeruleus* GTs showed little activity towards amphotericin substrates. The NypY extending GT or mutated versions of PegA might allow efficient low-cost production of water-soluble disaccharide-modified amphotericins by fermentation methods. Engineering of the GDP-sugar binding domain might lead to the production of analogs containing disaccharides composed of two mycosaminyl residues. These compounds should have improved therapeutic properties because additional positive charges increase antifungal activity.<sup>23)</sup>

### **Acknowledgment**

This work was conducted with the financial support of the Science Foundation Ireland, under grant no. 09/RFP/GEN2132.

### **References**

- 1) Aparicio JF, Caffrey P, Gil JA, and Zotchev SB, *Appl. Microbiol. Biotechnol.*, **61**, 179-188 (2003).
- 2) Caffrey P, Aparicio JF, Malpartida F, and Zotchev SB, *Curr. Top. Med. Chem.*, **8**, 639-653 (2008).
- 3) Volokhan O, Sletta H, Ellingsen TE, and Zotchev SB, *Appl. Environ. Microbiol.*, **72**, 2514-2519 (2006).
- 4) Mendez MV, Anton N, Martin JF, and Aparicio JF, *Biochem. J.*, **386**, 57-62 (2005).
- 5) Szlinder-Richert J, Mazerski J, Cybulska B, Grybowska J, and Borowski E, *Biochim. Biophys. Acta*, **1528**, 15-24 (2001).

- 6) Ehrenfreund-Kleinman T, Azzam T, Falk R, Polacheck L, Golenser J, and Domb AJ, *Biomaterials*, **23**, 1327-1335 (2002).
- 7) Volmer A, Szpilman AM, and Carreira EM, *Nat. Prod. Rep.*, **27**, 1329-1349 (2010).
- 8) Williams GJ, Zhang C, and Thorson JS, *Nat. Chem. Biol.*, **3**, 657 – 662 (2007).
- 9) Quiros LM, Aguirrezabalaga I, Olano C, Mendez C, and Salas JA, *Mol. Microbiol.*, **28**, 1177–1185 (1998).
- 10) Gantt RW, Goff RD, Williams GJ, and Thorson JS, *Angew. Chem. Int. Ed.*, **47**, 8889 –8892 (2008).
- 11) Barke J, Seipke RF, Gruschow S, Heavens D, Drou N, Bibb MJ, Goss RJM, Yu DW, and Hutchings MI, *BMC Biol.*, **8**,109-118 (2010).
- 12) Lee M-J, Kong D, Han K, Sherman DH, Bai L, Deng Z, Lin S, and Kim E-S, *Appl. Microbiol. Biotechnol.*, **95**,157–168 (2012).
- 13) Omura S and Tanaka H “Macrolide Antibiotics: Chemistry, Biology and Practice,” ed. Omura S, Academic Press Inc., Harcourt Brace Jovanovich publishers, New York, pp 351-414 (1984).
- 14) Tamura T, Nakagaito Y, Nishii T, Hasegawa T, Stackebrandt E, and Yokota A, *Int. J. Syst. Bacteriol.*, **44**, 193-203 (1994).
- 15) Byrne B, Carmody M, Gibson E, Rawlings B, and Caffrey P, *Chem. Biol.*, **10**, 1215-1224 (2003).
- 16) Hutchinson E, Murphy B, Dunne T, Breen C, Rawlings B, and Caffrey P, *Chem. Biol.*, **17**, 174-182 (2010).
- 17) Stephens N, Rawlings B, and Caffrey P, *Biosci. Biotechnol. Biochem.*, **76**, 384-387 (2012).
- 18) Murphy B, Anderson K, Borissow C, Caffrey P, Griffith G, Hearn J, Ibrahim O, Khan N, Lamburn N, Lee M, Pugh K and Rawlings B, *Org. Biomol. Chem.*, **8**, 3758-3770 (2010).
- 19) Chen S, Huang X, Zhou X, Bai L, He J, Jeong KJ, Lee SY, and Deng Z, *Chem. Biol.*, **10**, 1065-1076 (2003).

- 20) Zhang C, Moretti R, Jiang J, and Thorson JS, *ChemBioChem*, **9**, 2506-2514 (2008).
- 21) Carmody M, Murphy B, Byrne B, Power P, Rai D, Rawlings B, and Caffrey P, *J. Biol. Chem.*, **280**, 34420-34426 (2005).
- 22) Power P, Dunne T, Murphy B, Nic Lochlainn L, Rai D, Borissow C, Rawlings B, and Caffrey P, *Chem. Biol.*, **15**, 78 - 86 (2008).
- 23) Paquet V and Carreira EM, *Org. Lett.*, **8**, 1807-1809 (2006).

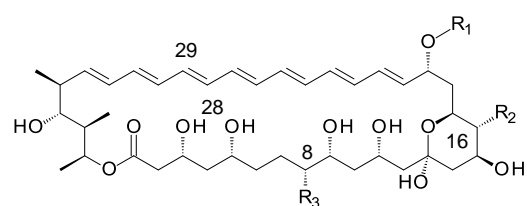
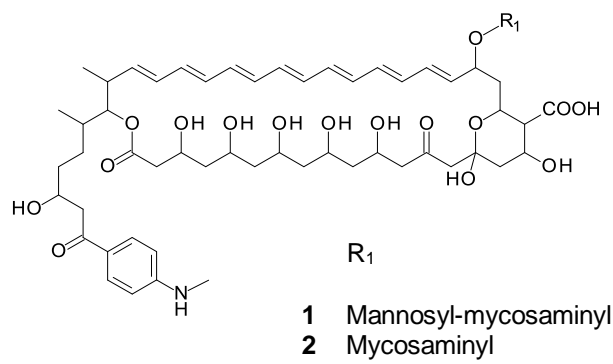
**Fig. 1.** Structures of Polyenes.

**1**, 67-121C; **2**, 67-121A; **3**, amphotericin B; **4**, 8-deoxy-16-descarboxyl-16-methyl-amphoteronolide B; **5**, 8-deoxyamphoteronolide B; **6**, mannosyl amphotericin B; **7**, amphotericin A; **8**, 8-deoxy-16-descarboxyl-16-methyl-amphoteronolide A; **9**, 8-deoxyamphoteronolide A; **10**, mannosyl amphotericin A.

**Fig. 2.** Alignment of *A. caeruleus* PegA with NypY Extending Glycosyltransferase from *Pseudonocardia* sp P1.

**Fig. 3.** Complementation of the *S. nodosus amphDI*, *DII*, *N*, and *M* Mutations by Plasmid-Borne *aceDI-DII-N-M* Genes.

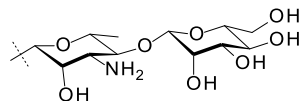
Polyene products were analyzed by HPLC. A C18 column (4.6 x 150 mm) was used with 0.1% (v/v) formic acid (solvent A) and methanol 0.1% (v/v) formic acid (solvent B). For separation of heptaenes, the gradient was 50%-100% B over 30 min at a flow rate of 1 mL/min. Separation of tetraenes (28-29 dihydro analogs) was achieved with a shallower gradient of 70%-100% B at the same run time and flow rate. The absorbance of heptaenes and tetraenes was monitored at 405 nm and 320 nm respectively. A, Tetraenes from *S. nodosus amphDI-DII-N-M* pIAGO. B, Tetraenes from *S. nodosus amphDI-DII-N-M* pIAGO-Ace1. C, Heptaenes from *S. nodosus amphDII* pIAGO. D, Heptaenes from *S. nodosus amphDII* pIAGO-Ace1. Polyene **9** was identified by co-chromatography with 8-deoxyamphoteronolide A isolated from *S. nodosus amphDIII*.<sup>15)</sup> **1** was identified by co-chromatography with commercially available amphotericin B. In addition to **5** and **9**, *S. nodosus amphDII* produced 19-*O*-deoxyhexosyl-amphoteronolide B. This peak is marked with a circle in panels C and D.



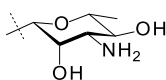
	R <sub>3</sub>	R <sub>2</sub>	R <sub>1</sub>
<b>3</b>	OH	COOH	Mycosaminyl
<b>4</b>	H	CH <sub>3</sub>	H
<b>5</b>	H	COOH	H
<b>6</b>	OH	COOH	Mannosyl-mycosaminyl

**7, 8, 9, 10**

= 28-29 dihydro analogs of **3, 4, 5, 6**



Mannosyl-mycosaminyl



Mycosaminyl

Fig. 1 Stephens et al.

```

PegA      VRVSLQGTGKRPSVGDDEMTSH----RPILFCCTHSTGQATSSSLVLAGEFAARGVKNLWFASDDNHRAAVAGLSGASEVGFVS
NypY      ---MEQTGTTRPADDAAQQADGAGAQPILFCCTHSTGEAATSLVLAGELARRGVPDLVFAADENLRGPVGELADRSAREFVS
          * **.**: . . . :*****:*:*****:* ** * * **.*: * **.*: * *. . * . * . * * **

PegA      TGPVNPRVAPTMMWDETYRSITQSRWKGNRARLRQLMDPVHADRFRQLDAAVQRIQPALMVINNLCIHGIQVAMKRGVPY
NypY      LGPVNPDALTMIDDATYARIHQRSRVRGLRARARQLFDVDHLMQRYQALDEVVERVRPALMVINRFATHAVQVALTRGIPY
          ***** : * ** * * * * * * * * * * : * ** * * * : * * . * : : * * . * : : * * * * * . : . * : * * * . * * : * *

PegA      VITGSSLASDIWQFDLPDDYPVPYSGLPRRMNLRQRIGNRFRHRNLLSMLDPALLRTVGAMLKATDELGLDRRQLQQRAWM
NypY      VITAPCLLSSLVEHDLPRGFPPSSGLPLHRTRRQELERIWFGIGTGLFLDRSVFRKAVRLHRGMGELGIDPRTLVRVPVQN
          * * . . * * * . : : * * * * . : * * * * * : . * * : . * . . * * : : * * . : : . * * * . * * * : .

PegA      EEAEVLVLCFSVFGLDYFFPHPPKLMVGAMVPLPQNVPVGETDRWLDHPSVVFIAFGTITRLTKAEVRAVVDVARRLGER
NypY      AGARSMCLCFSVPGVDYPLVPDRVRMVGALVPPVRHDERDAGVAEWLDAH PSTVYLAFGSITRMTADQVRSVVELARRLGDD
          * . : * * * * * * : * * * * * * * * : : . . . * * * * * * . * : * * * * * * * * : * * * * * * * * :

PegA      HHVLWKLPREQQRFLPPAAELPANLRVEDWLPSQYDVLHNSNRVVFHGGGNSFHEGIYFGKPSLVRPLWFDCLDHAVRAV
NypY      HGVLVLRADQQRFLPDADR PANLKVVDWLHSQHAVLEHPHVRAFFTHGGSNSIHESLWFGTPVLVRPTNVDQYDHAVRAT
          * * * * * * : * * * * * * . : * * * * * * * * * * * * * * : * * * * * * * * * * * * * * * * * * * * * * .

PegA      DSGVGLSVA-PGTMDPAEIH SKLTALLDDTSFRARA EHF AQIQHEAGGVRRADLILRCAAVTSAGQDEQPAVQR
NypY      DTGIGLAVDRPDLVDVDDVHAKLQRLLEPAFTERARELGDVQRAAGGLDTAADA VL-----AELG-NRHPATP-
          * : * * * * * * * . : * * * * * * * * : : * * * * * * * * * * * * * * * * * * * * * * * * :

```

Fig. 2 Stephens et al.

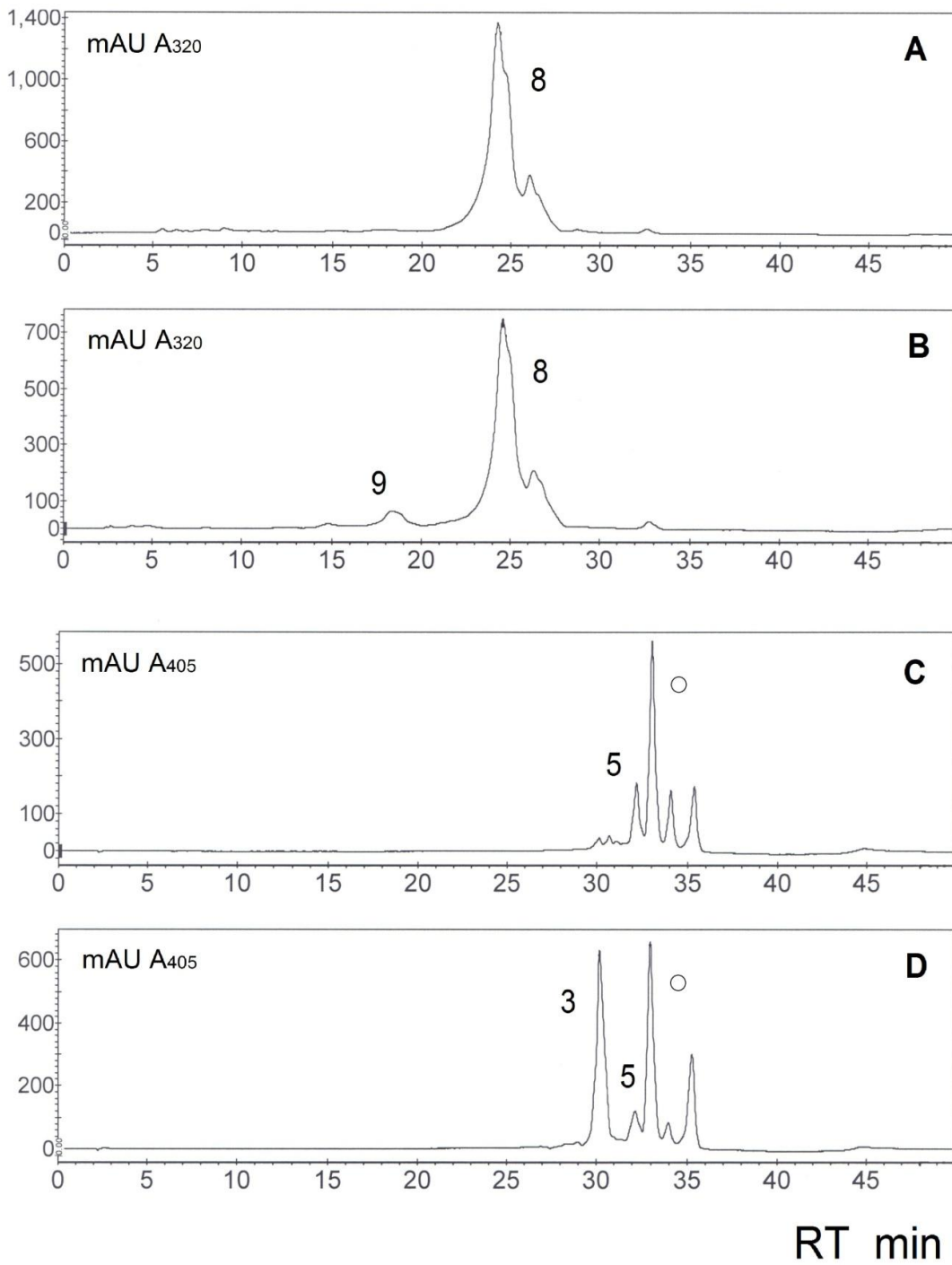


Fig. 3 Stephens et al.