

SCIENTIFIC REPORTS



OPEN

New Zealand glowworm (*Arachnocampa luminosa*) bioluminescence is produced by a firefly-like luciferase but an entirely new luciferin

Oliver C. Watkins^{1,2}, Miriam L. Sharpe¹, Nigel B. Perry² & Kurt L. Krause¹

The New Zealand glowworm, *Arachnocampa luminosa*, is well-known for displays of blue-green bioluminescence, but details of its bioluminescent chemistry have been elusive. The glowworm is evolutionarily distant from other bioluminescent creatures studied in detail, including the firefly. We have isolated and characterised the molecular components of the glowworm luciferase-luciferin system using chromatography, mass spectrometry and ¹H NMR spectroscopy. The purified luciferase enzyme is in the same protein family as firefly luciferase (31% sequence identity). However, the luciferin substrate of this enzyme is produced from xanthurenic acid and tyrosine, and is entirely different to that of the firefly and known luciferins of other glowing creatures. A candidate luciferin structure is proposed, which needs to be confirmed by chemical synthesis and bioluminescence assays. These findings show that luciferases can evolve independently from the same family of enzymes to produce light using structurally different luciferins.

Glowworms are found in New Zealand and Australia, and are a major tourist attraction at sites located across both countries. In contrast to luminescent beetles such as the firefly (Coleoptera), whose bioluminescence has been well characterised (reviewed by ref.¹), the molecular details of glowworm bioluminescence have remained elusive. These glowworms are the larvae of fungus gnats of the genus *Arachnocampa*, with eight species endemic to Australia and a single species found only in New Zealand². The New Zealand *Arachnocampa luminosa* (Skuse) (Diptera: Keroplatidae) is found in damp areas of forest, by streams and in caves throughout the country (the Māori name *titiwai* means “projected over water”). The larvae attract prey using blue-green light emitted from a specialised light organ found at the posterior end of the body, derived from modified Malpighian tubules^{3,4}.

Bioluminescent systems create light through the oxidation of a small molecule substrate, generically known as a luciferin, catalysing the reaction using an enzyme called a luciferase^{5,6}. The reaction produces a high-energy singlet state intermediate, which is able to release energy as visible light, or pass it onto a secondary emitter^{6,7}. There are over 40 different bioluminescent systems^{1,8}, with a great variety in the structure, mechanism and substrate specificity of the luciferase enzymes (reviewed by ref.⁹) and also in the structures of the luciferins. Nine different luciferin structures have been solved^{10,11}, most built around four natural product classes: aromatic amino acids, flavins, tetrapyrroles and quinones.

The fireflies mostly produce yellow-green light (λ_{\max} 552–582 nm) through the oxidative adenylation of firefly D-luciferin¹. This reaction is catalysed by firefly luciferase, a member of the ANL superfamily of adenyating enzymes (Acyl-CoA synthetases, Nonribosomal peptide synthetase adenylation domains, beetle Luciferases). ANL enzymes catalyse a two-step reaction¹²; initially activating a carboxylate substrate with ATP (adenosine triphosphate) to form an adenylate intermediate, followed by the formation of a thioester by the NRPS and

¹Department of Biochemistry, University of Otago, Dunedin, New Zealand. ²New Zealand Institute for Plant and Food Research Ltd., Department of Chemistry, University of Otago, Dunedin, New Zealand. Oliver C. Watkins and Miriam L. Sharpe contributed equally to this work. Nigel B. Perry and Kurt L. Krause jointly supervised this work. Correspondence and requests for materials should be addressed to N.B.P. (email: Nigel.Perry@plantandfood.co.nz) or K.L.K. (email: kurt.krause@otago.ac.nz)

acyl-CoA synthetases. In the reaction catalysed by firefly luciferase, the second partial reaction is the oxidative decarboxylation of the activated luciferyl-adenylate, producing a high-energy dioxetanone which decomposes to release carbon dioxide and a photon¹¹.

In *Arachnocampa*, studies by Lee (working with *A. richardsae*), Shimomura (*A. luminosa*) and Viviani *et al.* (*A. flava*) have established that the bioluminescence is also a luciferin-luciferase reaction that requires ATP-Mg²⁺ and oxygen^{13–15}. Silva *et al.*¹⁶ carried out a transcriptional profiling study on *A. luminosa*, sequencing 537 cDNAs that were constructed from light organ expressed sequence tags using the Sanger method (they did not include a comparison between tissues). Our research on *A. luminosa* using transcriptome sequencing and differential analysis of expression from the light organ compared with non-light organ tissue, revealed three candidate glowworm luciferases, all from the same superfamily of enzymes as the firefly luciferase¹⁷.

A recent study of bioluminescence in Australian glowworms reported the identification of the *A. richardsae* luciferase¹⁸. The authors isolated a sequence from a cDNA library made from light organ RNA that encoded an enzyme from the ANL superfamily. The study also reported the presence of firefly D-luciferin in the glow-worm light organs, and further, that the cloned enzyme expressed recombinantly in bacteria was able to produce light using firefly D-luciferin, which is in contrast to previous studies reporting no cross reaction between the glow-worm and firefly bioluminescent systems^{13,15}.

The glowworm and firefly bioluminescent systems have evolutionarily distinct origins: glowworms are true flies, of the order Diptera (family Keroplatidae), whereas the well-studied fireflies are beetles, members of the order Coleoptera (superfamily Elateroidea). Diptera and Coleoptera diverged about 330 million years ago, and do not have any known bioluminescent species intervening^{8,19,20}.

In this study we aimed to isolate and identify luciferin and luciferase from *A. luminosa* light organ tissue using biochemical techniques, including chromatographic fractionation of cell-free lysates, followed by assays and mass spectrometry (MS). We confirm that the New Zealand glowworm luciferase is a member of the ANL superfamily, similar to firefly luciferase, but reveal the *A. luminosa* luciferin to be a novel luciferin, containing tyrosine (Tyr) and xanthurenic acid (XA) moieties, and propose a candidate *A. luminosa* luciferin structure.

Results

Assay development and optimisation. We developed optimised assay methods for the investigation of the *Arachnocampa* bioluminescent system based on the ‘hot extract’, ‘cold extract’ assays of *Arachnocampa* luciferin-luciferase reported by¹⁵. We prepared hot extract containing luciferin without luciferase by heating a cell-free cold extract lysate at 95 °C for 5 min, thereby denaturing all proteins present, but leaving heat stable small molecules intact. Crude luciferase samples depleted in luciferin and related metabolites were prepared from the cell-free cold extract lysate using a centrifugation-based desalting size exclusion column, which separates large protein molecules from small molecules. An outline of our extract preparation is provided in Fig. 1A.

Combining hot extract and crude luciferase samples with ATP-Mg²⁺ produced bioluminescence (Fig. 1B). The luminescence emission spectrum produced by this reaction (Fig. 1C) had the same peak emission wavelength (487 nm) as that reported previously for *A. luminosa*¹⁴. Notably, the hot extract alone produced no light when mixed with ATP-Mg²⁺, showing it to be completely depleted in active luciferase. A small amount of light was produced by the combination of crude luciferase and ATP-Mg²⁺, so these samples were not entirely depleted in luciferin. However, significantly more luminescence (>100 times) was produced with the addition of hot extract to this combination, indicating that the luciferin content was low enough to use the crude luciferase preparations in a luciferin detection assay. As a control in all luciferin assays, the luminescence produced by the sample was compared with that produced by crude luciferase and ATP-Mg²⁺ without sample addition.

Other improvements to the assay involved additives, buffer optimisation, and careful light organ preparation. The stability of lysate was improved by the addition of reducing agents such as dithiothreitol (DTT), extending the time that light was produced by at least 5-fold (Supplementary Fig. S1), and luminescence intensity was increased over 8-fold using MOPS (3-(N-morpholino)propanesulfonic acid), pH 7.5 instead of Tris (Tris(hydroxymethyl) amino methane) pH 8.0 as the assay buffer (Supplementary Fig. S2). Care was taken to include only light organs in the lysate preparation, because the addition of non-luminescent glowworm tissue other than light organ significantly reduced lysate luminescence (Supplementary Fig. S3). Also, lysate was always kept at 4 °C when not stored at –80 °C. Even with these improvements, bioluminescence from freshly produced (or thawed) light organ lysate upon the addition of hot extract and ATP-Mg²⁺ declined over several hours until it was undetectable 4 to 6 hours after lysis (or thawing).

Time course of bioluminescence. We found different time courses of light emission from the reactions of different *A. luminosa* lysates (Fig. 1D). Some preparations produced light with a time course profile similar to those observed previously^{13,15}, where light intensity increased rapidly to a maximum within seconds of adding ATP-Mg²⁺, then decreased gradually over several minutes (‘fast peak’ or FP type). However, other lysates produced profiles in which the luminescent maximum was reached many minutes after ATP-Mg²⁺ was added (‘slow peak’ or SP type). Some lysates produced profiles that were intermediate between FP and SP luminescence. Initially it was unclear what factors during lysate preparation were associated with the different luminescence time course profiles, but below we report data suggesting that these different profiles may relate to the biosynthesis of luciferin from its precursors.

***A. luminosa* luciferase and luciferin do not cross react with firefly luciferase and luciferin.** *A. luminosa* crude luciferase preparations were tested to see if they would produce light following the addition of firefly luciferin and ATP-Mg²⁺. Additionally, firefly luciferase was tested to see if it would produce light when mixed with *A. luminosa* hot extract and ATP-Mg²⁺. When the luciferins and luciferases from glowworm and firefly species were mixed, we detected the same level of light as for the luciferases without luciferins added

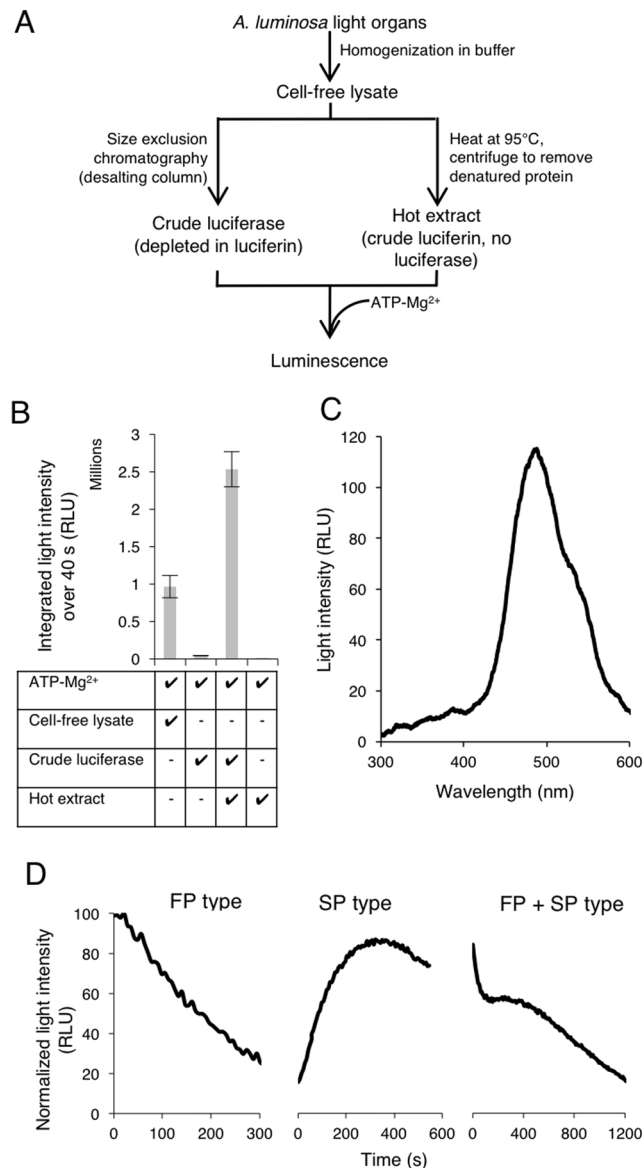


Figure 1. Preparation of components of luciferase and luciferin detection assays. **(A)** Overview of the production of *A. luminosa* lysate, crude luciferase and hot extract. **(B)** Luminescence of lysate, crude luciferase and hot extract for 40 s immediately after the addition of ATP-Mg²⁺. Data points represent the mean of three experiments; error bars indicate standard deviation. **(C)** *In vitro* bioluminescence spectrum of hot extract mixed with crude luciferase and ATP-Mg²⁺, with a peak emission wavelength of 487 nm. **(D)** Variations in the time course of luminescence reactions initiated by the addition of ATP-Mg²⁺ to hot extract and crude luciferase (added 5 s before measurements started). From left: ‘fast peak’ (FP) type, ‘slow peak’ (SP) type, and intermediate FP + SP type. RLU: relative luminescence units.

(Fig. 2). This is in contrast to the intense luminescence obtained when crude *A. luminosa* luciferase and *A. luminosa* hot extract preparations were mixed together, or when firefly luciferase and firefly luciferin were mixed. Therefore, firefly luciferin is not the substrate of the *A. luminosa* bioluminescent reaction. Furthermore, analysis of *A. luminosa* lysates using MS and liquid chromatography-MS (LC-MS) detected no firefly D-luciferin ([M-H]⁻ C₁₁H₇N₂O₃S₂⁻ (278.9895) or decarboxylated daughter ion C₁₀H₇N₂O₂⁻ (234.9996)²¹ (Supplementary Fig. S4 and Table S1).

Partial purification and identification of the glowworm luciferase enzyme. Protein-associated bioluminescent activity was partially purified from light organ lysate using fast protein liquid chromatography (FPLC). We used two columns sequentially during purification: a gel filtration column followed by an ion exchange column. Fractions eluted from both columns were assayed for bioluminescent activity using hot extract as luciferin substrate and analysed using SDS-PAGE.

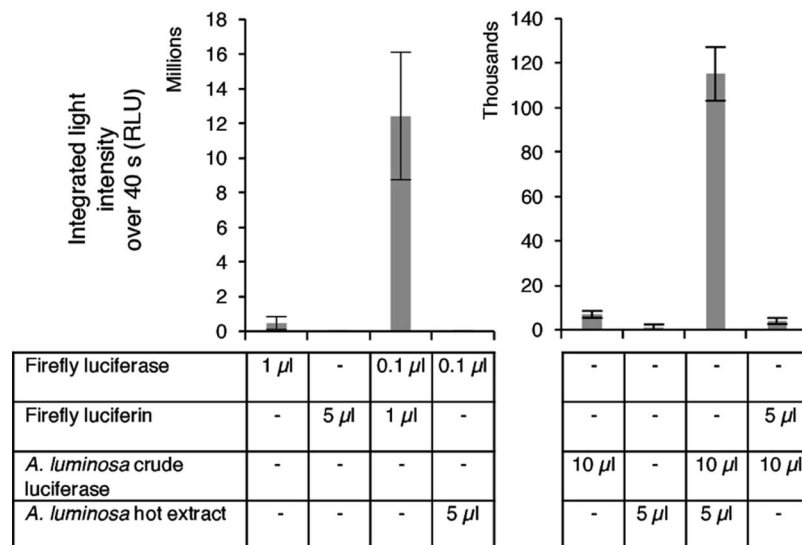


Figure 2. *A. luminosa* luciferase does not produce bioluminescence with firefly luciferin and vice versa. ATP-Mg²⁺ was added to all assays and luminescence measured for 40 s. Data points represent the mean of three experiments; error bars indicate standard deviation. RLU: relative luminescence units.

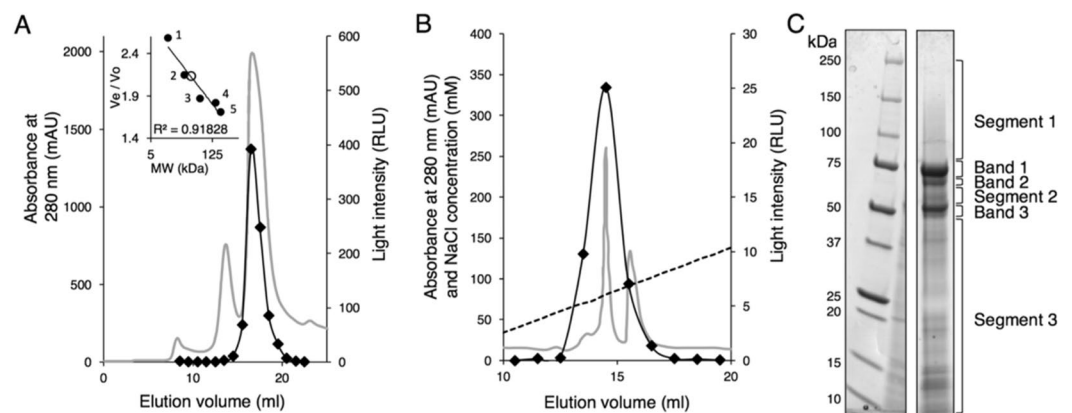


Figure 3. Purification of luciferase bioluminescent activity from *A. luminosa* light organ cell lysate. Absorbance at 280 nm is represented as a grey line and luminescence as a black solid line. Light intensity was the integrated relative light units recorded over the first 40 seconds of the activity assay with *A. luminosa* hot extract and ATP-Mg²⁺. (A) Gel filtration chromatography. See also: SDS-PAGE Gel G1 available in the supplement, Inset: calibration curve for the determination of size associated with glowworm luciferase activity, constructed using five standard proteins of known molecular weights (closed circles; 1 cytochrome C, 2 carbonic anhydrase, 3 bovine albumin, 4 alcohol dehydrogenase, 5 β -amylase). The V_e/V_o value of the fraction containing the highest luminescent activity is represented by an open circle. (B) Ion exchange chromatography on active fractions pooled from gel filtration chromatography, (gradient of NaCl represented as a dashed black line). (C) SDS-PAGE of the fraction eluted from the ion exchange column with the highest bioluminescent activity (right lane), indicating the different areas of the gel used for MS analysis. Left lane: molecular weight markers. Lanes are cropped from different parts of the same gel. See also: full SDS-PAGE gel G2 available in the supplementary information.

Five of the fractions eluting from the gel filtration column contained significant bioluminescent activity (Fig. 3A), with the highest level of activity eluting between 16 and 17 ml, correlating to a size of about 41 kDa. This peak was quite broad (size range of 11 to 63 kDa), and therefore the estimate of the molecular weight of the corresponding enzyme guided by calibration standards is very approximate. The five most active fractions were pooled and injected onto the ion exchange column.

Proteins binding to the ion exchange column were eluted with a shallow gradient, from 0 to 200 mM NaCl (Fig. 3B). Four fractions eluting at about 60 to 90 mM NaCl had detectable bioluminescent activity, and the fraction with the highest activity eluted early in the gradient. SDS-PAGE of the active fractions eluted from ion exchange chromatography showed that they contained a mixture of different proteins, including three prominent

bands and many faint bands. The gel of the most active fraction is pictured in Fig. 3C. Further purification was unable to be carried out because of loss of sample during chromatography steps, and decline of the bioluminescent signal.

To identify the proteins in the most active fraction using MS, we divided the gel lane in Fig. 3C into six distinct areas: three prominent bands, and three segments that encompassed the remaining faint bands in the lane. Bands and segments of interest were excised from the gel and digested using trypsin. The masses of resultant peptides were measured using MS (MALDI TOF/TOF) and then searched against a glowworm transcriptome assembly (Genbank accession GDQV00000000)¹⁷ that was translated into all possible reading frames. Details of the MS analysis, including protein sequence and peptide coverage are provided in Supplementary Tables S2 and S3. Annotated sequence homologs of the glowworm proteins were found using BLASTX²² and the Genbank protein sequence database (NCBI; <http://www.ncbi.nlm.nih.gov>).

Two types of protein dominate the list of identified proteins in the active fraction: proteins annotated as larval serum protein 2 (encoded by four transcripts) and proteins annotated as luciferin 4-monooxygenase (encoded by two transcripts). Annotation as luciferin 4-monooxygenase originates from sequence similarity to firefly luciferases. Alignments of the four larval serum protein 2 sequences (Supplementary Fig. S5) show that they are four similar but separate proteins; only one of these has a transcript that provides the sequence for the full-length protein. Only one of the two luciferin 4-monooxygenase transcripts encodes a full-length protein (62762); the other (55693) appears to be a fragment. Where the two sequences can be aligned, they are nearly identical except for two amino acids (Leu29 and Ile51 in 55693, Ile29 and Val51 in 62762; Supplementary Fig. S6). It is likely that they are separate isoforms, since both amino acid variations are captured by the peptide coverage of these proteins on the gel (Supplementary Table S3). Two aldehyde dehydrogenases also recur on the list; these also appear to be isoforms, with 76% sequence identity.

Alignment of the *A. luminosa* luciferase with the *A. richardsae* luciferase sequence isolated by Trowell *et al.*¹⁸ shows that the sequences are very similar, with 86% sequence identity (Supplementary Fig. S7). In contrast, the *A. luminosa* luciferase encoded by transcript 62762 has 30% identity with luciferases from the fireflies *Photinus pyralis* and *Luciola cruciata*. The ATP-binding motifs conserved throughout the ANL superfamily¹² are present in all four sequences, and the lysine residue that plays a key role in the firefly luciferase adenylation half reaction²³ is conserved in the *Arachnocampa* sequences. However, the luciferin-binding residues from the beetle luciferase^{24,25} are not as well conserved in the *Arachnocampa* luciferase, and the lysine that plays a key role in the oxidation (light-producing) half reaction in the beetle²³ is replaced by a methionine in *Arachnocampa*. We are now working on producing the *A. luminosa* luciferase in a recombinant form, to be assayed for bioluminescent activity using the native glowworm luciferin substrate as described in this report. Initial attempts at expression in multiple standard prokaryotic strains, as well as in insect and human cells, have so far not yielded soluble protein, but continued efforts in this regard are ongoing.

Xanthurenic acid (XA) and tyrosine (Tyr) contribute to glowworm bioluminescence. Initial attempts at isolating the glowworm luciferin were carried out on *A. luminosa* hot extract preparations using reverse phase (C18) bench column chromatography followed by reverse phase HPLC (high pressure liquid chromatography). Separations from three different hot extract preparations are reported here, carried out using slightly different elution strategies (Supplementary Table S4) and with hot extracts that exhibited different luminescent time course profiles (Fig. 4A). Fractions eluted from the bench column were concentrated and then tested for the presence of luciferin. Fractions that produced luminescence were further separated by HPLC. Eluted HPLC fractions were also concentrated and tested for bioluminescent activity; any of these fractions demonstrating luminescence were then characterised by LC-MS (liquid chromatography-mass spectrometry). Separation, concentration and handling steps were carried out under N₂ to minimise oxidation. Time course profiles for the starting material and luminescent fractions isolated for each separation are provided in Fig. 4A, and chromatograms for each semi-preparative HPLC fractionation are provided in Fig. 4B. LC-MS analysis details can be seen in Supplementary Fig. S8.

These three separations each led to the isolation of different components that produced light when assayed. Separation 1 yielded a fraction that eluted from the C18 bench column with 20% acetonitrile that exhibited an FP type luminescence reaction time course, and was selected for further fractionation by HPLC. The active fraction eluting during HPLC also produced FP luminescence, and corresponded to a peak that eluted immediately after tryptophan. The compound in the active peak could not be identified due to a lack of material. Further attempts to purify this compound by reverse phase HPLC produced no bioluminescent fractions. From separation 2, a luminescent fraction (SP-type) from the bench column was obtained following elution with water. This fraction was injected onto the HPLC column to give an early-eluting fraction that also produced SP luminescence. LC-MS analysis of this fraction showed that it mainly contained Tyr, with small amounts of 3-OH kynurenine and leucine. During separation 3, a fraction was eluted from the bench column with 20% acetonitrile that produced SP type luminescence. Following HPLC purification, a late-eluting SP luminescent fraction was obtained. LC-MS analysis showed that this fraction mainly contained XA, with trace amounts of riboflavin.

Since Tyr and XA were isolated in these separations, and appeared to be associated with luminescence activity, we decided to assay the combination of pure (commercial) L-Tyr and the active HPLC fraction containing XA from separation 3 with the crude luciferase. This combination produced 8-fold greater luminescence than XA-containing fractions tested without additional Tyr (Fig. 5A). Next, we assayed pure L-Tyr plus pure (commercial) XA for luciferin activity using our assay (Fig. 5B). Both L-Tyr and XA separately produced luminescence above that of the crude luciferase alone, but when they were assayed together, the light produced was about eighteen-fold more intense. These results suggest that Tyr and XA are both required for glowworm bioluminescence. Furthermore, LC-MS analysis of crude luciferase preparations indicated that both compounds are present

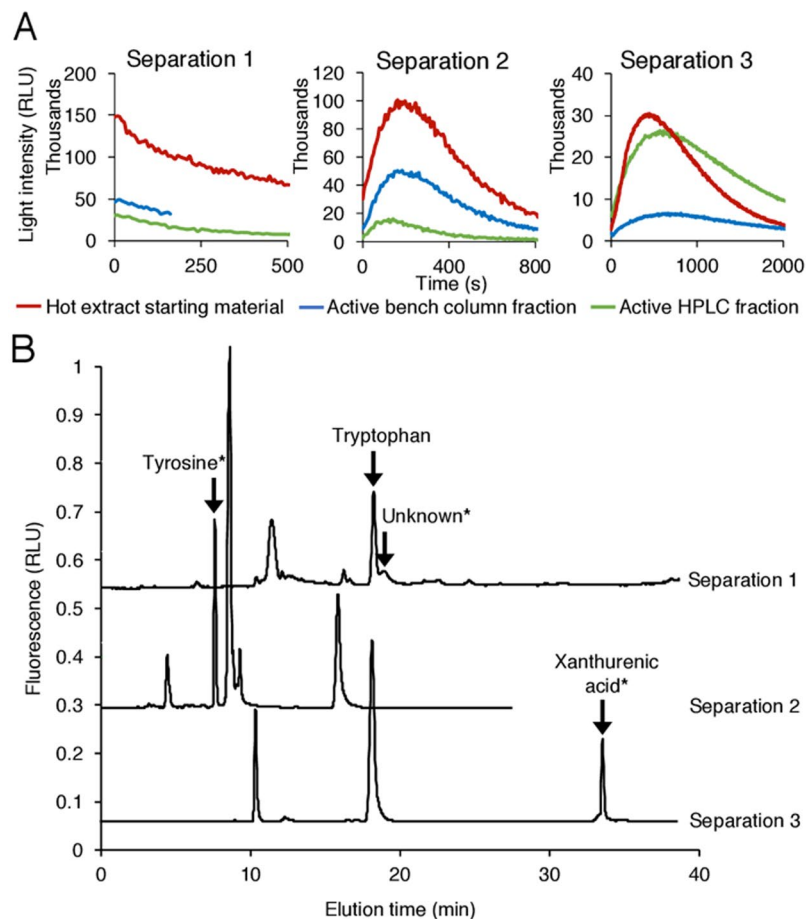


Figure 4. Isolation of Tyr and XA from *A. luminosa* hot extract. (A) Time courses of luminescence reactions with luciferin-containing fractions from chromatographic separations of *A. luminosa* hot extracts. In each assay, material to be assayed was added to crude luciferase, and ATP-Mg²⁺ was injected 5 s before start of measurement. (B) Chromatogram showing the fractionation of three preparations from *A. luminosa* hot extract by semi-preparative HPLC, monitored by fluorescence (excitation 280 nm, emission 415 nm). Peaks correlating with fractions that luminesced when added to crude luciferase-ATP-Mg²⁺ are indicated by *. Back injection of these fractions produced the same peak. Separation 1 produced light with FP profile; separations 2 and 3 produced light with SP profile. RLU: relative luminescence units.

(Supplementary Fig. S4 and Table S1), which could explain why luminescence is also observed when Tyr or XA were tested alone.

The bioluminescence emission spectrum produced by the reaction of XA and Tyr with luciferin-depleted luciferase and ATP-Mg²⁺ has the same peak wavelength as that produced by the reaction of hot extract with luciferin-depleted luciferase and ATP-Mg²⁺ (Fig. 5C), consistent with this being the same bioluminescent reaction.

Intriguingly, all of these assays produced light which took many minutes to reach peak intensity (SP type time course). One explanation for this could be that Tyr and XA are precursors of the actual glowworm luciferin, and the delay in peak light intensity is due to the extra time required to conjugate the compounds together into a luciferin by an unknown luciferin synthetase enzyme present in the crude luciferase preparation, before the bioluminescence reaction can occur. If this theory is correct, the addition of a ‘fully formed’ *Arachnocampa* luciferin to the crude luciferase would produce rapidly peaking light intensity (FP type). Therefore, the compound isolated during separation 1 that produced a FP type luminescent time profile (but was unable to be identified by LC-MS) may be the glowworm luciferin. A schematic of the putative two-step process that Tyr and XA require to produce SP type bioluminescence is provided in Fig. 5D. A similar two-stage bioluminescence scheme was proposed by Airth and Foerster for bioluminescent fungi²⁶, and recently confirmed by Purto *et al.*, who were also able to identify the luciferin precursor as hispidin and the luciferin as 3-hydroxy hispidin²⁷.

Isotope labeling identifies a glowworm luciferin-related compound (LRC) comprising tyrosine and xanthurenic acid. Since the actual *A. luminosa* luciferin proved difficult to isolate and identify from hot extract, and we now suspected that Tyr and XA were precursors to the *A. luminosa* luciferin, we decided to try extracting the luciferin or luciferin-related compounds (such as a luciferin precursor or the end product of the bioluminescence reaction) from crude luciferase preparation supplemented with Tyr and XA. We have shown

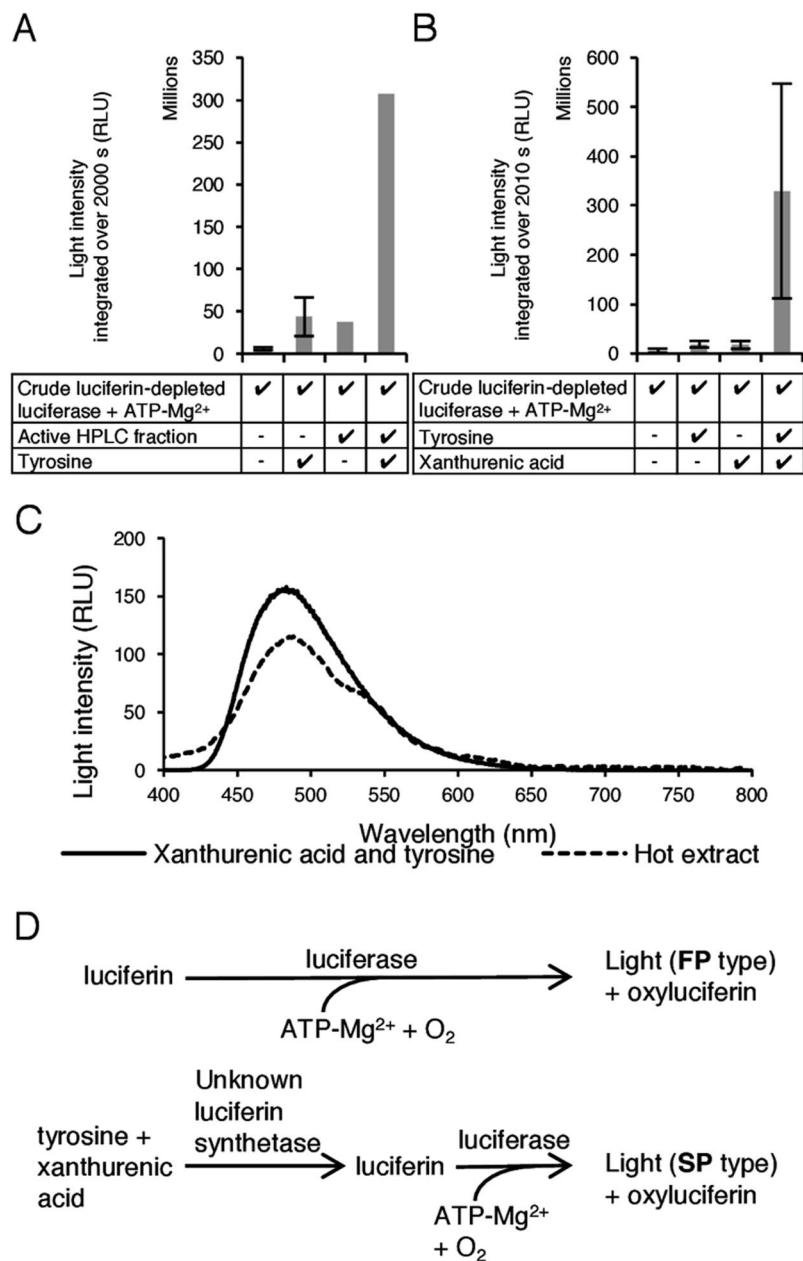


Figure 5. XA and Tyr produce light when added to crude luciferin-depleted *A. luminosa* luciferase. ATP-Mg²⁺ was injected 5 s prior to measurement beginning for all assays. RLU: relative luminescence units. **(A)** Active HPLC fraction from separation 3 mixed with Tyr produces more light than either alone when added to crude luciferin-depleted luciferase. **(B)** XA mixed with Tyr produce more light than either alone when added to crude luciferin-depleted luciferase. Where error bars are present, data points represent the mean of two **(A)** or three **(B)** experiments; error bars indicate standard deviation. No error bars indicate that replicate assays were not carried out due to lack of material. **(C)** Luminescence emission spectra of either hot extract or XA and Tyr added to crude luciferin-depleted luciferase and ATP-Mg²⁺. **(D)** A proposed model of glowworm luciferin metabolism. The luminescence time course profile produced (FP type or SP type) is dependent on the starting material provided.

above that such preparations appear to contain the necessary constituents, likely to be enzyme(s), required for luciferin synthesis.

To help identify such compounds, the Tyr used consisted of a mixture of isotope-labeled Tyr (¹³C₉H₁₁¹⁵NO₃) and unlabeled Tyr (¹²C₉H₁₁¹⁴NO₃) in a 1:1 ratio. Any Tyr metabolites biosynthesised from this mixture would then exist as a pair of compounds, with identical LC-MS elution times but different masses, one made from labeled Tyr and one made from unlabeled Tyr. Such pairs can be searched out automatically in LC-MS data (SIEVE 'perfect pair' program, Thermo Fisher Scientific). Any such Tyr metabolite found only in reactions which luminescence is likely to be the luciferin or a luciferin-related compound.

We therefore prepared four reaction mixtures: (1) crude luciferase, labeled and unlabeled Tyr, XA and ATP-Mg²⁺; (2) crude luciferase, labeled and unlabeled Tyr, and XA but no ATP-Mg²⁺; (3) crude luciferase and ATP-Mg²⁺; and (4) crude luciferase alone. The reactions were incubated for 5.5 hours and light production recorded before being prepared for LC-MS analysis using HPLC. Reaction mixture 1 contained one main ‘perfect pair’: co-eluting compounds with MH⁺ 369.1080 Da and 379.1350 Da, a mass difference indicating incorporation of C₉N from Tyr (Supplementary Table S5). This pair of compounds was only produced in the reaction mixture that produced luminescence, reaction mixture 1 (Fig. 6A).

In other experiments, either only labeled or only unlabeled Tyr was mixed with crude luciferase, XA and ATP-Mg²⁺ (Supplementary Fig S9). Both mixtures produced luminescence, but the unlabeled Tyr gave only the compound with MH⁺ 369.1080. Labeled Tyr gave the compound with MH⁺ 379.1350 Da, plus a small amount of MH⁺ 369.1080 Da compound, probably from the unlabeled Tyr in the crude luciferase. This is referred to as glowworm luciferin-related compound (LRC).

In order to produce enough LRC for further characterisation, we scaled up the preparation with unlabeled Tyr, and isolated about 10 µg of material. Initially this batch of LRC was 99% pure (by HPLC-UV and LC-MS), but it was unstable, giving oxidised products on storage. Portions of purified LRC were added to luciferin-depleted crude luciferase and ATP-Mg²⁺ to assay for bioluminescent activity. LRC produced intense FP type luminescence (Fig. 6B), which appears to be concentration-dependent (Fig. 6C). This evidence suggests that LRC is either the *A. luminosa* luciferin or a luciferin precursor.

Characterisation of luciferin-related compound (LRC). MS and MS fragmentation (MS²) analyses were carried out with unlabeled and stable isotope labeled LRC using LC-MS (Supplementary Table S5). High-resolution MS revealed that unlabeled and labeled LRC had molecular formulae of C₁₉H₁₆N₂O₆ and C₁₀¹³C₉H₁₆N¹⁵NO₆, respectively, consistent with compounds containing either unlabeled or labeled Tyr, plus XA, minus H₂O. This is most readily explained by LRC having an amide or an ester linkage between Tyr and XA moieties. The similarity of the UV-Vis spectra of LRC and XA (Fig. 6D) show that the XA chromophore was retained in LRC and that the ring was not further oxidised or reduced.

MS² of LRC gave fragments also observed in the MS² of XA and Tyr, as well as four further fragments that incorporated both XA and tyrosine moieties (Supplementary Table S5). The lack of an intact [Tyr + H]⁺ ion indicates that these two moieties are not linked by a weak phenol ester (depside bond)²⁸. Unlabeled LRC showed an initial loss of CO₂H₂, while labeled LRC [M + H]⁺ showed a corresponding initial loss of ¹³CO₂H₂ (Supplementary Table S5). This indicated that ^{Tyr}carboxylic acid was lost from LRC as formic acid without the cleavage of the XA residue from the Tyr residue, thus showing that Tyr was not attached to XA via the ^{Tyr}carboxylic acid. The presence of the [M-CO₂H₂-H₂O]⁺ ion is best explained by loss of water from the [M-CO₂H₂]⁺ ion; [XA + H]⁺ also fragments via the loss of water (m/z 206–188, Supplementary Table S5). The C₁₀^{Tyr}CH₉N^{Tyr}NO₃ ion meanwhile is likely formed via the cleavage of most of the tyrosyl residue from the XA leaving only a ^{Tyr}C and ^{Tyr}N attached, while the C₁₀H₈NO₄⁺ ion likely represents the [XA + H]⁺ ion or similar.

The LRC NMR (Nuclear magnetic resonance) spectrum showed signals similar to those of the 5, 6 and 7 protons of XA and the 2 + 6, 3 + 5 and β protons of Tyr, suggesting that these regions are similar in LRC. The XA H3 signal (singlet at 7.8 ppm) had no near equivalent in the spectrum of LRC, but there was a one-proton singlet at 9.1 ppm in the LRC NMR spectrum, absent from Tyr and XA spectra. ¹³C and 2D NMR spectra could not be collected on this sample due to the small amount (10 µg) available; the proton NMR took over a day to collect.

The only possible amide link, not involving a ^{Tyr}carboxylic acid linkage, is that of ^{XA}carboxylic acid to ^{Tyr}N amide (Fig. 7A). However, such a structure does not explain the high chemical shift singlet at 9.1 ppm (Supplementary Table S6 and Fig. S10). The few known amides involving a ^{XA}carboxylic acid show ^{XA}H3 around 7.3 ppm²⁹, effectively ruling out the structure Fig. 7A. The α proton in LRC is at 5.4 ppm, compared to 4.5 ppm in Tyr (Supplementary Table S6), suggesting linkage of XA to Tyr via the ^{Tyr}N group (given that the ^{Tyr}carboxylic acid is intact, see above, and N-phenyl Tyr has its α proton deshielded compared to Tyr)³⁰. Since we already dismissed the XA-Tyr amide, Fig. 7A (see above), we considered compounds with ^{Tyr}N attached directly to the XA quinoline ring at C3 or C4, with loss of the 4OH, i.e. the structure in Fig. 7B, or the C4-N isomer. However, direct infusion MS analysis of LRC dissolved in 1:1 D₂O:CD₃CN with 0.1% formic acid showed [MD]⁺ ions up to 372 Da, compared to [MH]⁺ of 369 Da for unlabelled LRC (Supplementary Table S7). LRC therefore contains two protons that can exchange with deuterium under acidic conditions (excluding the ionising proton required for [M + H]⁺). In contrast (excluding ionising protons), XA contains one acid exchangeable proton (COOH) while Tyr has three (NH₂ and COOH). This shows that the XA and Tyr phenols are not exchangeable under the conditions, and that structure Fig. 7B should contain three acid-exchangeable protons (excluding the ionising proton). Since LRC contains two acid exchangeable protons (excluding the ionising proton), structure Fig. 7B was dismissed.

In light of this evidence, we propose that the 9.1 ppm singlet in the ¹H NMR spectrum of LRC (Supplementary Table S6) is caused by an aldehyde proton, suggesting that the candidate luciferin has the structure in Fig. 7C or in 7D. Model compounds show singlet aldehyde proton signals at about 10 ppm³¹ or 9.6 ppm^{32,33}. Both these values are a little higher but they do indicate that an aldehydic proton could be responsible for the 9.1 ppm peak. The Fig. 7D structure may appear prohibited due to the lack of coupling between the aldehyde proton and the vicinal Tyr α proton. However, some similar aldehyde protons exhibit singlet ¹H NMR peaks at 9.6 ppm^{32,33}.

The proposed structure in Fig. 7C is preferred because it contains a free ^{Tyr}carboxylic acid, which makes the ^{Tyr}CO₂H₂ MS loss easy to rationalise. This structure in Fig. 7C is still problematic, because it is difficult to determine the origin of the C₁₀H₈NO₄⁺ ion [XA + H]⁺. However, some aldehydes with a vicinal group that can hydrogen bridge to the aldehyde (e.g. formylchromones and 3-formylcoumarins) do form fragment ions during electrospray ionisation MS² in which the aldehyde is converted into a carboxylic acid³⁴. The carboxylic acid group is thought to be produced from the aldehyde via loss of H₂ followed by water adduction within the collision cell. Such a process occurring to a suitable LRC fragment ion could possibly result in the [XA + H]⁺-like ion.

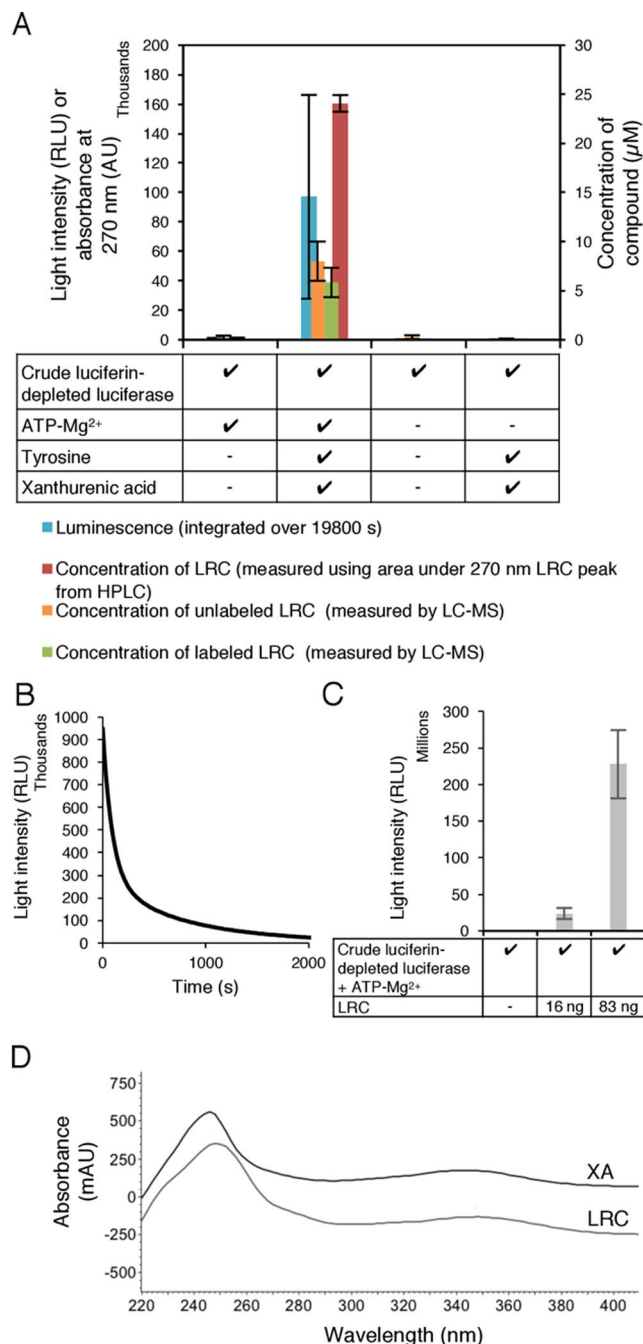


Figure 6. Isolation and identification of LRC. (A) Bioluminescence and the compound LRC are only produced when Tyr and XA are added to a reaction mix of luciferin-depleted crude luciferase and ATP-Mg²⁺. LRC concentration was measured using LC-MS and HPLC-UV. Data points represent the mean of three experiments; error bars indicate standard deviation. (B) Isolated LRC produces light with an FP time course when added to luciferin-depleted crude luciferase. ATP-Mg²⁺ was injected 5 s prior to measurement. (C) Isolated LRC produces light in a concentration-dependent manner. An increase in LRC added to the assay of about 5-fold resulted in an increase of light of nearly 10-fold. Data points represent the mean of three experiments; error bars indicate standard deviation. (D) Comparison of UV-Vis spectra of XA and LRC. Absorbance maxima: 248 and 348 nm for LRC, 244 and 348 nm for XA. RLU: relative luminescence units.

We therefore propose the LRC structure shown in Fig. 7C. To confirm this hypothesis, synthesis of the proposed LRC structure is being pursued, to be followed by spectral comparisons and bioluminescence assays.

Discussion

The purification and characterisation work reported above has shown directly that the *A. luminosa* glowworm luciferase is the 59 kDa protein encoded by transcript 62762, which we had identified previously in differential

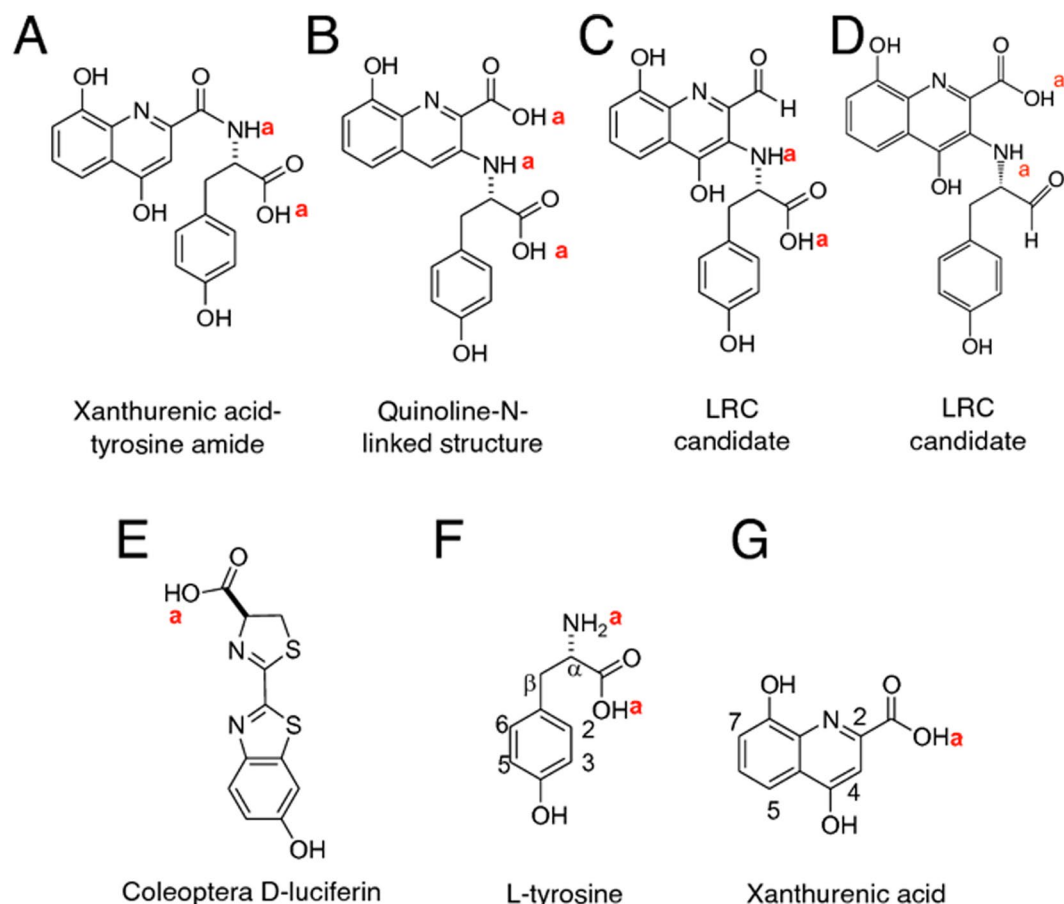


Figure 7. Small molecules involved in the bioluminescence of the glowworm and the firefly. (A) Possible XA-Tyr amide. (B) Possible quinoline-N-linked structure. (C), (D) Candidate structures of *A. luminosa* LRC. (E) Coleoptera D-luciferin, found in fireflies, click beetles and railroad worms. (F) L-Tyr. (G) XA. “a” indicates protons exchangeable with deuterium under acid conditions.

expression analyses¹⁷. Since the glowworm bioluminescence reaction requires ATP, it makes sense for the glowworm luciferase to be encoded by a member of the ANL superfamily of adenylating enzymes, which all require ATP for catalysis. The larval serum proteins and the aldehyde dehydrogenases found in the active fraction are unlikely to catalyse the glowworm reaction because such proteins are not known to require ATP. Because of the difficulty of collecting *A. luminosa* larvae, we are working on producing this protein recombinantly. It will then be assayed for bioluminescent activity using the native glowworm luciferin substrate described here.

The combination of XA and Tyr proposed here for glowworm luciferin formation is unprecedented in the luciferins of other bioluminescent systems^{10,11}. The other known insect luciferin, from the bioluminescent Coleoptera, i.e. fireflies, click beetles and railroad worms, Fig. 7E, is biosynthesised from 1,4-hydroquinone and two L-cysteines³⁵. XA and Tyr are reasonable precursors for the glowworm luciferin for several reasons. Tyr (Fig. 7F) is a component of both coelenterazine and of *Fridericia heliota* luciferin^{36,37}, and is also an important insect metabolite^{38–43} that can undergo a number of radical, redox and charge transfer processes^{44–47}. XA (Fig. 7G) is a well-studied insect metabolite^{48–50} and metabolic precursor^{51–53}. It is an antioxidant free radical scavenger, used by insects to prevent damage by reactive oxygen species^{54,55}. Furthermore, like other quinolines, XA is chemiluminescent when oxidised by free radicals^{54,56}. There is evidence of XA synthesis in *A. luminosa* light organs: in our study 3-OH kynurenine (converted into XA in insects by kynurenine transaminase)³⁴ was found in the *A. luminosa* light organ (Supplementary Figs S4, S8, Table S1); and we previously observed that kynurenine transaminase is more highly expressed in the *A. luminosa* light organ compared with the rest of the larval body¹⁷ (although not to a statistically significant extent).

To confirm the proposed LRC structure shown in Fig. 7C, synthesis is being pursued, to be followed by spectral comparisons and bioluminescence assays.

The LRC molecule isolated in these experiments is likely to be the *A. luminosa* luciferin, or rapidly converted to it, because LRC produces high intensity, fast peaking blue-green luminescence when mixed with crude luciferin-depleted luciferase and ATP-Mg²⁺. However it is not the same luminescent compound that we isolated from the glowworm (see Fig. 4B) as these molecules showed different retention times. We cannot rule out the possibility that LRC is a precursor to the *A. luminosa* luciferin, converted by an unknown enzyme from the light organ lysate into the active substrate.

Our findings contrast with those of Trowell *et al.*¹⁸ who recently reported that recombinant *A. richardsae* luciferase uses firefly D-luciferin to produce blue light. In *A. luminosa*, we could not detect a cross reaction between firefly and *A. luminosa* molecules (Fig. 2). Further, we did not detect any firefly D-luciferin in *A. luminosa* lysate (Supplementary Fig. S4 and Table S1). Our results are in agreement with Lee¹³, who reported that firefly luciferin did not stimulate light production when added to reconstituted lysate made from freeze-dried *A. richardsae* light organs. We note that the bioluminescent emission spectrum reported by Trowell *et al.* from recombinant luciferase and firefly D-luciferin¹⁸ is very much broader than that reported by Lee when adding hot extract made from *A. richardsae* light organ lysate to crude luciferase made from *A. richardsae* light organ lysate¹³. Therefore, either *A. luminosa* and *A. richardsae* use different luciferins, or the luminescence reported by Trowell *et al.* stems from another source. *A. richardsae* luciferase may be able to produce light with firefly luciferin, since even acyl-CoA enzymes from non-bioluminescent insects have been found to produce some light with firefly luciferin or an analog^{57,58}. It would be useful to compare the bioluminescence Trowell *et al.* reported with that produced when adding hot extract made from *A. richardsae* light organ lysate to firefly luciferase and to crude *A. richardsae* luciferase (as in Fig. 1A).

It was previously assumed that evolutionarily unrelated bioluminescent systems use luciferase enzymes with very different structures and mechanisms, evolved from different enzyme families^{1,7,59}. There are also a variety of luciferin substrates^{1,11}, although species with evolutionarily unrelated bioluminescent systems sometimes share a luciferin, especially when a luciferin is readily available from the food chain. For example, the Japanese firefly squid, some fish, Hydrozoa, and Anthozoa all use coelenterazine or coelenterazine derivatives as a luciferin substrate^{60,61}, which they do not synthesise, instead feeding upon coelenterazine-producing marine crustaceans such as the shrimp *Systellaspis debilis*⁶² and the copepod *Metridia pacifica*³⁷. In the *A. luminosa* bioluminescent system, we see the trend reversed: it has a novel luciferin not used by any other characterised system, and a luciferase belonging to an enzyme family seen already in the bioluminescent beetles (including fireflies). In fact, there is a growing group of bioluminescent systems that are evolutionarily distinct and are known to require ATP to produce light, but use entirely different luciferins. Three of these systems use enzymes related to acyl-CoA synthetases to carry out bioluminescence: the bioluminescent beetles/firefly⁶³, the Japanese firefly squid (*Watasenia scintillans*)^{64–66}, and now the *Arachnocampa* glowworm. The Siberian earthworm (*Fridericia heliota*) also uses ATP and an entirely different luciferin to produce light^{67,68}; we await further research to see if the *F. heliota* system also uses an enzyme related to acyl-CoA synthetases. Recently, another possible example of convergent evolution of similar enzymes into luciferases has been identified. A luminous echinoderm, the ophiuroid *Amphiura filiformis*, appears to have a luciferase homologous to that of the sea pansy *Renilla* (Cnidaria)⁶⁹. It is therefore evident that these bioluminescent systems are an example of evolved functional convergence of enzymes at the molecular level. It appears that the ability of the acyl-CoA synthetase enzymes to adenylate their substrate molecules provides them with the capacity to evolve bioluminescent activity using a range of substrates.

In conclusion, this investigation has revealed key components of the bioluminescent system of the New Zealand glowworm. The purified luciferase enzyme is from the same adenylating enzyme superfamily as firefly luciferase. Two *A. luminosa* light organ metabolites, Tyr and XA, form a luciferin candidate compound which produces light when added to a crude *Arachnocampa* luciferase. Neither the proposed luciferin candidate structure nor similar structures have been reported from any other bioluminescent organism. This novel glowworm bioluminescent system could be suitable for the development of new research tools and analytical methods.

Methods

Sample collection. *A. luminosa* is an endemic species of Aotearoa/New Zealand, therefore Māori were consulted about this research through Te Komiti Rakahau ki Kāi Tahu (the Ngāi Tahu Research Consultation Committee). Glowworms were collected at night from locations including Dunedin, Te Anau, Palmerston North and Auckland (voucher specimens of each collection have been stored in 75% ethanol, details are available from the authors). Live glowworms were snap frozen on foil placed on dry ice, and stored at -80°C until required. Using a razor, light organs were removed from the glowworm bodies while still frozen, taking care to excise only white light organ tissue, and not the darker non-light organ tissue.

Chemicals. Stable isotope-labeled L-Tyr ($^{13}\text{C}_9\text{H}_{11}^{15}\text{NO}_3$; 99% pure) was purchased from Cambridge Isotope Laboratories. Unlabeled L-Tyr, XA and other amino acids were purchased from Sigma Aldrich and used without further purification. Labeled and unlabeled Tyr and XA were dissolved in NaHCO_3 (1.2 mg/ml) before use. Water (Type 1) was produced using a Milli-Q ultrapure system. LiChrosolv LC grade solvents were purchased from Merck and were used for all chromatography steps.

Production of lysate, hot extract and crude luciferin-depleted luciferase. *Lysate*: frozen light organs (usually 20–40 or 0.023–0.107 g) were homogenised in a B type glass douncer (ThermoFisher Scientific) with 1–4 ml of 100 mM MOPS pH 7.4, 2 mM DTT (unless mentioned otherwise) on ice. The lysed solution was centrifuged at 15800 rcf for 5 min at 4°C to remove cell debris and the supernatant retained. *Hot extract*: lysate was heated at 95°C for 5 min, then centrifuged at 15800 rcf for 5 min at 4°C to remove denatured protein and the supernatant retained. *Crude luciferin-depleted luciferase*: a ZebaTM spin 2 ml desalting column (ThermoFisher Scientific) was equilibrated by centrifugation using 100 mM MOPS pH 7.4, 2 mM DTT. Lysate as prepared above was loaded onto the spin column and the column centrifuged for 1 min, eluting crude luciferin-depleted luciferase. Lysate, hot extract and crude luciferin-depleted luciferase preparations were either used immediately or stored at -80°C and thawed on ice before use.

Bioluminescence assays. Assays were carried out in white, flat bottomed 96 well plates (Nunc®, ThermoFisher Scientific). Typically, assays included 50–60 μl of assay buffer (100 mM MOPS pH 7.4, 2 mM DTT, unless mentioned otherwise), 5 μl of hot extract and 5 μl of crude luciferin-depleted luciferase preparation.

ATP-Mg²⁺ solution (40 mM ATP, 80 mM MgSO₄, 0.1 M MOPS pH 7.4, 2 mM DTT) was then added simultaneously to all wells being assayed (30 µl per well) using a multi-channel pipette, assay reactions mixed, and monitoring of light intensity of each well started immediately afterward. Light intensity was measured using a microplate reader in luminescence mode; either the POLARstar (BMG Labtech) (gain 4095, measurement time for each well within a cycle was 0.58 s with an overall cycle time of 7 s for the measurement of 12 wells) or the CLARIOstar (BMG Labtech) (gain 3500, focal height 11 mm, rapid speed function, measurement time for each well within a cycle was 0.42 s with an overall cycle time of 5 s for the measurement of 12 wells). Luminescence of samples tested non-concurrently or with different crude luciferin-depleted luciferase preparations could not be compared directly, because crude luciferase activity was highly variable between different preparations, and because the activity of any preparation reduced over time. Assay mixtures used to assess the effect of bioluminescence included either 50 µl L-Tyr solution (0.94 µg/ml) and/or 50 µl XA solution (58 µg/ml).

Assay results were analysed using Microsoft Excel. Light intensity in relative luminescence units (RLU) were integrated between each time point using the trapezoid method [(time A – time B) * (height A + height B)/2] and the results summed over the time of interest. A sample was considered to contain luciferin substrate (or luciferase enzyme) if the integrated light intensity over a set time for the sample mixed with crude luciferin-depleted luciferase (or hot extract) and ATP-Mg²⁺ was at least three times the integrated light intensity for a crude luciferin-depleted luciferase and ATP-Mg²⁺ only control, tested at the same time under the same conditions. Statistical analyses: error bars show one standard deviation above and below the mean (average of three replicate reaction mixtures unless stated otherwise). Statistical significance was tested using a student T test (two tailed, two sample unequal variance).

Bioluminescence emission spectra were measured using a Cary Eclipse fluorescence spectrophotometer (Varian) in chemiluminescence mode from a reaction mixture in a 100 µl or 1 ml cuvette (settings: gate time 200 ms, scan mode, emission 400–600 nm, slit 20 nm, data interval 0.1500 nm, emission filter open, PMT voltage high, filter size 101, temperature 10 °C).

Fast protein liquid chromatography (FPLC). A Superose 6 10/300 GL column (approximately 24 ml; from Pharmacia, now GE Healthcare) was used for gel filtration, and a 1 ml Mono Q 5/50 GL column (GE Healthcare) was used for ion exchange on an Äkta Purifier FPLC system (GE Healthcare). Prior to purification, columns were equilibrated with 4–6 column volumes of buffer A (30 mM MOPS pH 7.5, 10% glycerol, 1 mM EDTA, 2 mM DTT). Flow rates were 0.4 mL/min for gel filtration and 0.5 ml/min for ion exchange. 1 mL fractions were collected and 5 µl of each fraction assayed for luciferase activity using 5 µl of a hot extract preparation as substrate. Gel filtration was carried out on 2 ml of glowworm light organ lysate, part of a 4 ml batch previously prepared using 0.753 g of glowworm light organ tissue from 186 glowworms, and stored at –80 °C. Before chromatography, the 2 ml lysate aliquot was thawed on ice, filtered through a 0.2 µm filter and assayed to check for bioluminescent activity before being loaded onto the gel-filtration column. Fractions with bioluminescent activity eluting from the gel filtration column were pooled and injected onto the ion-exchange column, which was then washed with 5 column volumes of buffer A, and programmed to carry out elution gradient steps of 0–20% buffer B (30 mM MOPS pH 7.5, 10% glycerol, 1 mM EDTA, 2 mM DTT, 1 M NaCl) over 20 ml, and 30–100% buffer B over 0 ml. Calibration standards (Sigma) were loaded onto the gel filtration column under the same conditions as for the glowworm lysate.

Protein identification and functional annotation. *SDS-PAGE analysis.* 200 µl of each fraction eluted from FPLC was precipitated by adding 800 µl of acetone chilled to –20 °C, then centrifuged at 16,100 rcf for 10 min, the supernatant discarded and pellet left at 37 °C for 20 min to dry. The pellet was resuspended in 10 µl of loading buffer containing β-mercaptoethanol before being loaded onto the gel. Samples were run on Mini-PROTEAN TGXTM 4–20% polyacrylamide precast SDS-PAGE gels (BioRad) and stained using Coomassie blue. Gels were scanned using a Gel DocTM XR + system (BioRad), and molecular weight estimates for gel bands were calculated using Image Lab software version 5.0 (BioRad).

MS analysis. Each band or segment of interest was excised from the SDS-PAGE gel, subjected to tryptic digestion, and peptides analysed on a 4800 MALDI tandem Time-of-Flight Analyzer (MALDI TOF/TOF, Applied Biosystems, MA) at the Centre for Protein Research (University of Otago, Dunedin, New Zealand). The resulting peptide mass data were searched against the combined glowworm transcriptome assembly translated into all possible reading frames by Proteome Discoverer 1.4, using the Mascot search engine (<http://www.matrixscience.com>). Identified proteins were annotated by searching the GenBank non-redundant database at the NCBI (<http://www.ncbi.nlm.nih.gov>) using the BLASTX software²² with default parameters. An E-value cut-off of 1E-6 was used to identify similar annotated proteins from which function could be inferred.

Bench column and high pressure liquid chromatography (HPLC). *Instrumentation.* Agilent 1100 with photodiode array and Agilent 1260 Infinity with fluorescence detector, UV-Vis detector and automatic fraction collection unit. Columns: Phenomenex Synergi 4 µM Hydro RP 80 Å: analytical (150 × 4.6 mm) and semi-preparative (150 × 10 mm). *Mobile phase:* A, H₂O; B, acetonitrile. *HPLC method 1:* 0% B–6% B over 5 min, 6% B for 23 min, 6–25% B over 14 min, 25–100% B over 10 min, 100% B for 4 min, 100–0% B over 1 min, 0% B for 5 min. *HPLC method 2:* 0% B–6% B over 5 min, 6–100% B for 10 min, 100% B for 7 min, 100–0% B over 1 min, 0% B for 5 min with A and B both containing 0.1% formic acid. *Flow rate:* 0.5 ml/min for hydro C18 analytical column, 2.3 ml/min for hydro C18 semi-prep column. *Fluorescence detection:* excitation 280 nm, emission 415 nm. *UV Detection:* 210, 250, 270, 280, 300, 320, 350, 380 and 400 nm. The identities of compounds eluted using HPLC were confirmed by comparison to purchased standards analysed using LC retention time, MS, MS2 and UV-Vis spectra.

Anaerobic technique. Most handling of *A. luminosa* extracts during bench chromatography and HPLC was done under N₂ in a glove bag. When this was not possible, samples were kept in vials under N₂. Liquids were transferred between vials using a gas tight Hamilton syringe. All solvents were degassed by bubbling N₂ through the sample. Samples were concentrated by introducing frozen samples (−80 °C) to a desiccator inside a N₂ filled glove bag. The desiccator was then sealed, removed from the glove bag and attached to high vacuum until the samples were dried. Samples were kept under vacuum until the desiccator was opened inside a N₂-filled gas bag. They were then dissolved in degassed solvent and refrozen using dry ice.

Chromatography. Each lysate, crude luciferase and hot extract preparation used was first assayed for luminescent activity. Hot extract was then loaded onto a pre-equilibrated 0.5 g (unless stated otherwise) Isolute C18 column under N₂, eluted with a gradient of water to acetonitrile. Fractions were eluted using positive N₂ pressure. Eluted fractions were frozen on dry ice, concentrated using the method described above, then dissolved in water or buffer (0.1–1 ml), and samples assayed for the presence of luminescent compounds. Fractions with luminescent activity were combined, concentrated, and then further separated using HPLC method 1 under a N₂ atmosphere. HPLC fractions were also concentrated and assayed for luminescence activity.

Luminescence yield. The light intensity integrated over a selected time was compared between eluted fractions and the original column input material and hot extract and these values, along with concentration and subsampling factors, were used to determine the luminescence yield compared with the initial hot extract. Fractions were also analysed using HPLC method 1.

Isolation of LRC for stable isotope labeling analysis. Four reaction mixtures were prepared in triplicate as described in Results. The reactions were incubated inside the Clariostar well plate luminometer at room temperature for 5.5 h and light production measured. In preparation for LC-MS analysis, each reaction mixture was added to a pre-equilibrated Sorenson Bioscience Multiguard II Filter tip containing 30 mg of C18 resin. The filter tips were positioned in holes drilled through the lids of 1.5 ml plastic tubes. Centrifugation was then used to elute reaction mixtures through the C18-filled tips, which were then further washed with 6 × 200 µl of acetonitrile with 0.1% formic acid. For each reaction mixture, all eluted material was combined, and the resulting solution dried under vacuum, then stored at −80 °C. Each reaction mixture was dissolved in 100 µl of 50% acetonitrile, 50% water and 0.1% formic acid. 80 µl of this mixture was injected onto a hydro C18 semi-preparative column and eluted using HPLC gradient method 2 using 0.1% formic acid. Six fractions were collected: F1 (0–3.06 min), F2 (3.06–6.05 min), F3 (6.05–9.05 min), F4 (9.05–12.05 min), F5 (12.05–15.05 min) and F6 (15.05–18.05 min). Fractions 3 and 4 were pooled and all fractions were dried by vacuum. Fractions 3, 4 and 5 were analysed by Orbitrap LC-MS (supplementary methods). SIEVE software⁷⁰, was used to analyse LC-MS data, identifying labeled LRC by searching for co-eluting perfect pairs with expected mass differences (Supplementary Table S5).

LRC characterisation. A reaction mixture containing lysate (made from 30 light organs, 0.0231 g of tissue), unlabeled Tyr, XA and ATP-Mg²⁺ was set up at 10 °C and monitored for luminescence, with further Tyr, XA and ATP-Mg²⁺ added until no further luminescence was produced. Incubation continued for three days, after which LRC was isolated by reverse phase bench column chromatography and HPLC. Isolated LRC was dissolved in 200 µl of 50% acetonitrile, 50% water and 0.1% formic acid, then analysed by HPLC-UV and LC-MS. Initial analysis by Orbitrap LC-MS showed the sample to contain about 10 µg LRC (99% pure). UV-Vis spectra were taken using the HPLC photodiode array. The sample was dried and re-dissolved in 200 µl of 50% d3-acetonitrile and 50% deuterium oxide (D₂O) and placed into a D₂O Shigimi tube for analysis on a 500 MHz Varian NMR spectrophotometer using ¹H Presat (D₂O) settings for 12 h at 25 °C. Tyr (0.2 mg/ml), and XA (0.2 mg/ml) were analysed under the same conditions. Chemical shifts were referenced to the residual solvent signal for acetonitrile at 1.94 ppm. These solutions were then injected without a column into the Orbitrap MS so the number of exchangeable protons could be studied. The LRC solution was reanalysed by nanoflow orbitrap LC-MS to determine the amount of LRC degradation, then assayed for bioluminescent activity.

References

- Shimomura, O. *Bioluminescence: chemical principles and methods*. (World Scientific Publishing Co. Pte. Ltd., 2006).
- Baker, C. H. *et al.* Distribution and phylogenetic relationships of Australian glow-worms *Arachnocampa* (Diptera, Keroplatidae). *Mol Phylogenet Evol* **48**, 506–514 (2008).
- Green, L. F. The fine structure of the light organ of the New Zealand glow-worm *Arachnocampa luminosa* (Diptera: Mycetophilidae). *Tissue Cell* **11**, 457–465 (1979).
- Broadley, R. A. & Stringer, I. A. N. Prey attraction by larvae of the New Zealand glowworm *Arachnocampa luminosa* (Diptera: Mycetophilidae). *Invert. Biol* **120**, 170–7 (2001).
- Herring, P. J. Systematic distribution of bioluminescence in living organisms. *J Biolumin Chemilumin* **1**, 147–163 (1987).
- McCapra, F. Chemical generation of excited states: the basis of chemiluminescence and bioluminescence. *Methods Enzymol* **305**, 3–47 (2000).
- Silva, L. Pinto da. & Silva, J. C. G. Esteves da. Firefly luciferin as a multifunctional chemiluminescence molecule. *Photochem Photobiol Sci* 1615–1621 <https://doi.org/10.1039/c3pp50086a> (2013).
- Haddock, S. H. D., Moline, M. A. & Case, J. F. Bioluminescence in the sea. *Ann Rev Mar Sci* **2**, 443–493 (2010).
- Sharpe, M. L., Hastings, J. W. & Krause, K. L. Luciferases and Light-emitting Accessory Proteins: Structural Biology (2014), <https://doi.org/10.1002/9780470015902.a0003064.pub2> at <http://onlinelibrary.wiley.com/doi/10.1002/9780470015902.a0003064.pub2/abstract>. eLS (John Wiley & Sons, Ltd, 2001).
- Kaskova, Z. M., Tsarkova, A. S. & Yampolsky, I. V. 1001 lights: luciferins, luciferases, their mechanisms of action and applications in chemical analysis, biology and medicine. *Chem Soc Rev* **45**, 6048–6077 (2016).
- Navizet, I., Liu, Y.-J., Ferré, N., Roca-Sanjuán, D. & Lindh, R. The chemistry of bioluminescence: An analysis of chemical functionalities. *ChemPhysChem* **12**, 3064–3076 (2011).

12. Gulick, A. M. Conformational dynamics in the acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem Biol* **4**, 811–827 (2009).
13. Lee, J. Bioluminescence of the Australian glow-worm *Arachnocampa richardsae* Harrison. *Photochem Photobiol* **24**, 279–285 (1976).
14. Shimomura, O., Johnson, F. H. & Haneda, Y. In *Bioluminescence in progress* (eds. Johnson, F. H. & Haneda, Y.) 487–494 (Princeton University Press, 1966).
15. Viviani, V. R., Hastings, J. W. & Wilson, T. Two bioluminescent diptera: the North American *Orfelia fultoni* and the Australian *Arachnocampa flava*. Similar niche, different bioluminescence systems. *Photochem Photobiol* **75**, 22–27 (2002).
16. Silva, J. R., Amaral, D. T., Hastings, J. W., Wilson, T. & Viviani, V. R. A transcriptional and proteomic survey of *Arachnocampa luminosa* (Diptera: Keroplatidae) lanterns gives insights into the origin of bioluminescence from the Malpighian tubules in Diptera. *Luminescence* **30**, 996–1003 (2015).
17. Sharpe, M., Dearden, P., Gimenez, G. & Krause, K. Comparative RNA seq analysis of the New Zealand glowworm *Arachnocampa luminosa* reveals bioluminescence-related genes. *BMC Genomics* **16**, 825 (2015).
18. Trowell, S. C., Dacres, H., Dumancic, M. M., Leitch, V. & Rickards, R. W. Molecular basis for the blue bioluminescence of the Australian glow-worm *Arachnocampa richardsae* (Diptera: Keroplatidae). *Biochem Biophys Res Commun* **478**, 533–9 (2016).
19. Lloyd, J. E. Bioluminescence and communication in insects. *Annu Rev Entomol* **28**, 131–160 (1983).
20. Misof, B. *et al.* Phylogenomics resolves the timing and pattern of insect evolution. *Science* **346**, 763–767 (2014).
21. Harvey, E. N. *Bioluminescence*. (Academic Press, 1952).
22. Altschul, S. F., Gish, W., Miller, W., Myers, E. W. & Lipman, D. J. Basic local alignment search tool. *J Mol Biol* **215**, 403–410 (1990).
23. Branchini, B. R. *et al.* Mutagenesis evidence that the partial reactions of firefly bioluminescence are catalyzed by different conformations of the luciferase C-terminal domain. *Biochemistry* **44**, 1385–1393 (2005).
24. Branchini, B. R., Magyar, R. A., Murtiashaw, M. H. & Portier, N. C. The role of active site residue arginine 218 in firefly luciferase bioluminescence. *Biochemistry* **40**, 2410–2418 (2001).
25. Branchini, B. R., Southworth, T. L., Murtiashaw, M. H., Boije, H. & Fleet, S. E. A mutagenesis study of the putative luciferin binding site residues of firefly luciferase. *Biochemistry* **42**, 10429–10436 (2003).
26. Airth, R. L. & Foerster, G. E. The isolation of catalytic components required for cell-free fungal bioluminescence. *Arch Biochem Biophys* **97**, 567–573 (1962).
27. Purtov, K. V. *et al.* The chemical basis of fungal bioluminescence. *Angew Chem Int Ed Engl* **127**, 8242–8246 (2015).
28. Huneck, S. & Yoshimura, I. *Identification of Lichen Substances*. <https://doi.org/10.1007/978-3-642-85243-5> (Springer Berlin Heidelberg, 1996).
29. Theodoropoulos, S., inventor; Theodoropoulos, S., original assignee. Bifunctional chelating agents, their chelates and process of preparation. United States patent US 5,294,717. 1994 Mar 15.
30. McKerrow, J. D., Al-Rawi, J. M. A. & Brooks, P. Use of Diphenyliodonium Bromide in the Synthesis of Some N-Phenyl α -Amino Acids. *Synth Commun* **40**, 1161–1179 (2010).
31. Xue, L. *et al.* Ratiometric fluorescent sensor based on inhibition of resonance for detection of cadmium in aqueous solution and living cells. *Inorg Chem* **50**, 3680–3690 (2011).
32. Cui, P., McCalmont, W. F., Tomsig, J. L., Lynch, K. R. & Macdonald, T. L. Alpha- and beta-substituted phosphonate analogs of LPA as autotaxin inhibitors. *Bioorganic Med Chem* **16**, 2212–2225 (2008).
33. Diness, F., Beyer, J. & Meldal, M. Synthesis of 3-Boc-(1,3)-oxazinane-protected amino aldehydes from amino acids and their conversion into urea precursors. Novel building blocks for combinatorial synthesis. *QSAR Comb Sci* **23**, 130–144 (2004).
34. Neta, P., Simón-Manso, Y., Liang, Y. & Stein, S. E. Loss of H₂ and CO from protonated aldehydes in electrospray ionization mass spectrometry. *Rapid Commun Mass Spectrom* **28**, 1871–1882 (2014).
35. Oba, Y. Biosynthesis of firefly luciferin in adult lantern: decarboxylation of L-cysteine is a key step for benzothiazole ring formation in firefly luciferin synthesis. *PLoS ONE* **8**, e84023 (2013).
36. Petushkov, V. N. *et al.* A novel type of luciferin from the Siberian luminous earthworm *Fridericia heliota*: structure elucidation by spectral studies and total synthesis. *Angew Chem Int Ed* **53**, 5566–5568 (2014).
37. Oba, Y. & Kato, S. -ichi, Ojika, M. & Inouye, S. Biosynthesis of coelenterazine in the deep-sea copepod, *Metridia pacifica*. *Biochem Biophys Res Commun* **390**, 684–688 (2009).
38. Raabe, M. *Insect neurohormones and female reproduction*. **3**, (Elsevier, 1984).
39. Sugumaran, M. Comparative biochemistry of eumelanogenesis and the protective roles of phenoloxidase and melanin in insects. *Pigment Cell Research* **15**, 2–9 (2002).
40. Brunet, P. C. J. Tyrosine metabolism in insects. *Ann N Y Acad Sci* **100**, 1020–1034 (1963).
41. Huang, Z. *et al.* Cutaneous melanin exhibiting fluorescence emission under near-infrared light excitation. *J Biomed Opt* **11**, 34010–34016 (2006).
42. Gallas, J. M. & Eisner, M. Fluorescence of melanin-dependence upon excitation wavelength and concentration. *Photochem Photobiol* **45**, 595–600 (1987).
43. Kayatz, P. *et al.* Oxidation causes melanin fluorescence. *Invest Ophthalmol Vis Sci* **42**, 241–246 (2001).
44. Harriman, A. Further comments on the redox potentials of tryptophan and tyrosine. *J Phys Chem* **91**, 6102–6104 (1987).
45. Zhang, C., Boussac, A. & Rutherford, A. W. Low-temperature electron transfer in photosystem II: a tyrosyl radical and semiquinone charge pair. *Biochemistry* **43**, 13787–13795 (2004).
46. Sjödin, M., Styring, S., Åkermark, B., Sun, L. & Hammarström, L. Proton-coupled electron transfer from tyrosine in a tyrosine-ruthenium-tris-bipyridine complex: comparison with tyrosine₂ oxidation in photosystem II. *J Am Chem Soc* **122**, 3932–3936 (2000).
47. Whittaker, M. M., DeVito, V. L., Asher, S. A. & Whittaker, J. W. Resonance raman evidence for tyrosine involvement in the radical site of galactose oxidase. *J Biol Chem* **264**, 7104–7106 (1989).
48. Real, M. D. & Ferré, J. Distribution of xanthurenic acid glucoside in species of the genus *Drosophila*. *Insect Biochem* **19**, 111–115 (1989).
49. Real, M. D. & Ferré, J. Biosynthesis of xanthurenic acid 8-O- β -D-glucoside in *Drosophila*. Characterization of the xanthurenic acid: UDP-glucosyltransferase activity. *J Biol Chem* **265**, 7407–7412 (1990).
50. Ishiguro, I., Ikeno, T. & Matsubara, H. Studies on the fluorescent substances in the body of bees. *Yakugaku Zasshi* **94**, 116–123 (1974).
51. Smânia, A. *et al.* Toxicity and antiviral activity of cinnabarin obtained from *Pycnoporus sanguineus* (Fr.) Murr. *Phytother Res* **17**, 1069–1072 (2003).
52. Le Roes-Hill, M., Goodwin, C. & Burton, S. Phenoxazinone synthase: what's in a name? *Trends Biotechnol* **27**, 248–258 (2009).
53. Seligy, V. L. Ommochrome pigments of spiders. *Comp Biochem Physiol Part A* **42**, 699–709 (1972).
54. Hardeland, R. *et al.* In *Tryptophan, Serotonin, and Melatonin* (eds. Huether, G., Kochen, W., Simat, T. J. & Steinhart, H.) 389–395 (Springer US, 1999).
55. Lima, V. L. *et al.* The antioxidant role of xanthurenic acid in the *Aedes aegypti* midgut during digestion of a blood meal. *PLoS ONE* **7**, e38349. <https://doi.org/10.1371/journal.pone.0038349> (2012).
56. Sławiński, J., Elbanowski, M. & Sławińska, D. Spectral characteristics and mechanism of chemiluminescence from tryptophan solutions induced by UV irradiation. *Photochem Photobiol* **32**, 253–260 (1980).
57. Mofford, D. M., Reddy, G. R. & Miller, S. C. Latent luciferase activity in the fruit fly revealed by a synthetic luciferin. *Proc Natl Acad Sci USA* **111**, 4443–4448 (2014).

58. Viviani, V. R., Prado, R. A., Arnoldi, F. C. G. & Abdalla, F. C. An ancestral luciferase in the malpighi tubules of a non-bioluminescent beetle! *Photochem Photobiol Sci* **8**, 57–61 (2009).
59. Hastings, J. W. Biological diversity, chemical mechanisms, and the evolutionary origins of bioluminescent systems. *J Mol Evol* **19**, 309–321 (1983).
60. Shimomura, O., Inoue, S., Johnson, F. H. & Haneda, Y. Widespread occurrence of coelenterazine in marine bioluminescence. *Comp Biochem Physiol B* **65**, 435–437 (1980).
61. Shimomura, O. Presence of coelenterazine in non-bioluminescent marine organisms. *Comp Biochem Physiol B Biochem Mol Biol* **86**, 361–363 (1987).
62. Thomson, C. M., Herring, P. J. & Campbell, A. K. Evidence for de novo biosynthesis of coelenterazine in the bioluminescent midwater shrimp, *Systellaspis debilis*. *J Mar Biol Assoc U.K.* **75**, 165–171 (1995).
63. Day, J. C., Goodall, T. I. & Bailey, M. J. The evolution of the adenylate-forming protein family in beetles: multiple luciferase gene paralogues in fireflies and glow-worms. *Mol Phylogenet Evol* **50**, 93–101 (2009).
64. Gimenez, G., Metcalf, P., Paterson, N. G. & Sharpe, M. L. Mass spectrometry analysis and transcriptome sequencing reveal glowing squid crystal proteins are in the same superfamily as firefly luciferase. *Sci Rep* **6**, 27638 (2016).
65. Tsuji, F. I. Bioluminescence reaction catalyzed by membrane-bound luciferase in the “firefly squid”. *Watasenia scintillans*. *Biochim Biophys Acta* **1564**, 189–197 (2002).
66. Tsuji, F. I. Role of molecular oxygen in the bioluminescence of the firefly squid, *Watasenia scintillans*. *Biochem Biophys Res Commun* **338**(1), 250–3 (2005).
67. Petushkov, V. N., Rodionova, N. S. & Bondar, V. S. Study of the luminescence system of the soil enchytraeid *Fridericia heliota* (Annelida: Clitellata: Oligochaeta: Enchytraeidae). *Dokl Biochem Biophys* **391**, 204–207 (2003).
68. Rodionova, N. S., Bondar, V. S. & Petushkov, V. N. ATP is a cosubstrate of the luciferase of the earthworm *Fridericia heliota* (Annelida: Clitellata: Oligochaeta: Enchytraeidae). *Dokl Biochem Biophys* **392**, 253–5 (2003).
69. Delroisse, J. *et al.* A puzzling homology: a brittle star using a putative cnidarian-type luciferase for bioluminescence. *Open Biology* **7** (2017).
70. Thermo Fisher Scientific. Sieve User Guide version 2.2, Version A September 2014. 30 (2014).

Acknowledgements

We are very grateful to all those who assisted studied fireflies are beetles, members of the order with glowworm collection, including S. Nielsen, C. Wallace, Y. Cottam, N. and L. Collinson, L. Deverick, P. R. Armitage, B. Drayton, P. Ramritu, C. Rickard, M. Brown, G. Golding, S. Ferreira, A. Barl, R. Mortuza, M. Peng, S. Liu, T. Kim, P. Shantti, M. Pearce, C. Warren, S. Bell, R. Phillips, D. Badan, G. Whitting, S. De Silva, S. Carter, D. Real and Te Warena Taua, and Ngai Tahu, Dunedin City Council, Fiordland Department of Conservation and West Auckland Council for collecting permits. We thank D. Carne and T. Kleffmann (Centre for Protein Research, University of Otago), and I. Billingham, J. Pentelow, N. Joyce and L. Stevenson for assistance with MS and instrument access; and E. Ledgerwood and C. Marshall for luminometer access. We would also like to acknowledge the advice, assistance, and inspiration provided by Professor J. Woodland Hastings and Dr. Thérèse Wilson. Funding for this study was provided by the Foundation for Research, Science and Technology (now the New Zealand Ministry of Business, Innovation, and Employment), the Division of Health Sciences, University of Otago, New Zealand, a University of Otago Doctoral Scholarship to OCW, and the New Zealand Marsden Research Fund.

Author Contributions

O.C.W. and M.L.S. carried out sample collection and experiments and wrote the paper. N.B.P. and K.L.K. proposed the study, contributed to experiment design, and revised the manuscript. All authors read and approved the final manuscript.

Additional Information

Supplementary information accompanies this paper at <https://doi.org/10.1038/s41598-018-21298-w>.

Competing Interests: The authors declare no competing interests.

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2018