

## Accepted Article

**Title:** Euglenatides, Potent Antiproliferative Cyclic Peptides Isolated from the Freshwater Photosynthetic Microalga *Euglena gracilis*

**Authors:** A Ganesan, Mohammed Aldholmi, Rizwan Ahmad, Daniel Carretero-Molina, Ignacio Pérez-Victoria, Jesús Martín, Fernando Reyes, Olga Genilloud, Léa Goubeyre, Thierry Gefflaut, Hanne Carlsson, Alexei Maklakov, Robert A. Field, Ellis O'Neill, Barrie Wilkinson, and Maria O'Connell

This manuscript has been accepted after peer review and appears as an Accepted Article online prior to editing, proofing, and formal publication of the final Version of Record (VoR). The VoR will be published online in Early View as soon as possible and may be different to this Accepted Article as a result of editing. Readers should obtain the VoR from the journal website shown below when it is published to ensure accuracy of information. The authors are responsible for the content of this Accepted Article.

**To be cited as:** *Angew. Chem. Int. Ed.* **2022**, e202203175

**Link to VoR:** <https://doi.org/10.1002/anie.202203175>

# Euglenatides, Potent Antiproliferative Cyclic Peptides Isolated from the Freshwater Photosynthetic Microalga *Euglena gracilis*

Mohammed Aldholmi,<sup>\*[a]</sup> Rizwan Ahmad,<sup>[a]</sup> Daniel Carretero-Molina,<sup>[b]</sup> Ignacio Pérez-Victoria,<sup>[b]</sup> Jesús Martín,<sup>[b]</sup> Fernando Reyes,<sup>[b]</sup> Olga Genilloud,<sup>[b]</sup> Léa Goubeyre,<sup>[c]</sup> Thierry Gefflaut,<sup>[c]</sup> Hanne Carlsson,<sup>[d]</sup> Alexei Maklakov,<sup>[d]</sup> Ellis O'Neill,<sup>[e]</sup> Robert A. Field,<sup>[f]</sup> Barrie Wilkinson,<sup>[g]</sup> Maria O'Connell,<sup>[h]</sup> A. Ganesan<sup>\*[h]</sup>

- [a] Dr. M. Aldholmi, Dr. R. Ahmad  
Natural Products and Alternative Medicine, College of Clinical Pharmacy, Imam Abdulrahman Bin Faisal University  
Dammam, 31441, Saudi Arabia  
E-mail: mjalldholami@iau.edu.sa
- [b] Dr. D. Carretero-Molina, Dr. I. Pérez-Victoria, Dr. J. Martín, Dr. F. Reyes, Dr. O. Genilloud  
Fundación MEDINA, Centro de Excelencia en Investigación de Medicamentos Innovadores en Andalucía  
Avenida del Conocimiento 34, 18016 Armilla (Granada), Spain  
E-mail: olga.genilloud@medinaandalucia.es
- [c] Dr. L. Goubeyre, Prof. T. Gefflaut  
Université Clermont Auvergne, Clermont Auvergne INP, CNRS, Institut Pascal  
F-63000 Clermont-Ferrand, France  
E-mail: thierry.gefflaut@uca.fr
- [d] Dr. H. Carlsson, Prof. A. Maklakov  
School of Biological Sciences, University of East Anglia, Norwich Research Park  
Norwich NR4 7TJ, United Kingdom  
E-mail: a.maklakov@uea.ac.uk
- [e] Dr. E. O'Neill  
School of Chemistry, University of Nottingham  
Nottingham NG7 2RD, United Kingdom  
E-mail: ellis.o'neill@nottingham.ac.uk
- [f] Prof. R. A. Field  
Manchester Institute of Biotechnology, University of Manchester  
Manchester M1 7DN, United Kingdom  
E-mail: robert.field@manchester.ac.uk
- [g] Prof. B. Wilkinson  
John Innes Centre, Norwich Research Park  
Norwich NR4 7UH, United Kingdom  
E-mail: barrie.wilkinson@jic.ac.uk
- [h] Prof. M. O'Connell, Prof. A. Ganesan  
School of Pharmacy, University of East Anglia, Norwich Research Park  
Norwich NR4 7TJ, United Kingdom  
E-mail: a.ganesan@uea.ac.uk

Supporting information for this article is given via a link at the end of the document

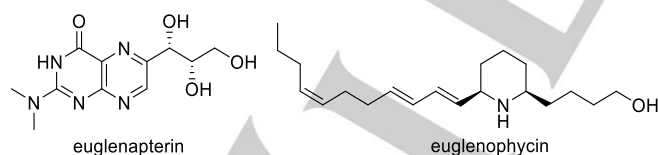
**Abstract:** By limiting the nitrogen source to glutamic acid, we isolated cyclic peptides from *Euglena gracilis* containing asparagine and non-proteinogenic amino acids. Structure elucidation was accomplished through spectroscopic methods, mass spectrometry and chemical degradation. The euglenatides potently inhibit pathogenic fungi and cancer cell lines e.g., euglenatide B exhibiting IC<sub>50</sub> values of 4.3 μM in ketoconazole resistant *Aspergillus fumigatus* and 0.29 μM in MCF-7 breast cancer cells. In an unprecedented convergence of non-ribosomal peptide synthetase and polyketide synthase assembly-line biosynthesis between unicellular species and the metazoan kingdom, euglenatides bear resemblance to nemamides from *Caenorhabditis elegans* and inhibited both producing organisms *E. gracilis* and *C. elegans*. By molecular network analysis, we detected over forty euglenatide-like metabolites in *E. gracilis*, *E. sanguinea* and *E. mutabilis*, suggesting an important biological role for these natural products.

## Introduction

Microorganisms are a bountiful source of biologically active secondary metabolites, especially those with antimicrobial or cytotoxic properties.<sup>[1,2]</sup> Indeed, the treatment of infectious disease and cancer relies heavily on natural products or their semisynthetic or fully synthetic analogues.<sup>[3]</sup> Nevertheless, we have only scratched the surface of microbial metabolite diversity, as a tiny fraction of species are successfully isolated and cultured.<sup>[4,5]</sup> Furthermore, even within this minority, the vast number of secondary metabolite biosynthetic gene clusters (BGCs) remain silent under artificial laboratory conditions.<sup>[6]</sup> Many approaches are being studied for the activation of BGCs including the alteration of fermentation parameters, addition of

small molecule elicitors, expression or deletion of transcriptional modulators and reconstitution in a heterologous host.<sup>[7,8]</sup>

As a case in point, the common freshwater unicellular microalga *Euglena gracilis* is capable of plant-like photoautotrophic growth with light as an energy source and carbon dioxide as the carbon source, animal-like heterotrophic feeding with an external carbon source, or mixotrophically combining the two modes. *E. gracilis* is produced commercially as a food supplement,<sup>[9]</sup> investigated as a potential biofuel source of lipids, carbohydrates and vitamins,<sup>[10]</sup> and even grown with complete replacement of hydrogen by deuterium,<sup>[11]</sup> or in outer space as a potential bioregenerative life support system.<sup>[12]</sup> Despite such intensive scrutiny, the heterocycle euglenapterin (Figure 1) is the only secondary metabolite reported from *E. gracilis*.<sup>[13,14]</sup> In fact, the only natural product of note from the entire *Euglena* genus (>250 species) is the cytotoxic alkaloid euglenophycin, isolated from *E. sanguinea* and detected in several other species but absent in *E. gracilis*.<sup>[15,16]</sup> Nonetheless, partial genome sequencing of *E. gracilis* and transcriptome analysis identified a large number of enzymes involved in secondary metabolism.<sup>[17,18]</sup> This included 19 non-ribosomal peptide synthetases (NRPSs) and 14 polyketide synthases (PKSs), multidomain enzymes involved in assembly-line biosynthesis,<sup>[19,20]</sup> suggesting that *E. gracilis* is a potential source of hitherto undiscovered natural products.



**Figure 1.** Previous secondary metabolites isolated from *Euglena*.

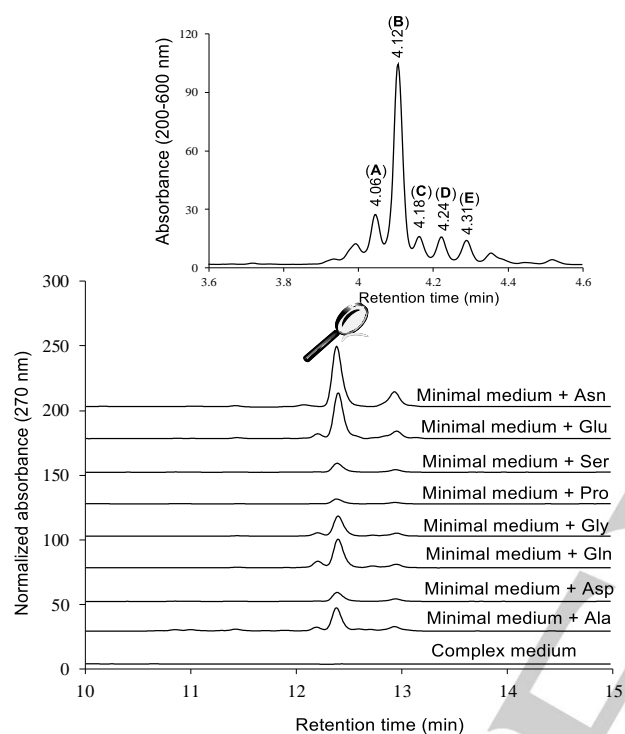
## Results and Discussion

### 1. Discovery of euglenatides

We cultured *Euglena gracilis* var. *saccharophila* Klebs (strain 1224/7A) under photosynthetic and heterotrophic conditions (media composition in Table S1, S2). Disappointingly, only common chlorophyll and xanthophyll photosynthetic pigments were perceived in extracts of either the supernatant or the algal biomass. The addition of small molecule epigenetic modulators, that we used to activate BGCs in *Aspergillus* fungi,<sup>[21]</sup> was also to no avail. A breakthrough came about by considering the role of our nitrogen source. While the standard culture medium contains a complex mixture of amino acids arising from tryptone and yeast extract protein hydrolysates, Oda et al. reported that amino acids can either promote or inhibit growth in *E. gracilis*, with glutamic acid being the most effective nutrient.<sup>[22]</sup> Consequently, we removed tryptone and yeast extract from our medium and instead used a minimal composition supplemented by glutamic acid. Gratifyingly, new metabolites were now observed in the biomass extract of *E. gracilis* grown photosynthetically, and the effect was replicated by several other amino acids, notably asparagine and glutamine (Figure 2). Time course experiments indicated a maximal metabolite production after ten days of cultivation in minimal medium with the addition of 30 mM glutamic acid (Figure S1). Under higher resolution in the HPLC, the main induced peak split into a number of closely overlapping signals (Figure 2). We named the compounds corresponding to the five major peaks as euglenatides A-E in order of chromatographic elution.

From UV-Vis absorption maxima at 260, 269 and 280 nm (Figure S2), we surmised the euglenatides contain a conjugated triene substructure.<sup>[23]</sup> In negative mode electrospray mass spectrometry (ESI-MS, Figure S3), the molecular ions and predicted molecular formulae did not correspond to any known natural products: euglenatide A,  $m/z$  825.4612 [M-H]<sup>-</sup> (calcd. for C<sub>39</sub>H<sub>65</sub>N<sub>6</sub>O<sub>13</sub>, 825.4615,  $\Delta = -0.4$  ppm); euglenatide B  $m/z$  808.4464 [M-H]<sup>-</sup> (calcd. for C<sub>38</sub>H<sub>62</sub>N<sub>7</sub>O<sub>12</sub>, 808.4462,  $\Delta = 0.3$  ppm); euglenatide C  $m/z$  822.4615 [M-H]<sup>-</sup> (calcd. for C<sub>39</sub>H<sub>64</sub>N<sub>7</sub>O<sub>12</sub>, 822.4618,  $\Delta = -0.4$  ppm); euglenatide D  $m/z$  809.4659 [M-H]<sup>-</sup>

(calcd. for  $C_{39}H_{65}N_6O_{12}$ , 809.4666,  $\Delta = -0.9$  ppm); euglenatide E  $m/z$  792.4516  $[M-H]^-$  (calcd. for  $C_{38}H_{62}N_7O_{11}$ , 792.4513,  $\Delta = 0.4$  ppm). In positive mode ESI-MS (Figure S4), in addition to the molecular ion, all euglenatides produced prominent fragments of  $[M-32]^+$  and  $[M-50]^+$  that we attributed to the loss of methanol (32 Da) or methanol + water (50 Da).



**Figure 2.** Overlay of HPLC traces at 270 nm from *E. gracilis* extracts cultured in complex or minimal media supplemented with 30 mM of individual amino acids. Under optimized HPLC conditions, the major induced peak split into multiple overlapping signals.

## 2. Isolation and structure elucidation

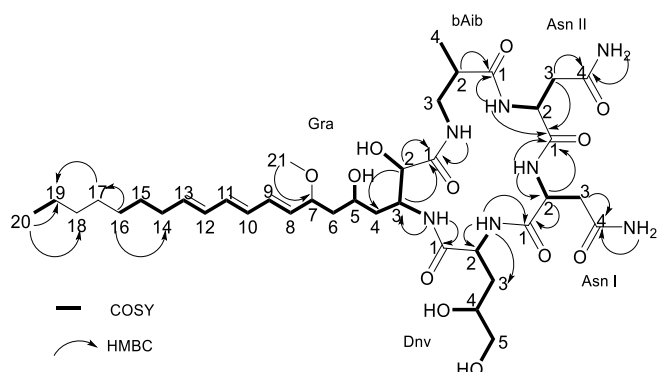
We cultured *E. gracilis* in larger 18 L scale, with minimal medium + 30 mM Glu, at ambient temperature under photosynthetic conditions (irradiation by daylight lamps, 2000 lumens). Following cell lysis and extraction, the residue of 1.7 g was initially purified by silica flash chromatography to give 187 mg of a mixture of euglenatides. Repeated cycles of semi-preparative HPLC finally afforded 2.8 mg of pure euglenatide A, 5.7 mg of euglenatide B, 2.8 mg of euglenatide C, 2.3 mg of euglenatide D

and 2.4 mg of euglenatide E. IR absorptions at  $3347\text{ cm}^{-1}$  (NH) and  $1651\text{ cm}^{-1}$  (CO) revealed the presence of amide carbonyl groups while the broad peak at  $2873\text{ cm}^{-1}$  (CH) was characteristic of an aliphatic chain.

The molecular connectivity of euglenatides was established by extensive 1D and 2D-NMR experiments (Figures S5-S40). For euglenatide B (Table S3), the presence of nine amide protons at  $\delta_H$  6.88 (1H, s), 6.90 (1H, d), 7.06 (1H, s), 7.29 (1H, s), 7.36 (1H, s), 7.65 (1H, d), 7.77 (1H, d), 7.78 (1H, s) and 9.02 (1H, d) in the  $^1H$  NMR spectrum, and seven carbonyl carbon signals at  $\delta_C$  176.3, 173.4, 172.4, 172.3, 172.1, 171.1 and 169.6 implied a peptide backbone. In addition, overlapping signals at  $\delta_H$  1.25 in the  $^1H$  NMR spectrum and a cluster of methylene carbon signals at  $\delta_C$  28.5-31.2 in the HSQC spectrum, and a methyl group at  $\delta_H$  0.85 (3H, t) and  $\delta_C$  13.9 indicated the presence of a long aliphatic chain. The signals at  $\delta_H$  4.31, 4.38, 4.44 and 5.13 were diagnostic of four hydroxyls while the singlet at  $\delta_H$  3.13, accompanied by  $\delta_C$  55.5, indicated a methoxy group. A triene, as inferred from the UV-Vis spectrum, was confirmed by six proton signals at  $\delta_H$  5.51 (1H, dd), 5.72 (1H, dt), 6.07 (1H, dd), 6.15 (1H, dd), 6.20 (1H, dd), 6.24 (1H, dd), and six carbon signals at  $\delta_C$  135.1, 133.9, 132.8, 131.3, 130.3, 130.0. Analysis

of  $^1H$ ,  $^{13}C$ , and HSQC data, in addition to key COSY and HMBC correlations, unveiled the presence of two asparagine residues (Asn I and Asn II) and three non-proteinogenic amino acids:  $\beta$ -aminoisobutyric acid ( $\beta$ Aib), 4,5-dihydroxynorvaline (Dnv) and the novel  $\beta$ -amino-2,5-dihydroxy-7-methoxy-8,10,12-ecosatrienoic acid residue that we named graciline (Gra) after its origin from *E. gracilis*. The sequence of amino acid residues was established by correlations of the NH protons to the amide carbonyls and  $CH_\alpha$  and  $CH_\beta$  protons in the amino acid side chains in HMBC and COSY spectra, resulting in the complete 2D structure of euglenatide B (Figure 3). The allylic loss of methanol and water assisted by the conjugated triene is

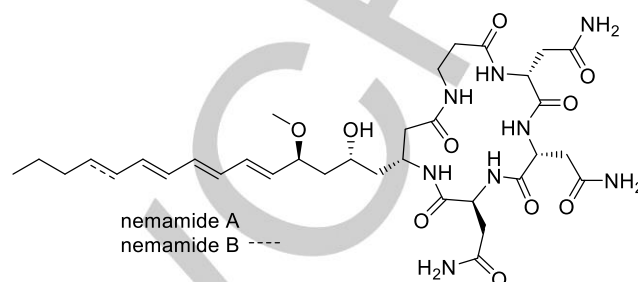
presumably a facile process, accounting for the resulting fragment ions observed in positive mode mass spectrometry.



**Figure 3.** Key COSY and HMBC correlations used to establish the molecular connectivity of euglenatide B.

The non-proteinogenic amino acid  $\beta$ Aib is rarely found within natural products, previously reported in only the guanidine phascolosimine from the sipunculid worm *Phascolion strombi*,<sup>[24]</sup> and the lipopeptides YM-170320 and fusaristatins from fungal *Candida* and *Fusarium* sp. respectively.<sup>[25,26]</sup> To the best of our knowledge, euglenatide B is the first natural product identified to contain Dnv, although the *O*-carbamoyl derivative is present in the antimicrobial *Streptomyces* polyoxin nucleoside peptides and GE81112 tetrapeptides.<sup>[27,28]</sup> Meanwhile, to our surprise, the long chain C20 lipophilic  $\beta$ -amino acid Gra bears a resemblance to the C16 amino acid found within the nemamides isolated by the Butcher group from the nematode *Caenorhabditis elegans* (Figure 4).<sup>[29]</sup> The nemamides are the product of the multimodule megasynthases PKS-1 and NRPS-1 in *C. elegans*, and the first example of such assembly-line biosynthesis in a metazoan. Furthermore, the presence of PKS-1 and NRPS-1 homologs in many nematode genomes implies an important physiological function for these peptides. It is remarkable that nemamides, the only known metazoan PKS-NRPS metabolites, should possess structural similarity in the lipophilic region to euglenatides of microalgal origin. While the

rest of the nemamide backbone contains three Asn residues, the euglenatides are more complex and contain additional stereocenters with one Asn replaced by Dnv and  $\beta$ Ala substituted by the chiral  $\beta$ Aib in euglenatide B.



**Figure 4.** The structures of nemamides A and B.

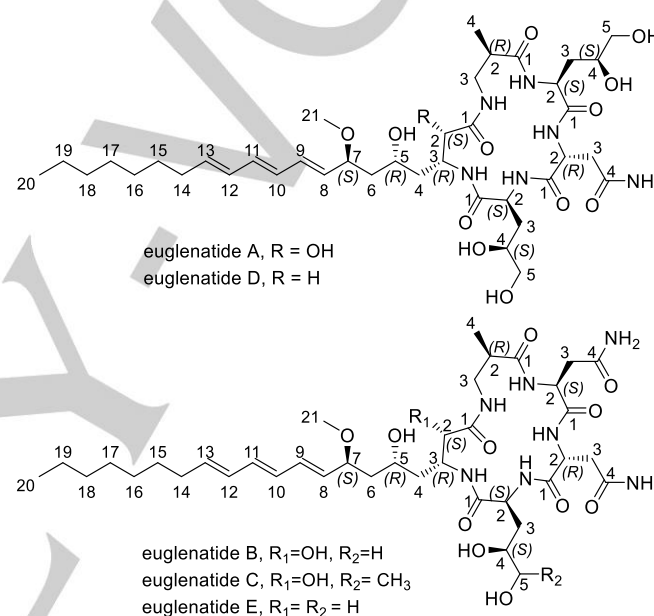
The structures of euglenatides A, C, D and E were elucidated through their relationship to euglenatide B. In brief (Figure S41, Tables S4-S7), euglenatide A contained the same amino acids as euglenatide B, except that Asn II was replaced by a second Dnv residue. Euglenatides D and E were deoxy analogues of euglenatides A and B, respectively, without the C2-hydroxyl group of the graciline residue. The NMR spectrum of euglenatide C was unique relative to the others, as it contained an extra methyl group ( $^1\text{H}$ ,  $\delta$  1.02 ppm;  $^{13}\text{C}$ ,  $\delta$  19.5 ppm), and COSY/HMBC correlations assigned this to the replacement of Dnv by 4,5-dihydroxynorleucine (Dnl).

Numerous attempts to obtain diffraction quality X-ray crystals of euglenatide B were unsuccessful. The stereochemistry of Asn residues was instead determined through their conversion to aspartic acid by microscale acidic degradation. The hydrolysate was subjected to Marfey's analysis by chemical derivatization with *N* $\alpha$ -(2,4-dinitro-5-fluorophenyl)-L-valinamide (L-FDVA).<sup>[30]</sup> Both L-FDVA-L-Asp (RT= 20.66 min) and L-FDVA-D-Asp (RT= 25.52 min) were found, indicating one each of L-Asn and D-Asn in euglenatide B, whereas only the D amino acid was detected from euglenatide A (Figure S42).

Previously, the configurational assignment of  $\beta$ Aib by Marfey analysis was complicated by difficult HPLC separation of the two diastereomers and required mathematical peak fitting by Gaussian deconvolution.<sup>[31]</sup> By careful optimization of conditions including mobile and stationary phase, temperature, flow rate and isocratic solvent system, we devised a new protocol with sufficient chromatographic resolution: L-FDVA-L- $\beta$ Aib, RT= 59.47 min and L-FDVA-D- $\beta$ Aib, RT= 56.98 min (Figure S43). This confirmed that both euglenatides A and B contain the (*R*)-enantiomer, identical in stereochemistry to naturally occurring  $\beta$ Aib isolated from bacterial, fungal or plant sources. On the other hand, in humans,  $\beta$ Aib is a catabolite of thymine and valine metabolism and both enantiomers are detected in plasma.<sup>[32]</sup>

Determination of the Dnv stereochemistry was challenging due to the lack of commercial standards and we employed a bi-enzymatic recycling cascade synthesis from Ala involving pyruvate-aldolase and L- or D- $\alpha$ -transaminase.<sup>[33]</sup> This resulted in two scalemic samples: the L-series contained L-Ala (70%) + L-*syn*-Dnv (26%) + L-*anti*-Dnv (4%) while the D-series contained D-Ala (25%) + D-*syn*-Dnv (9%) + D-*anti*-Dnv (66%). Marfey analysis with these samples resulted in clear separation of the four Dnv diastereomers: L-FDVA-L-*anti*-Dnv, RT= 17.03 min; L-FDVA-L-*syn*-Dnv, RT= 32.38 min; L-FDVA-D-*anti*-Dnv, RT= 24.92 min; L-FDVA-D-*syn*-Dnv, RT= 29.90 min (Figure S44). Our hydrolysates were found to produce L-FDVA-L-*anti*-Dnv, and further confirmation came from enantiopure Boc-L-*anti*-Dnv(4-TBDMS)-OEt, an intermediate in GE81112A total synthesis kindly provided by Dr. Armin Bauer at Sanofi-Aventis.<sup>[34]</sup> Global deprotection of this material to Dnv followed by Marfey's analysis gave an identical result to our sample. We note that of the four diastereomers of Dnv, only L-*anti*-Dnv or its *O*-carbamoyl derivative are currently observed in natural products. Neither the intact Gra residue nor simpler fragments were recognized from our acid hydrolysis, due to the lability of this complex amino acid under these conditions. The <sup>1</sup>H and <sup>13</sup>C

NMR chemical shifts, coupling constants and NOESY correlations (Tables S8 – S11) of 2-deoxyGra in euglenatides D and E closely matched the data for nemamide A, as did the CD spectra (Figure S45). We conclude that 2-deoxyGra is identical in relative and absolute stereochemistry with the nemamide side chain. The chirality of the additional C2 hydroxyl group in graciline was deduced from comparison of *J* values with nemamide A. Combining all the above information completed the structure elucidation of euglenatides A-E (Figure 5).



**Figure 5.** Structures of euglenatides A-E.

### 3. Antiproliferative activity of euglenatides

The euglenatides were tested for antibacterial activity with methicillin-sensitive *Staphylococcus aureus* (MSSA ATCC 29213), methicillin-resistant *S. aureus* (MRSA MB5393) and *Escherichia coli* ATCC 25922 but were inactive at the highest tested concentration of 128  $\mu$ g/mL (Figure S46). On the other hand, antifungal activity was observed against the yeast *Candida albicans* ATCC 64124. Euglenatides B, C and E had IC<sub>50</sub> values of 13-14  $\mu$ M, whereas euglenatide A (27  $\mu$ M) and D (40  $\mu$ M) were less active (Figure S47). All five euglenatides had a higher activity against the mold *Aspergillus fumigatus* ATCC 46645, with similar IC<sub>50</sub> values of 4-9  $\mu$ M. Since this strain is

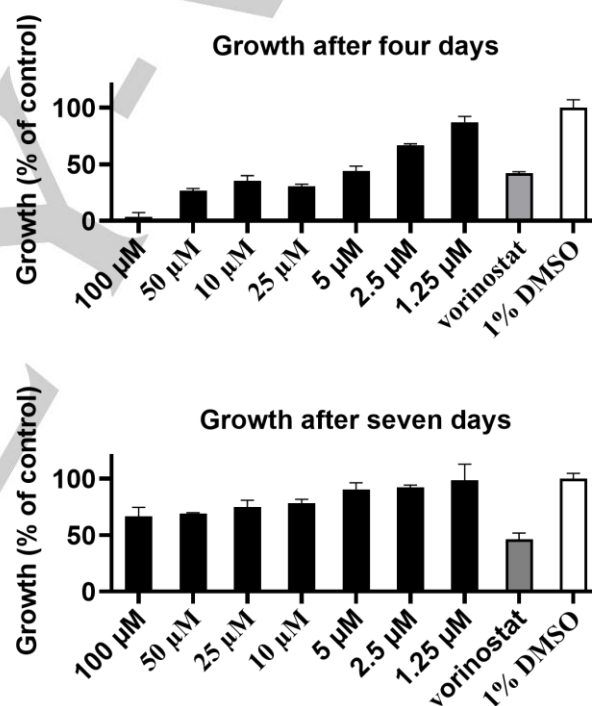
ketoconazole resistant, it signifies the euglenatides are not cross-resistant with the clinically important azole class of antifungal agents. The lack of activity against bacteria suggested a specific target in eukaryotic cells absent in prokaryotes, and we screened the most abundant member, euglenatide B, in several human cancer cell lines. Potent antiproliferative activity was noted with respective  $IC_{50}$  values of 533, 773 and 292 nM against the THP-1 acute monocytic leukemia, A-549 lung adenocarcinoma and MCF-7 breast cancer cell lines (Figure S48).

The Butcher group did not report bioassays with the worm nemamides, perhaps due to the paucity of material as only 0.07 mg of nemamide A and even less of nemamide B was isolated from 50 L of nematode culture. Instead, *C. elegans* mutants that do not produce nemamides were discerned to be defective in recovery from arrest of the L1 larval stage under starvation conditions. Based on the similarity of the lipophilic side chain between nemamides and euglenatides, we predict the former would possess antiproliferative properties against fungal and mammalian cancer cells.

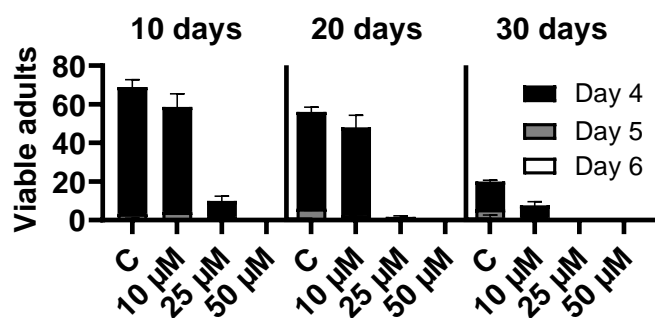
#### 4. Effects on producing organisms

Given the biological activity against eukaryotes, we were curious if euglenatides would have an effect on the producer *E. gracilis* itself. After four days of algal incubation with euglenatide B, growth was significantly inhibited (Figure 6). The effect appears to be cytostatic rather than cytotoxic, as the algae recovered after a longer period of seven days. The question then arises whether the nemamides have similar bioactivity against their nematode producer *C. elegans*. Although the worm nemamides are unavailable for testing, we indirectly addressed this issue by using our euglenatides. At concentrations of 10 and 25  $\mu$ M, euglenatide B was strongly inhibitory to starvation recovery of the L1 larval stage of *C. elegans* (Figure 7).

It is tempting to speculate that the nemamides in *C. elegans* and the euglenatides in *E. gracilis* perform similar biological functions. While low concentrations might serve regulatory purposes and facilitate recovery from starvation conditions without negative effects on the producing organism, the external addition of higher amounts is apparently deleterious and overcomes mechanisms of self-resistance.<sup>[35,36]</sup> In our *Euglena* cultures, we did not detect euglenatides in the supernatant, implying that they are not secreted into the external environment. It is possible the natural products are sequestered in a specific cellular compartment to reduce self-exposure and serve as antifeedants that discourage predators.



**Figure 6.** Algastatic effects of euglenatide B (1.25  $\mu$ M to 100  $\mu$ M) on the growth of *E. gracilis* after four and seven days of treatment. Vorinostat (1 mM) and DMSO (1%) were used as positive and negative controls respectively. The data represent the average  $\pm$  standard error of three replicates.



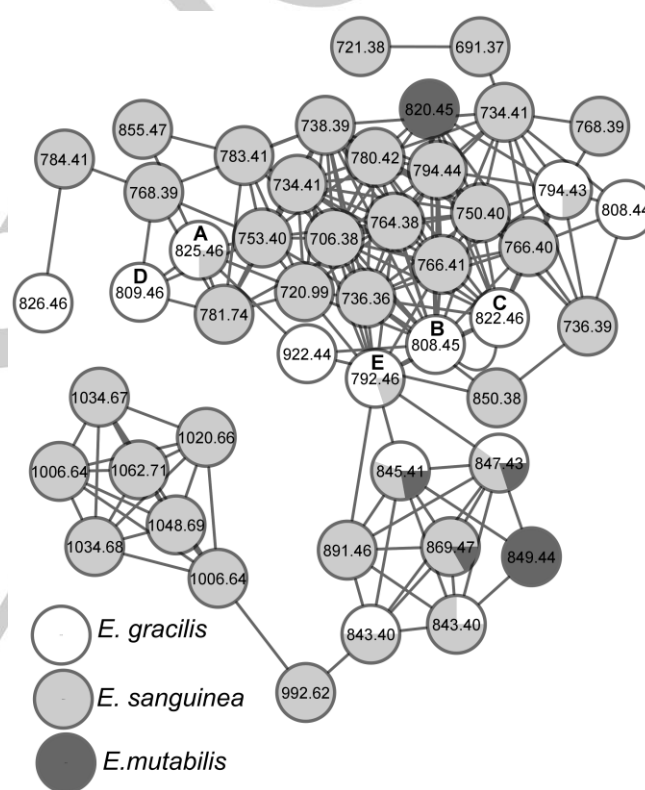
**Figure 7.** Inhibition of L1 *C. elegans* recovery by euglenatide B (10  $\mu$ M to 50  $\mu$ M, C = control) after starvation for 10, 20 or 30 days. The data represent three replicates with 100 eggs per treatment, based on the total number of sexually mature adults recovered by day four, five and six, respectively.

### 5. A common biosynthetic pathway in *Euglena*

We cultivated *E. sanguinea* and *E. mutabilis* strains under a variety of conditions and subjected extracts from cultures with good or excellent growth to metabolomic profiling by LC-MS. The UV-Vis profile of the *E. sanguinea* extract was similar to that of *E. gracilis*, while the *E. mutabilis* extract did not have significant UV-Vis absorptions in the 200-600 nM range (Figures S49, S50). This suggests that both *E. gracilis* and *E. sanguinea* produce euglenatides with conjugated trienes, while they are either absent or in very low amounts in *E. mutabilis*.

MS/MS spectra of *E. gracilis*, *E. sanguinea* and *E. mutabilis* extracts were obtained in negative ion mode and used to construct a molecular network of related metabolites (Figure 8).<sup>[37]</sup> Each node in the molecular network is labelled with the precursor mass of the molecule deduced from the corresponding MS/MS spectrum. The constructed molecular network groups the structurally similar metabolites from three *Euglena* species in a cluster. The higher the similarity, the closer they are in the cluster. For example, precursor masses 808 and 822 represent the MS/MS spectra of euglenatides B and C which differ only by the addition of a methyl group, while precursor masses of 809 and 825 represent euglenatides A and D that differ by an additional hydroxyl group. Our network analysis indicates the presence of over 40 euglenatide-like metabolites within these

three species, including ubiquitous examples such as 845 and 847 and others such as 849 and 1006, detected only in *E. mutabilis* and *E. sanguinea*, respectively. The broad distribution, and the multitude of family members, point to an important function for euglenatides and we searched the available genomic *Euglena gracilis* data for homology with the nemamide BGC.<sup>[38]</sup> However, due to the incomplete assembly, and the likely absence of gene clustering in *Euglena*, we did not identify a candidate BGC for euglenatide biosynthesis.



**Figure 8.** Molecular network analysis of euglenatides produced by *E. gracilis* (white circles), *E. sanguinea* (gray) and *E. mutabilis* (black). Each node represents the precursor mass of a single metabolite. The color of the nodes indicates the relative abundance of each metabolite between strains based on the MS ion intensities of precursors. The isolated euglenatides are represented by A, B, C, D and E respectively.

### Conclusion

We demonstrate that even a species like *Euglena gracilis*, subject to scientific study for centuries, harbors unusual and novel secondary metabolites. Ultimately, our successful unmasking of the euglenatides relied on the traditional 'one



strain, many compounds (OSMAC)' approach of altering fermentation conditions.<sup>[39]</sup> Specifically, culture in minimal media with glutamic acid as the sole nitrogen source served to activate a normally quiescent biosynthetic pathway. The resulting euglenatides contain L- and D-asparagine, as well as (*R*)- $\beta$ -aminoisobutyric acid, a rare constituent in peptide natural products. In addition, the euglenatides are the first peptide natural products disclosed to incorporate 4,5-dihydroxynorvaline, 4,5-dihydroxynorleucine and the novel C20 polyketide derived  $\beta$ -amino acid graciline with a conjugated triene and two oxygenated stereocenters in the side chain. Interestingly, the *E. sanguinea* metabolite euglenophycin (Figure 1) also contains 20 carbons, and it is plausible that it shares a common C20 polyunsaturated fatty acid precursor with the euglenatides.

Euglenatides exhibited significant antiproliferative activity against *Candida* and *Aspergillus*, the two major pathogenic fungal species, and a panel of human cancer cells. By molecular network analysis, we detected over 40 euglenatide-like metabolites in three *Euglena* species, implying a widespread occurrence within the genus. The euglenatides share a structural similarity to the nemamides, the only known metazoan natural product of NRPS-PKS origin, and inhibited the growth of both producing organisms *E. gracilis* and *C. elegans*. The astonishing structural congruence between protist and animal kingdom secondary metabolites is unique in chemical ecology. The ability to harvest multimg quantities of euglenatides from *E. gracilis* will expedite further research into the biological ramifications of these cyclic peptides.

## Acknowledgements

This work was funded through a scholarship to MA from Imam Abdulrahman Bin Faisal University, Saudi Arabia. We are incredibly grateful to Dr Armin Bauer (Sanofi-Aventis, Germany) for the kind gift of Boc-L-*anti*-4,5-dihydroxynorvaline(4-TBDMS)-OEt. We thank Dr Mercedes de la Cruz and Dr Caridad Díaz (Fundación MEDINA, Spain) for their assistance with the antimicrobial assays.

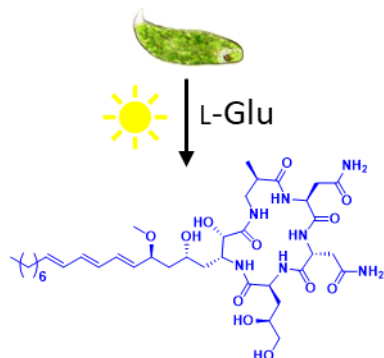
**Keywords:** natural products • cyclic peptides • microalgae • nematodes • antiproliferative

- [1] F. Ntie-Kang, D. Svozil, *Phys. Sci. Rev.* **2020**, *5*, 20180121.
- [2] A. S. Abdel-Razek, M. E. El-Naggar, A. Allam, O. M. Morsy, S. I. Othman, *Processes* **2020**, *8*, 470.
- [3] D. J. Newman, G. M. Cragg, *J. Nat. Prod.* **2020**, *83*, 770–803.
- [4] L. C. Vitorino, L. A. Bessa, *Divers.* **2018**, *10*, 46.
- [5] W. H. Lewis, G. Tahon, P. Geesink, D. Z. Sousa, T. J. G. Ettema, *Nat. Rev. Microbiol.* **2020**, *19*, 225–240.
- [6] P. N. Tran, M. R. Yen, C. Y. Chiang, H. C. Lin, P. Y. Chen, *Appl. Microbiol. Biotechnol.* **2019**, *103*, 3277–3287.
- [7] H. A. Tomm, L. Ucciferri, A. C. Ross, *J. Ind. Microbiol. Biotechnol.* **2019**, *46*, 1381–1400.
- [8] B. C. Covington, F. Xu, M. R. Seyedsayamdost, *Annu. Rev. Biochem.* **2021**, *90*, 763–788.
- [9] K. Suzuki, *Adv. Exp. Med. Biol.* **2017**, *979*, 285–293.
- [10] A. Gissibl, A. Sun, A. Care, H. Nevalainen, A. Sunna, *Front. Bioeng. Biotechnol.* **2019**, *7*, 108.
- [11] S. E. Mandeville, H. L. Crespi, J. J. Katz, *Science* **1964**, *146*, 769–770.
- [12] D. Häder, *Front. Plant Sci.* **2020**, *10*, 1621.
- [13] M. Böhme, W. Pfeleiderer, E. F. Elstner, W. J. Richter, *Angew. Chem. Int. Ed. Engl.* **1980**, *19*, 473–474.
- [14] S. Sasso, G. Pohnert, M. Lohr, M. Mittag, C. Hertweck, *FEMS Microbiol. Rev.* **2012**, *36*, 761–785.
- [15] P. V. Zimba, P. D. Moeller, K. Beauchesne, H. E. Lane, R. E. Triemer, *Toxicol.* **2010**, *55*, 100–104.
- [16] P. V. Zimba, I.-S. Huang, D. Gutierrez, W. Shin, M. S. Bennett, R. E. Triemer, *Harmful Algae* **2017**, *63*, 79–84.
- [17] E. C. O'Neill, M. Trick, L. Hill, M. Rejzek, R. G. Dusi, C. J. Hamilton, P. V. Zimba, B. Henrissat, R. A. Field, *Mol. Biosyst.* **2015**, *11*, 2808–2820.
- [18] T. E. Ebenezer, M. Zoltner, A. Burrell, A. Nenarokova, A. M. G. Novák Vanclová, B. Prasad, P. Soukal, C. Santana-Molina, E. O'Neill, N. N. Nankissoor, et al., *BMC Biol.* **2019**, *17*, 11.
- [19] R. D. Süßmuth, A. Mainz, *Angew. Chem. Int. Ed.* **2017**, *56*, 3770–3821.
- [20] A. Nivina, K. P. Yuet, J. Hsu, C. Khosla, *Chem. Rev.* **2019**, *119*, 12524–12547.
- [21] M. Aldholmi, B. Wilkinson, A. Ganesan, *J. Antibiot.* **2020**, *73*, 410–413.
- [22] Y. Oda, Y. Nakano, S. Kitaoka, *J. Gen. Microbiol.* **1982**, *128*, 853–858.
- [23] K. A. Martínez, C. Lauritano, D. Druka, G. Romano, T. Grohmann, M. Jaspars, J. Martín, C. Díaz, B. Cautain, M. de la Cruz, et al., *Mar. Drugs* **2019**, *17*, 385.
- [24] Y. Guillou, Y. Robin, *J. Biol. Chem.* **1973**, *248*, 5668–5672.
- [25] T. Sugawara, A. Tanaka, K. Tanaka, K. Nagai, K. Suzuki, T. Suzuki, *J. Antibiot.* **1998**, *51*, 435–438.
- [26] Y. Shiono, M. Tsuchinari, K. Shimanuki, T. Miyajima, T. Murayama, T. Koseki, H. Laatsch, T. Funakoshi, K. Takanami, K. Suzuki, *J. Antibiot.* **2007**, *60*, 309–316.
- [27] K. Isono, K. Asahi, S. Suzuki, *J. Am. Chem. Soc.* **1969**, *91*, 7490–7505.

- [28] L. Brandi, A. Lazzarini, L. Cavaletti, M. Abbondi, E. Corti, I. Ciciliato, L. Gastaldo, A. Marazzi, M. Feroggio, A. Fabbretti, et al., *Biochemistry* **2006**, *45*, 3692–3702.
- [29] Q. Shou, L. Feng, Y. Long, J. Han, J. K. Nunnery, D. H. Powell, R. A. Butcher, *Nat. Chem. Biol.* **2016**, *12*, 770–772.
- [30] S. Sethi, J. Martens, R. Bhushan, *Biomed. Chromatogr.* **2021**, *35*, e4990.
- [31] H. Vemula, Y. Kitase, N. J. Ayon, L. Bonewald, W. G. Gutheil, *Anal. Biochem.* **2017**, *516*, 75–85.
- [32] J. Stautemas, A. B. P. Van Kuilenburg, L. Stroomer, F. Vaz, L. Blancquaert, F. B. D. Lefevere, I. Everaert, W. Derave, *Front. Physiol.* **2019**, *10*, 1240.
- [33] C. Guérard-Hélaine, E. Heuson, M. Ndiaye, L. Gourbeyre, M. Lemaire, V. Hélaine, F. Charmantray, J.-L. Petit, M. Salanoubat, V. De Berardinis, et al., *Chem. Commun.* **2017**, *53*, 5465–5468.
- [34] G. Jürjens, S. M. M. Schuler, M. Kurz, S. Petit, C. Couturier, F. Jeannot, F. Nguyen, R. C. Wende, P. E. Hammann, D. N. Wilson, et al., *Angew. Chem. Int. Ed.* **2018**, *57*, 12157–12161.
- [35] S. Li, P. Wang, W. Yuan, Z. Su, S. H. Bullard, *Sci. Rep.* **2016**, *6*, 29315.
- [36] K. H. Almabruk, L. K. Dinh, B. Philmus, *ACS Chem. Biol.* **2018**, *13*, 1426–1437.
- [37] A. T. Aron, E. C. Gentry, K. L. McPhail, L. F. Nothias, M. Nothias-Esposito, A. Bouslimani, D. Petras, J. M. Gauglitz, N. Sikora, F. Vargas, et al., *Nat. Protoc.* **2020**, *15*, 1954–1991.
- [38] L. Feng, M. T. Gordon, Y. Liu, K. B. Basso, R. A. Butcher, *Nat. Commun.* **2021**, *12*, 4912.
- [39] H. B. Bode, B. Bethe, R. Höfs, A. Zeeck, *ChemBioChem* **2002**, *3*, 619.

## Entry for the Table of Contents

Graphic for Table of Contents:



Insert text for Table of Contents.

Light and glutamic acid induced *E. gracilis* to produce euglenatides with antifungal and anticancer activity. Euglenatides were also detected in *E. sanguinea* and *E. mutabilis*. They resemble the nemamides isolated from *C. elegans*, inhibited both producing species, and represent uniquely convergent assembly-line biosynthesis between unicellular organisms and metazoans.

Institute and/or researcher Twitter usernames: uniofeastanglia, maldholmi