Phosphonated glycans as post-translational modifications of proteins in velvet worm slime

Alexandre Poulhazan[1]†, Alexander Baer[2]†, Gagan Daliaho[3], Frederic Mentink-Vigier[4], Alexandre A. Arnold[1], Darren Browne[5], Lars Hering[2], Stephanie Archer-Hartmann[6], Lauren E. Pepi[6], Parastoo Azadi[6], Stephan Schmidt[7], Georg Mayer[2], Isabelle Marcotte[1]*, Matthew J. Harrington[3]*

[1]Department of Chemistry, University of Quebec at Montreal, Montreal, H2X 2J6, Canada.

[2]Department of Zoology, Institute of Biology, University of Kassel, D-34132, Germany.

[3]Department of Chemistry, McGill University, Montreal, QC H3A 0B8, Canada.

[4]National High Magnetic Field Laboratory, Tallahassee, FL 32310, U.S.A.

[5]Department of Biological and Chemical Sciences, University of the West Indies, Cave Hill Campus, Barbados, BB11000, West Indies.

[6]Complex Carbohydrate Research Center, University of Georgia, Athens, GA 30602, USA.

[7]Chemistry Department, Heinrich-Heine-Universität Düsseldorf, D-40225, Germany.

† These authors contributed equally to this work

* Corresponding authors. Emails: marcotte.isabelle@uqam.ca; matt.harrington@mcgill.ca

Abstract:

To capture prey, onychophorans (velvet worms) expel a slime that forms stiff fibers upon shearing and dehydration. The high quantities of phosphorus previously found in the slime of the velvet worm *Euperipatoides rowelli* were ascribed to protein phosphorylation. We provide clear evidence, instead, that it is primarily present as phosphonate moieties in the slime of representative from both major onychophoran subgroups which diverged ~380 MYA. Advanced NMR and mass spectrometry demonstrate that 2-aminoethyl phosphonate (2-AEP) is associated with high molecular weight slime proteins as phosphonoglycans. Biogenic phosphonates are a substantial component of the organophosphorus cycle in marine environments but were not previously reported in terrestrial invertebrates. The evolutionary conservation of this rare protein modification suggests a potential role in the formation and function of these biological adhesive fibers with implications for bio-inspiration.

Introduction

Onychophorans (commonly referred as velvet worms) represent an ancient group of terrestrial invertebrates with ~200 described species. Peripatidae and Peripatopsidae, the two major subgroups, diverged about 380 MYA [1]. Velvet worms capture preys with a unique strategy: they eject a sticky slime from dedicated papillae on either side of the head [2] (**Movie S1**). The adhesivity of the slime combined with its ability to form recyclable fibers from a liquid protein solution upon mechanical shearing open a promising avenue in bio-inspired and reusable material research [3].

Biochemical analysis of slime from several species has been performed, revealing a primarily proteinaceous composition with components of different sizes. Mid-molecular-weight (MMW) proteins together with small quantities of lipids (<1 %) were proposed to form nanoscale condensed liquid nanodroplets [4], while low molecular weight (LMW) proteins may act as antimicrobial components [5]. However, several high molecular weight (HMW) proteins were shown to be the major structural component of slime fibers [3, 5-9]. Based on a positive phosphostaining and elemental analysis of the HMW proteins from the species *Euperipatoides rowelli* and the high content of divalent cations (Mg^{2+}, Ca^{2+}), electrostatic interactions between phosphorylated amino acids complexing free cations were previously proposed to initiate fiber formation [6, 10]. In addition, phosphorylated amino acids were directly identified.

Here, we investigated the chemical nature of the phosphorus found in slime from two distant species of velvet worms: *Euperipatoides rowelli* (Peripatopsidae) and *Epiperipatus barbadensis* (Peripatidae). Phosphorus is ubiquitous in living organisms and typically found as phosphate esters (C-O-P bond) [11], but rarely as phosphonates (C-P bond) in natural organophosphorus compounds [12]. Phosphonates occur mostly in marine environments, representing a substantial part of marine dissolved organic phosphorus [13]. While biogenic phosphate esters have been well studied, the prevalence of natural phosphonates is underestimated [12-14]. Overall, only about 40 natural phosphonates have been reported in the literature, mostly as small antimicrobial molecules [15]. In some occurrences, they have been found decorating larger molecules such as lipids, polysaccharides, and proteins in marine microorganisms [12, 13, 15].

Results and Discussion

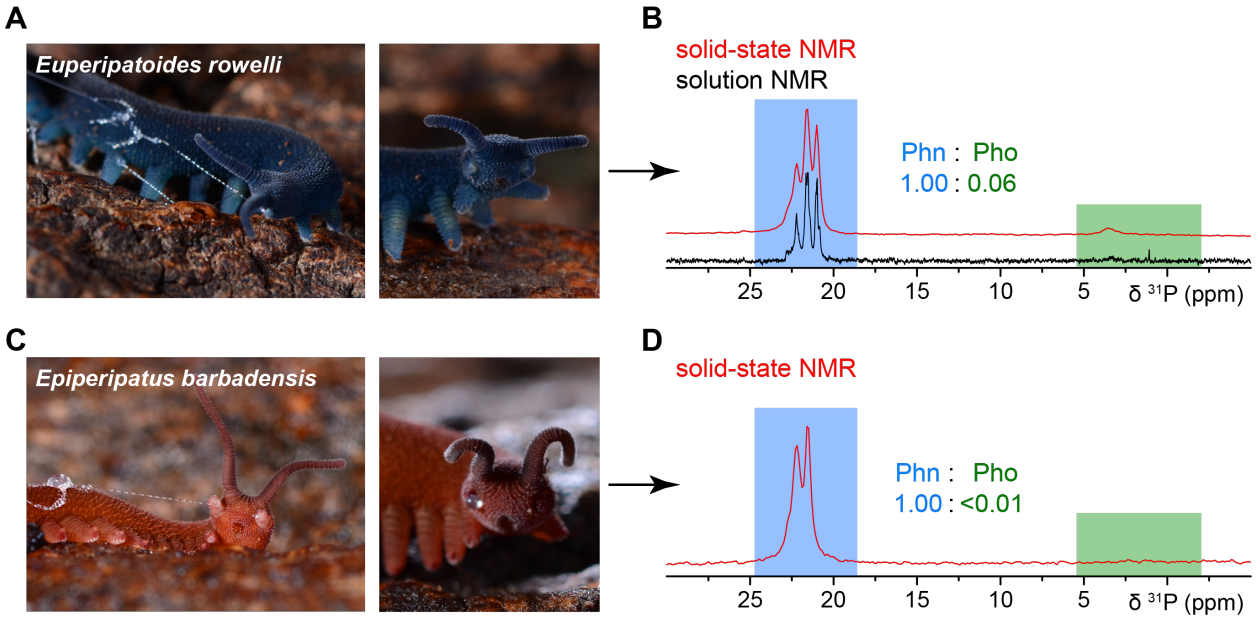

Fig. 1. ^{31}P ssNMR reveals phosphonates in slime of two distantly related species of velvet worms. (**A**) Photographs of the peripatopsid *Eu. rowelli*. (**B**) Solution (*black* line) and ssNMR (*red* line) ^{31}P NMR spectra indicate that phosphorus is present in slime of *Eu. rowelli* predominantly in the form of phosphonates (Phn, highlighted in *blue*), while the occurrence of phosphate is low (Pho, indicated in *green*). (**C**) Photographs of the peripatid *Ep. barbadensis.* (**D**) ^{31}P ssNMR spectrum of slime of *Ep. barbadensis* indicates only phosphonates.

Solution and solid-state nuclear magnetic resonance spectroscopy (ssNMR) were used to identify the molecular composition of slime produced by *Eu. rowelli* (**Fig. 1A** and **B**) and *Ep. barbadensis* (**Fig. 1C** and **D**) [16, 17]. Several intense ^{31}P NMR signals appeared at 20–23 ppm (**Fig. 1B** and **D**, **fig. S1**), which are unambiguously assigned to phosphonates. Additional weaker peaks between 0–5 ppm, only found in *Eu. rowelli*, are ascribed to phosphate esters involved in phospholipids or more likely slime phosphoproteins (**Fig. 1B** and **D**) [6]. Quantitative NMR reveals that the slime of *Eu. rowelli* contains 17 times more phosphonate than phosphate (**Fig. 1B**), and phosphate is quasi absent in *Ep. barbadensis* (**Fig. 1D**). This difference is supported by Fourier-transform infrared (FTIR) spectroscopic analyses, which revealed highly similar amide I protein bands but slight differences in the band corresponding to P-O vibrations typical of phosphates and phosphonates (900–1150 cm^{-1}) [18-20] (**fig. S2**). These results provide strong evidence that

the high amount of phosphorus primarily interpreted as phosphorylated proteins [6, 10] is, in fact, associated with phosphonate-rich molecules. Furthermore, similarities of ^{31}P solution and ssNMR spectra in *Eu. rowelli* slime (**Fig. 1B, fig. S3** and **table S1**) indicate that our findings apply to both fibrilized and non-fibrilized slime. The spectra from both species differ in the phosphonate region, as three peaks were detected in the slime of *Eu. rowelli* at 22.2/21.6/21.0 ppm while the 21.0 ppm peak is absent in *Ep. barbadensis*. Thus, while the overall compositions of the two velvet worm subgroups slimes appear similar (**fig. S3**), subtle differences between their phosphonates remain to be explored.

The occurrence of phosphonate in the slime and body (**fig. S3**) of representatives from both onychophoran subgroups suggests that phosphonate production evolved over at least 380 MYA and is thus shared by all slimes of existing velvet worm species. Given the large metabolic cost [13], the evolutionary conservation of phosphonate production hints at a functionally important role in slime fiber formation and/or function. To the best of our knowledge, this is the first occurrence of phosphonates in terrestrial invertebrates. Given that phosphonate producers are mainly found in marine environments, and velvet worms comprise the only major invertebrate "phylum" of exclusively terrestrial animals [21], this result is unexpected. However, it is noteworthy that extinct lobopodian lived in marine habitats and are the ancestors of velvet worms, water bears and arthropods [22]. Phosphonate production therefore might be an ancestral feature which could stil be present in water bears and arthropods.

To identify the chemical structure of the phosphonate moiety in the slime, we first conducted high-resolution ^{31}P 1D and ^{1}H-^{31}P 2D solution NMR experiments on diluted non-fibrilized *Eu. rowelli* slime (**figs. S4** and **S5**, and **table S1**). The peak assignments were compared to several phosphonate standards that have been described as precursors of natural phosphonate [23]. A good agreement was obtained with 2-aminoethylphosphonate (2-AEP). In marine microorganisms, the biosynthesis pathway of 2-AEP is catalyzed by phosphoenolpyruvate mutase (PEPm), phosphoenolpyruvate decarboxylase (Ppd) and 2-aminoethyl phosphonate transaminase (AEPt) [13, 24-26]. Local BLAST searches of published protein sequences of these three enzymes [25] against different transcriptomes of *Eu. rowelli* and *Principapillatus hitoyensis* (another peripatid species) revealed that the PEPm-, Ppd- and AEPt-encoding genes are indeed expressed

in the slime glands of representatives of both major onychophoran subgroups (**tables S2-5**), providing evidence for the natural ability of velvet worms to produce phosphonate moieties in the form of 2-AEP. Interestingly, these genes are expressed in several tissues in addition to the slime gland, consistent with the detection of phosphonates not only in slime, synthesized by the slime gland, but also in other parts of the body (**table S3** and **fig. S3**), suggesting that 2-AEP may fulfill various biological functions in velvet worms.

Further evidence for the existence of 2-AEP moieties in velvet worm slime was obtained using Magic-Angle Spinning of the sample combined with Dynamic Nuclear Polarization (MAS-DNP), a high-resolution NMR-related technique that allows tremendous sensitivity enhancement (**Fig. 2A** and **fig. S6**). In addition to carbon detection and identification, structural elucidation was enabled by probing the proximity between carbon and phosphorus atoms. Briefly, magnetization is transferred from ^{31}P to ^{13}C using cross-polarization (CP); the closer the nuclei, the faster the magnetization transfer. By increasing the duration of the CP, carbons that are further from the phosphorus gradually appear on the spectra (**Fig. 2A** and **B**). 2D MAS-DNP ^{31}P-^{13}C CP experiments of slime confirm that ^{13}C-^{31}P contacts arise from phosphonates and not phosphorylation (**fig. S7**). The carbons in closer proximity to the phosphonates were found to be those at chemical shifts of 24.8 ppm and 35.4 ppm (**Fig. 2B**, **fig. S8**), which corresponds well to the structure of 2-AEP (**table S1**), and this is further correlated by a signal at ~31 ppm in the ^{15}N spectrum (**fig. S9**). Carbons away from the phosphorus atom appeared at 63.3 ppm and 75.0 ppm in *Eu. rowelli* and 61.1 ppm, 66.5 ppm and 75.0 ppm for *Ep. barbadensis*, typical of glycan signals (**Fig. 2A** and **B** and **table S6**) [27], giving strong indication that phosphonated glycans are in the slime, similar to modifications observed in bacteria capsules [28].

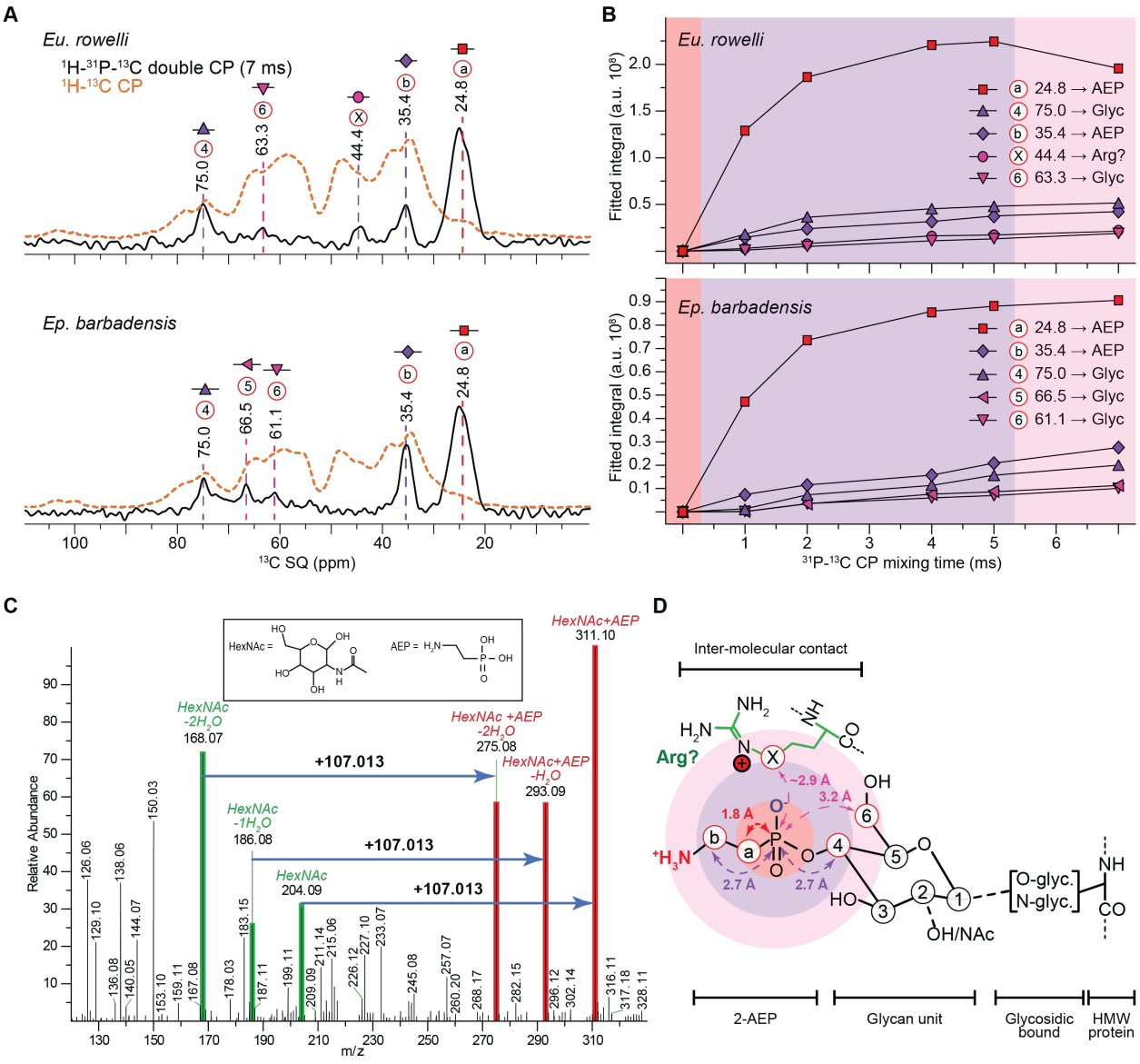

Fig. 2. MAS-DNP and MS identify phosphonoglycans in velvet worm slime. (**A**) MAS-DNP ^{1}H-^{31}P-^{13}C double cross-polarization (CP) spectra demonstrated carbon atoms close to ^{31}P atoms (*black* line) compared to direct ^{1}H-^{13}C CP spectra (*orange* dashed line) that reveal all carbons in the slime. (**B**) Plotting of ^{13}C peak intensities in the slime of both species as function of CP mixing time allows to sequentially assign the carbon close to phosphorus with 24.8 ppm, then 75.0 ppm and 35.4 ppm peaks further away, followed by 44.4 ppm (66.5 ppm in *Ep. barbadensis*) and finally 63.3 ppm (61.1 ppm in *Ep. barbadensis*). (**C**) *Eu. rowelli* slime MS/MS fragmentation of putative peptide with 2-AEP modified glycan attached. Mass range is from *m/z* 120–330 to show oxonium ions indicative of unmodified and 2-AEP modified HexNAc. (**D**) Combining proton, carbon, and

nitrogen chemical shifts assignments (**table S1**), and mass spectrometry allows tentative determination of phosphonoglycans decorating proteins in velvet worms. In slime of *Eu. rowelli*, intermolecular contacts are detected even between phosphonates and presumably arginine sidechain (*X atom*) (**fig. S11**). Disks around the phosphonate phosphorus correspond to 1.8 Å (red), 2.7 Å (purple) and 3.2 (pink) Å.

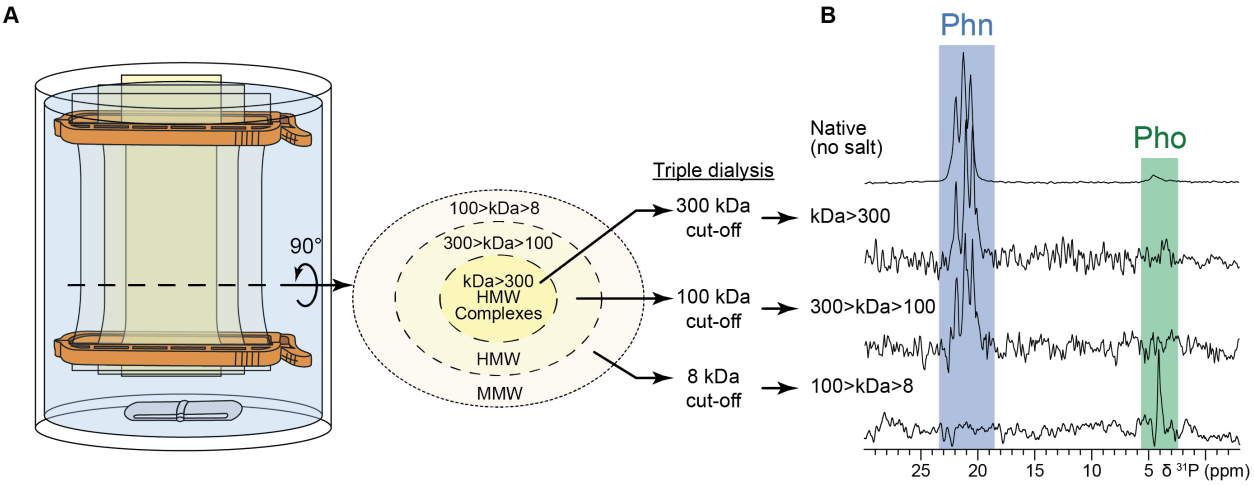

Fig. 3. Triple dialysis of slime from *Eu. rowelli* followed by ^{31}P ssNMR. (**A**) Simultaneous triple dialysis set-up. (**B**) The resulting dialysis fractions are analyzed by ^{31}P ssNMR showing that HMW proteins (>100 kDa) contain most of the phosphonates (Phn), while LMW proteins (<100 kDa) lack phosphonates but contain phosphates (Pho).

Higher-energy collision induced dissociation (HCD) tandem mass spectrometry (MS/MS) analysis [29] of trypsin-digested *Eu. rowelli* slime confirmed the presence of a 2-AEP modification on glycans. The results showed oxonium ions revealing both unmodified and 2-AEP-modified N-acetylhexosamines (HexNAc) decorating tryptic peptides (**Fig. 2C** and **fig. S10**) indicating that slime proteins are modified with phosphonoglycans. Previous biochemical analyses of the peripatopsid species *Eu. kanangrensis* ─ closely related to *Eu. Rowell* ─ assumed that slime carbohydrates mostly occur as N-acetyl galactosamine (GalNAc) [8] probably bonded to proteins via O-glycosylation. However, the exact linking pattern as well as the carbohydrate units in velvet worm slime need to be further investigated (**Fig. 2D**). Finally, while the presence of 2-AEP and its association to proteins are supported for both species, an intermolecular contact with a carbon at 44.4 ppm (**Fig. 2A** and **fig. S8**) was detected only in *Eu. rowelli*'s slime. Although this resonance can also

correspond to a glycine Cα, it is more likely due to arginine's Cδ as identified through correlations detected on 2D ^{13}C spectra (**fig. S11**). DFT calculations using our final molecular model (**fig. S12**) support a ~2.9 Å distance between 2-AEP phosphorus and arginine in the slime. This intermolecular contact of phosphonate with arginine that bears a positive charge at the slime pH of 5.2 is reminiscent of the zwitterionic phosphonoglycans in bacteria [30] (**fig. S13**). A zwitterionic charge state suggests the promotion of non-covalent electrostatic interactions between slime proteins during fiber formation, as previously proposed for *Eu. rowelli* [6]. This local addition of charges to the HMW proteins is also expected to increase solubility in the slime storage and recycling.

^{1}H and ^{31}P solution NMR diffusion experiments revealed that phosphonates are most likely associated with large molecules (**fig. S14**), while lipid extraction and both ^{13}C ssNMR or phenol-sulfuric acid assay showed low amounts of lipids or glycans (2.0 ± 0.5 in *Eu. rowelli* [6] and 0.300 ± 0.005 mg/mL and *Ep. barbadensis* slime) (**figs. S15** and **S16**), therefore suggesting that phosphonate modification mainly occurs on HMW proteins. Further support for phosphonoglycan-modified proteins was provided by NMR experiments on fractionated slime of *Eu. rowelli* into HMW (>300 kDa), MMW (100–300 kDa) and LMW (8–100 kDa) fractions obtained from a triple dialysis (**Fig. 3A**). Indeed, ^{31}P ssNMR revealed that phosphonates are linked to molecules with molecular weights higher than 100 kDa, whereas LMW compounds are only phosphorylated (**Fig. 3B**). According to previously published SDS-PAGE data on velvet worm slime [7, 9], our HMW and MMW fractions correspond to HMW monomers/complexes (232–429 / 478–634 kDa) proteins in *Eu. rowelli* slime, and MMW (110 kDa) to some extent. Altogether, our results showed that 2-AEP is part of a phosphonoglycan modification of HMW proteins in onychophoran slime (**Fig. 2D**).

Conclusions

Our results provide insight into the molecular composition and assembly of velvet worm slime – an important source of bio-inspiration for sustainable polymers and adhesives. Although the exact chemical structure and function of this highly unusual protein modification requires further investigation, the evolutionary conservation of phosphonates in the velvet worm over a period of over 380 million years argues for an important functional role, given the high energetic cost to produce this modification. The discovery

of phosphonates in a terrestrial invertebrate once again underlines the necessity to consider phosphonates as a potential source of organophosphorus in understudied groups of animals and their functions in biological materials.

Materials and Methods

Materials. All chemicals, including enzymes, chemicals, phosphate and phosphonate standards, were purchased from Sigma-Aldrich Co. (Oakville, ON, Canada). All chemicals were used without further purification. Dialysis tubings (300 kDa, 100 kDa and 3 kDa cutoffs) were obtained from VWR international (Mississauga, ON, Canada) and rinsed in Nanopure® water prior to use.

Specimen collection and maintenance. Experiments were performed on the peripatopsid species *Euperipatoides rowelli*, Reid 1996 and the peripatid species *Epiperipatus barbadensis*, (Froehlich, 1962 and *Principapillatus hitoyensis*, Oliveira *et al.*, 2012. Specimens were obtained from decaying wood and leaf litter at the corresponding type localities and maintained in the laboratory as described previously [3, 31]. Specimens were collected and exported under the following permits: (i) SL101720/2016, issued by NSW National Parks and Wildlife Service, Australia, and PWS2016-AU-001023 and Department of Sustainability, Environment, Water, Population and Communities, Australia, (ii) 8434/56/1, Ministry of Environment and National Beautification, Green and Blue Economy, Barbados and (iii) 123-2005-SINAC and 014950, Gerencia Manejo y Uso Sostenible de RR NN–Ministerio del Ambiente y Energia, Costa Rica. All animal treatments complied with the principles of laboratory animal care and the local laws on the protection of animals. Worms were fed with living banded cricket (*Gryllodes sigillatus*). Slime samples for each experiment were obtained from several specimens by stimulating them to eject the slime into 1.5-mL Eppendorf tubes. Collected slime was stored in the fridge for no longer than four days at 4 °C to avoid bacterial growth or potential degradation of proteins. Samples for shipments were lyophilized and rehydrated in milliQ water under gentle movements and additional heat treatments at 30 °C.

Sample collection and preparation. *Eu. rowelli* slime was harvested in Germany and shipped to Canada for chemical treatment, dialysis and NMR analysis using express

courier. Native slime was preferred for shipping, but dry samples proved to provide similar results for both solid-state and solution NMR using 35 °C bath followed by circular rotation using Glas-Col (Terre-Haute, IN, USA) hardware to rehydrate dry slime. *Ep. barbadensis* slime was harvested and analyzed in Canada without long-distance transportation. All NMR analysis on native slime were performed on at least three batches of slime for biological consistence.

Triple dialysis. To separate HMW, MMW and LMW slime proteins, we performed three (300 kDa, 100 kDa and 8 kDa cutoffs) simultaneous dialyses in 1M NaCl (and 0.02 % sodium azide) to prevent protein aggregation and nanoglobule collapse [6] (and bacterial growth). By keeping these conditions during dialysis, all protein molecules/complexes are expected to be free in solution and efficiently separated based on their hydrodynamic size. After seven days, salt was slowly removed by dialysis in 0.5 M NaCl for five days before using Nanopure® water (with 0.02 % sodium azide). After five days, the different dialysis tubings were emptied and freeze-dried, showing protein accumulation on the 300 kDa dialysis tubing indicating that, despite NaCl, most of the slime proteins are HMW and some protein aggregation happened.

Nuclear Magnetic Resonance. NMR measurements were performed on slime in the liquid state, as well as on aggregated slime and slime fibers in the solid state. Liquid slime samples were successfully filled into an NMR tube (diameter 5 mm) without forming fibers. This was performed first by diluting the native slime in D_2O increasing the fluidity and preventing partial aggregate formation. The slime appeared to aggregate during insertion into standard 7 inches-long NMR tubes. The aggregation could be significantly reduced by shortening the NMR tubes to ca. 4 inches. To form aggregates amenable to solid state NMR (ssNMR), the slime was simply agitated and mixed using a spatula. Aggregated slime appeared as a white sticky, gum-like material which was recovered and packed into an ssNMR rotor. All solution NMR experiments were conducted on a 600 MHz (14.1 T) Bruker Avance III HD NMR spectrometer using a Bruker BBFO probe at 25 °C. SsNMR experiments were conducted on the same spectrometer using a Varian 3.2 mm HXY probe at a Magic-Angle Spinning (MAS) frequency of 15 kHz at 25 °C or a Bruker 4 mm double resonance HX probe at a spinning frequency of 10 kHz. Additional

ssNMR spectra were recorded on a 400 MHz (9.4 T) Bruker Avance III HD spectrometer using a 4 mm double resonance HX probe or a Bruker 1.9 mm HXY triple-resonance probe when the sample amount was limited. All experiments at this field were carried out at a spinning frequency of 10 kHz. ^{31}P NMR spectra of solid samples were recorded using a single 90° pulse followed by spinal-64 ^{1}H decoupling at a radio-frequency field of 38 kHz, up to 30,000 scans were recorded with a recycle delay of 3 s. ^{31}P NMR spectra of soluble samples were recorded using a single 90° pulse followed by waltz-16 decoupling at a ^{1}H radio-frequency field of 3.5 kHz, 400 scans were recorded with a recycle delay of 5 s. The ^{31}P-^{1}H heteronuclear TOCSY were recorded using a 70 ms-long DIPSI2 spin lock. The spectral widths were 50 and 10 ppm for ^{31}P and ^{1}H respectively. Spectra of the phosphonate standards were recorded with 16 scans while the slime samples required 128, in both cases the indirect dimension consisted of 128 increments. Additional NMR parameters can be found following the references given in the figure captions.

For MAS-DNP enhanced experiments, the samples were prepared similar to the ssNMR, except that aggregates were mixed with a radical that contained unpaired electrons. The sample also contained a cryoprotectant to avoid formation of ice crystals in order to preserve potentially fragile structures. For our samples, a radical stock solution of AMUPol or AsymPol-POK [32-35] (for *Eu. rowelli* and *Ep. barbadensis* aggregates slime, respectively) was freshly prepared in d_8-glycerol/D_2O/H_2O (60/30/10 vol%), referred to as the *DNP juice* or the DNP matrix, at a final radical concentration of 10 mM. To prepare the MAS-DNP slime sample, 50 µL of the stock solution was added to ~50 µL of the slime and mixed for 5–10 min to allow the radical solution to penetrate the sample. Around 35 mg of well-hydrated samples were transferred into a 3.2-mm sapphire rotor. All MAS-DNP experiments were performed on a 600 MHz/395 GHz MAS-DNP spectrometer equipped with a gyrotron microwave source [36, 37] using a 3.2 mm, custom-made, HXY probe at 8 kHz MAS frequency. The cathode currents of the gyrotron were 160 mA. The temperature was ~98 K with the microwave (MW) off and ~100 K with the MW on. The radical in the *DNP juice* contained free electrons that had a 658-fold higher spin polarization than ^{1}H nucleus. Under the microwave radiation, the electron spin polarization is transferred to the nuclei and the sample remains at a cryogenic temperature. Typically, a 60-fold enhancement factor of NMR sensitivity with and without microwave irradiation was achieved (**fig. S6B**). Relatively short buildup time constants (1.4 s) indicated a good

mixing of the radicals and biomolecules in these cellular samples. 2D correlation experiments were implemented with 50-ms DARR, 1.5-s PDSD, NCa, and 100-ms NCaCx experiments. The total time for each 2D spectrum was 6 h, 8 h, and 2 h for the DARR, PDSD, and NC experiments respectively.

All ssNMR ^{13}C chemical shifts (including those obtained by MAS-DNP) were externally referenced to adamantane's CH_2 downfield signal set to 38.48 ppm [38], while ^{31}P chemical shifts were referenced using an 85 % aqueous solution of phosphoric acid (H_3PO_4) set to 0 ppm [23], and ^{15}N chemical shifts were referenced using an NH_4Cl powder set to 24.9 ppm [39].

Density Functional Theory calculations. DFT calculations have been performed to obtain the local geometry of the phosphonates. A truncated structure was used as input and its geometry was optimized using Density Functional Theory (DFT). We used the composite electronic -structure method based on r2SCAN [40] called r2SCAN-3c [41]. The computations were carried out using ORCA v5.0 [42].

The optimized resulting geometry is:

43
Coordinates from ORCA-job r2SCAN-3c

C	-6.28617023264502	2.38139257254500	0.09243181387826
C	-6.23319597485045	2.05327301370780	-1.38556430750551
N	-7.27550585414025	2.87841114782770	-2.14580934073389
H	-6.45842569746577	1.00363945817936	-1.59586426540360
H	-5.25985091576539	2.30364389033574	-1.81777404582538
P	-4.85694149664255	1.58725438452218	0.92302550383774
H	-7.21390017197148	2.04681727584072	0.57229967294640
H	-6.17813314256961	3.45819189104783	0.28289351500408
O	-3.63485870684023	1.67799529632120	0.08139108375378
H	-8.22105262005466	2.67905642835347	-1.80843775303848
H	-7.25085178661586	2.68537651805075	-3.15081994348680
O	-5.45034616040685	0.12161862425147	1.19069124995645
O	-4.77945536760181	2.21003189301643	2.38541050875373
H	-6.76208634838106	5.41292958969770	2.60973760554928
O	-6.27137323736572	4.64150058819208	2.29104721501860
N	-5.55637861957412	6.59330711900119	4.30134812242205
H	-4.22278539502546	6.26693046614186	2.76805314277530
C	-5.22043373738608	4.41827973660760	3.23504736354171
C	-4.57274285981645	5.73429613465205	3.66302047498798
H	-2.58163442753997	7.20089721397166	4.19223277010100
H	-2.10132820374292	1.52660210370726	2.75477190350441
C	-3.34481020733187	5.46294034004159	4.54734859661472

O	-2.71403278248025	6.62107603527916	4.95654440348768
O	-2.43344848011701	4.66279410434161	3.79514101659690
C	-2.94048661914480	3.37177617888495	3.48282893620705
C	-1.82132251284830	2.58991397609621	2.79525907713400
H	-3.63189204407957	4.92452594300130	5.47047891361482
H	-0.93131320421292	2.67741006501889	3.42887374835569
H	-2.00077233047284	2.59996779948339	0.85640169214354
O	-1.50115456015412	3.10244480059304	1.51587011448705
C	-4.17419137744821	3.53402170162016	2.59249141512786
H	-5.62133990259181	3.89132171624697	4.12082274848909
H	-3.23203287104229	2.83649887141554	4.40737965995110
H	-3.85518810619690	3.96134067580849	1.63122233858001
H	-7.10612484482561	3.87941502559830	-2.01265106192430
H	-4.77845883448240	-0.54680182638191	1.39311357462068
C	-5.47734222755223	7.96940102499645	4.11921886104275
H	-5.91778084125749	6.28382552862959	5.19471192476024
C	-6.30095825198778	8.79179488514959	5.07753396247000
H	-6.69432489610850	9.66258116779924	4.55050201535245
H	-5.63891098885725	9.15620943847166	5.87120907124848
H	-7.12058375354267	8.23243785038924	5.53697924386092
O	-4.79491940686278	8.46103935154501	3.23712745774188

Transcriptomic analysis. Specimens of *Eu. rowelli* and *P. hitoyensis* were anesthetized with chloroform vapor and dissected in physiological saline [43]. Total RNA from various tissues of both species (**table S2**) was extracted using TRIzol reagent (Invitrogen, Carlsbad, CA) and purified with RNeasy MinElute Cleanup Kit (Qiagen, Hilden, Germany) according to the manufacturers' specifications. Total RNA samples were shipped to the sequencing company for library preparation and sequencing (Eurofins Genomics Europe Sequencing GmbH, Konstanz, Germany). Paired end RNAseq runs (2 x 150 bp) were performed on a Illumina NovaSeq 6000 S4 system (Illumina, San Diego, CA). Raw reads were quality filtered and trimmed using Trimmomatic v.0.39 ([44], ILLUMINACLIP:TruSeq3-PE-2.fa:2:30:10 LEADING:30 TRAILING:30 SLIDINGWINDOW:4:30 MINLEN:50). To validate the filtering step, the raw reads were quality checked before and after trimming using FastQC v.0.11.9 and assembled afterwards using IDBA-Tran v.1.1.1 ([45], --mink 19 --maxk 124 --step 5 --min_contig 200 --max_isoforms 1[--*max_isoforms 3* in case of assembly *Eu. rowelli Slime gland: endpieces*]). Completeness of tissue-specific transcriptome assemblies were assessed using the metazoan dataset of BUSCO v.5.4.4 in transcriptome mode [46].

Obtained sequenced raw reads were submitted to SRA archive of GenBank (see **table S2** for assembly statistics and accession numbers).

Known enzymes involved in biocatalytic pathways of phosphonates [25] were searched in tissue-specific transcriptomes of *Eu. rowelli* and *P. hitoyensis* using BLAST v.2.13 [47](**table S3**). Resulting transcripts were confirmed by reciprocal BLAST searches against the NCBI database resulting in high-quality hits for enzymes catalyzing phosphonates (**table S4**). In addition, sequences were verified by cloning, which includes first strand cDNA synthesis from total RNA of both species using SuperScript™ IV First-Strand Synthesis System (Thermo Fischer, Steingrund, Germany), amplifying according genes using Phusion™ Plus Green PCR Master Mix (Thermo Fischer, Steingrund, Germany) and Invitrogen Custom Primers (Thermo Fischer, Steingrund, Germany), ligation and transformation using CloneJET PCR Cloning Kit (Thermo Fischer, Steingrund, Germany). Resulting clones were amplified and selected for sequencing (Sanger sequencing, LightRun, GATC, Eurofins Genomics, Ebersberg, Germany) (**table S5**). Preliminary assembled tissue-specific transcriptomes, and obtained nucleotide and protein sequences are given on tables **tables S2** and **S5**.

Mass spectrometry. Prior to mass spectrometry, slime was digested using trypsin enzyme as previously described [48]. Briefly, dry slime is slowly solubilized in a buffer containing 6M guanidine, 500 mM Tris-HCl (pH 8.0) and 3 mM DTT. The mixture is heated at 37 °C for an hour before adding 50 mM ammonium bicarbonate, 1mM calcium chloride solution to lower guanidine concentration to 1M, finally adjusting pH for best trypsin activity. Protein:protease ratio was set to 20:1 with final protein concentration of 0.3 mg/mL. Digestion is initiated and kept at 37 °C overnight. Digested slime is further separated according to their charges using SP (cation exchange) and Q (anion exchange) sepharose peptide ion-exchange columns (Cytiva columns) [49]. Tryptic peptides were resuspended in 0.1 % formic acid and injected into a Thermo Fisher Orbitrap Eclipse Tribrid MS coupled with an Ultimate RSLCnano 3000 system. Samples were run using a 180-minute gradient with low to high acetonitrile containing 0.1 % formic acid. Samples were analyzed in positive ion mode using a data dependent MS/MS method of the highest intensity peaks. Higher-energy collision induced dissociation (HCD) fragmentation was collected, and if the presence of HexNAc (*m/z* 204.0866) was found, a second scan using

collisional-induced dissociation (CID) was also collected. Samples were analyzed using Byonic Software (v4.0.12) and Thermo Fisher Xcalibur. Ion traces were used by extracting the m/z value of HexNAc oxonium (*m/z* 204.0866) and HexNAc+2-AEP (*m/z* 311.099). Fragmentation data was analyzed by determining neutral loss mass shifts.

Phenol-sulfuric acid assay. Slime carbohydrate analysis were performed using published method [50]. Briefly, native slime was diluted to reach 2 µM overall glycan concentration. 50µL of the sample was placed in a 96-well plate well before adding 150µL of concentrated sulfuric acid and 50µL of 5% phenol solution. The plate was then heated at 90°C for 5 min before measuring 490 nm absorbance to deduce glycan concentration.

References

(1) Oliveira, I. S.; Bai, M.; Jahn, H.; Gross, V.; Martin, C.; Hammel, Jörg U.; Zhang, W.; Mayer, G., Earliest Onychophoran in Amber Reveals Gondwanan Migration Patterns. *Curr. Biol.* **2016,** *26*, 2594–2601.

(2) Mayer, G.; Oliveira, I. S.; Baer, A.; Hammel, J. U.; Gallant, J.; Hochberg, R., Capture of Prey, Feeding, and Functional Anatomy of the Jaws in Velvet Worms (Onychophora). *Integr. Comp. Biol.* **2015,** *55*, 217–227.

(3) Baer, A.; Schmidt, S.; Haensch, S.; Eder, M.; Mayer, G.; Harrington, M. J., Mechanoresponsive Lipid-Protein Nanoglobules Facilitate Reversible Fibre Formation in Velvet Worm Slime. *Nat. Commun.* **2017,** *8*, 974.

(4) Baer, A.; Hoffmann, I.; Mahmoudi, N.; Poulhazan, A.; Harrington, M. J.; Mayer, G.; Schmidt, S.; Schneck, E., The Internal Structure of the Velvet Worm Projectile Slime: A Small-Angle Scattering Study. *Small* **2022**, 2300516.

(5) Haritos, V. S.; Niranjane, A.; Weisman, S.; Trueman, H. E.; Sriskantha, A.; Sutherland, T. D., Harnessing Disorder: Onychophorans Use Highly Unstructured Proteins, Not Silks, for Prey Capture. *Proc. Royal Soc. B* **2010,** *277*, 3255–3263.

(6) Baer, A.; Hänsch, S.; Mayer, G.; Harrington, M. J.; Schmidt, S., Reversible Supramolecular Assembly of Velvet Worm Adhesive Fibers Via Electrostatic Interactions of Charged Phosphoproteins. *Biomacromolecules* **2018,** *19*, 4034–4043.

(7) Lu, Y.; Sharma, B.; Soon, W. L.; Shi, X.; Zhao, T.; Lim, Y. T.; Sobota, R. M.; Hoon, S.; Pilloni, G.; Usadi, A., et al., Complete Sequences of the Velvet Worm Slime Proteins Reveal That Slime Formation Is Enabled by Disulfide Bonds and Intrinsically Disordered Regions. *Adv. Sci.* **2022,** *9*, e2201444.

(8) Benkendorff, K.; Beardmore, K.; Gooley, A. A.; Packer, N. H.; Tait, N. N., Characterisation of the Slime Gland Secretion from the Peripatus, Euperipatoides Kanangrensis (Onychophora: Peripatopsidae). *Comp. Biochem. Physiol. B, Biochem. Mol. Biol.* **1999,** *124*, 457–465.

(9) Baer, A.; Hoffmann, I.; Mahmoudi, N.; Poulhazan, A.; Harrington, M. J.; Mayer, G.; Schmidt, S.; Schneck, E. The Internal Structure of the Velvet Worm Projectile Slime: A Small-Angle Scattering Study *ChemRxiv* [Online], 2022.

(10) Baer, A.; Horbelt, N.; Nijemeisland, M.; Garcia, S. J.; Fratzl, P.; Schmidt, S.; Mayer, G.; Harrington, M. J., Shear-Induced B-Crystallite Unfolding in Condensed Phase Nanodroplets Promotes Fiber Formation in a Biological Adhesive. *ACS Nano* **2019,** *13*, 4992–5001.
(11) Mackey, K. R. M.; Paytan, A., Phosphorus Cycle. In *Encyclopedia of Microbiology*, 3 ed.; Schaechter, M., Ed. Academic Press: Oxford, 2009; pp 322–334.
(12) Horsman, G. P.; Zechel, D. L., Phosphonate Biochemistry. *Chem. Rev.* **2017,** *117*, 5704–5783.
(13) Acker, M.; Hogle, S. L.; Berube, P. M.; Hackl, T.; Coe, A.; Stepanauskas, R.; Chisholm, S. W.; Repeta, D. J., Phosphonate Production by Marine Microbes: Exploring New Sources and Potential Function. *Proc. Natl. Acad. Sci.* **2022,** *119*, e2113386119.
(14) Lockwood, S.; Greening, C.; Baltar, F.; Morales, S. E., Global and Seasonal Variation of Marine Phosphonate Metabolism. *ISME J.* **2022,** *16*, 2198–2212.
(15) Peck, S. C.; Gao, J.; van der Donk, W. A., Discovery and Biosynthesis of Phosphonate and Phosphinate Natural Products. *Meth. Enzymol.* **2012,** *516*, 101–123.
(16) Murienne, J.; Daniels, S. R.; Buckley, T. R.; Mayer, G.; Giribet, G., A Living Fossil Tale of Pangaean Biogeography. *Proc. Royal Soc. B* **2014,** *281*, 20132648.
(17) Oliveira, I. S.; Bai, M.; Jahn, H.; Gross, V.; Martin, C.; Hammel, Jörg U.; Zhang, W.; Mayer, G., Earliest Onychophoran in Amber Reveals Gondwanan Migration Patterns. *Curr. Biol.* **2016,** *26*, 2594-2601.
(18) Sanchez-Ruiz, J. M.; Martinez-Carrion, M., A Fourier-Transform Infrared Spectroscopic Study of the Phosphoserine Residues in Hen Egg Phosvitin and Ovalbumin. *Biochemistry* **1988,** *27*, 3338–3342.
(19) Sanchez-Ruiz, J. M.; Martinez-Carrion, M., A Fourier-Transform Infrared Spectroscopic Study of the Phosphoserine Residues in Hen Egg Phosvitin and Ovalbumin. *Biochemistry* **1988,** *27*, 3338-3342.
(20) Barja, B. C.; Herszage, J.; dos Santos Afonso, M., Iron(Iii)–Phosphonate Complexes. *Polyhedron* **2001,** *20*, 1821–1830.
(21) Nielsen, C., *Animal Evolution: Interrelationships of the Living Phyla*. 3rd ed.; Oxford University Press: 2012.
(22) Ortega-Hernández, J., Homology of Head Sclerites in Burgess Shale Euarthropods. *Curr. Biol.* **2015,** *25*, 1625–1631.
(23) Cade-Menun, B. J., Improved Peak Identification in ^{31}p-Nmr Spectra of Environmental Samples with a Standardized Method and Peak Library. *Geoderma* **2015,** *257-258*, 102–114.
(24) Bartlett, C.; Bansal, S.; Burnett, A.; Suits, M. D.; Schaefer, J.; Cegelski, L.; Horsman, G. P.; Weadge, J. T., Whole-Cell Detection of C-P Bonds in Bacteria. *Biochemistry* **2017,** *56*, 5870–5873.
(25) Bartlett, C.; Bansal, S.; Burnett, A.; Suits, M. D.; Schaefer, J.; Cegelski, L.; Horsman, G. P.; Weadge, J. T., Whole-Cell Detection of C-P Bonds in Bacteria. *Biochemistry* **2017,** *56*, 5870-5873.
(26) Rice, K.; Batul, K.; Whiteside, J.; Kelso, J.; Papinski, M.; Schmidt, E.; Pratasouskaya, A.; Wang, D.; Sullivan, R.; Bartlett, C., et al., The Predominance of Nucleotidyl Activation in Bacterial Phosphonate Biosynthesis. *Nat. Commun.* **2019,** *10*, 3698.
(27) Kang, X.; Zhao, W.; Dickwella Widanage, M. C.; Kirui, A.; Ozdenvar, U.; Wang, T., Ccmrd: A Solid-State Nmr Database for Complex Carbohydrates. *J. Biomol. NMR* **2020,** *74*, 239–245.

(28) Baumann, H.; Tzianabos, A. O.; Brisson, J. R.; Kasper, D. L.; Jennings, H. J., Structural Elucidation of Two Capsular Polysaccharides from One Strain of Bacteroides Fragilis Using High-Resolution Nmr Spectroscopy. *Biochemistry* **1992,** *31*, 4081–4089.
(29) Hensbergen, P. J.; de Ru, A. H.; Friggen, A. H.; Corver, J.; Smits, W. K.; van Veelen, P. A., New Insights into the Type a Glycan Modification of *Clostridioides Difficile* Flagellar Protein Flagellin C by Phosphoproteomics Analysis. *J. Biol. Chem.* **2022,** *298*, 101622.
(30) Paschinger, K.; Wilson, I. B. H., Anionic and Zwitterionic Moieties as Widespread Glycan Modifications in Non-Vertebrates. *Glycoconj. J.* **2020,** *37*, 27–40.
(31) Baer, A.; Mayer, G., Comparative Anatomy of Slime Glands in Onychophora (Velvet Worms). *J. Morphol.* **2012,** *273*, 1079–1088.
(32) Lilly Thankamony, A. S.; Wittmann, J. J.; Kaushik, M.; Corzilius, B., Dynamic Nuclear Polarization for Sensitivity Enhancement in Modern Solid-State Nmr. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017,** *102-103*, 120–195.
(33) Sauvée, C.; Rosay, M.; Casano, G.; Aussenac, F.; Weber, R. T.; Ouari, O.; Tordo, P., Highly Efficient, Water-Soluble Polarizing Agents for Dynamic Nuclear Polarization at High Frequency. *Angew. Chem. Int. Ed. Engl.* **2013,** *52*, 10858–10861.
(34) Mentink-Vigier, F.; Marin-Montesinos, I.; Jagtap, A. P.; Halbritter, T.; van Tol, J.; Hediger, S.; Lee, D.; Sigurdsson, S. T.; De Paëpe, G., Computationally Assisted Design of Polarizing Agents for Dynamic Nuclear Polarization Enhanced Nmr: The Asympol Family. *J. Am. Chem. Soc.* **2018,** *140*, 11013–11019.
(35) Harrabi, R.; Halbritter, T.; Aussenac, F.; Dakhlaoui, O.; van Tol, J.; Damodaran, K. K.; Lee, D.; Paul, S.; Hediger, S.; Mentink-Vigier, F., et al., Highly Efficient Polarizing Agents for Mas-Dnp of Proton-Dense Molecular Solids. *Angew. Chem. Int. Ed.* **2022,** *61*, e202114103.
(36) Lilly Thankamony, A. S.; Wittmann, J. J.; Kaushik, M.; Corzilius, B., Dynamic Nuclear Polarization for Sensitivity Enhancement in Modern Solid-State Nmr. *Prog. Nucl. Magn. Reson. Spectrosc.* **2017,** *102-103*, 120-195.
(37) Dubroca, T.; Smith, A. N.; Pike, K. J.; Froud, S.; Wylde, R.; Trociewitz, B.; McKay, J.; Mentink-Vigier, F.; van Tol, J.; Wi, S., et al., A Quasi-Optical and Corrugated Waveguide Microwave Transmission System for Simultaneous Dynamic Nuclear Polarization Nmr on Two Separate 14.1 T Spectrometers. *J. Magn. Res.* **2018,** *289*, 35–44.
(38) Morcombe, C. R.; Zilm, K. W., Chemical Shift Referencing in Mas Solid State Nmr. *J. Magn. Reson.* **2003,** *162*, 479–486.
(39) Wishart, D. S.; Bigam, C. G.; Yao, J.; Abildgaard, F.; Dyson, H. J.; Oldfield, E.; Markley, J. L.; Sykes, B. D., ^{1}h, ^{13}c and ^{15}n Chemical Shift Referencing in Biomolecular Nmr. *J. Biomol. NMR* **1995,** *6*, 135–140.
(40) Furness, J. W.; Kaplan, A. D.; Ning, J.; Perdew, J. P.; Sun, J., Accurate and Numerically Efficient R2scan Meta-Generalized Gradient Approximation. *J. Phys. Chem.* **2020,** *11*, 8208–8215.
(41) Grimme, S.; Hansen, A.; Ehlert, S.; Mewes, J.-M., R2scan-3c: A “Swiss Army Knife” Composite Electronic-Structure Method. *J. Chem. Phys.* **2021,** *154*, 064103.
(42) Neese, F., The Orca Program System. *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012,** *2*, 73–78.
(43) Robson, E. A.; Lockwood, A. P. M.; Ralph, R., Composition of the Blood in Onychophora. *Nature* **1966,** *209*, 533–533.

(44) Bolger, A. M.; Lohse, M.; Usadel, B., Trimmomatic: A Flexible Trimmer for Illumina Sequence Data. *Bioinformatics* **2014,** *30*, 2114–2120.
(45) Peng, Y.; Leung, H. C.; Yiu, S. M.; Lv, M. J.; Zhu, X. G.; Chin, F. Y., Idba-Tran: A More Robust De Novo De Bruijn Graph Assembler for Transcriptomes with Uneven Expression Levels. *Bioinformatics* **2013,** *29*, i326–334.
(46) Seppey, M.; Manni, M.; Zdobnov, E. M., Busco: Assessing Genome Assembly and Annotation Completeness. *Methods Mol. Biol.* **2019,** *1962*, 227–245.
(47) Altschul, S. F.; Madden, T. L.; Schäffer, A. A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D. J., Gapped Blast and Psi-Blast: A New Generation of Protein Database Search Programs. *Nucleic Acids Res.* **1997,** *25*, 3389–3402.
(48) Zheng, Y. Z.; DeMarco, M. L., Manipulating Trypsin Digestion Conditions to Accelerate Proteolysis and Simplify Digestion Workflows in Development of Protein Mass Spectrometric Assays for the Clinical Laboratory. *Clin. Mass Spectrom.* **2017,** *6*, 1–12.
(49) Shekhawat, L. K.; Tiwari, A.; Yamamoto, S.; Rathore, A. S., An Accelerated Approach for Mechanistic Model Based Prediction of Linear Gradient Elution Ion-Exchange Chromatography of Proteins. *J. Chromatogr. A.* **2022,** *1680*, 463423.
(50) Masuko, T.; Minami, A.; Iwasaki, N.; Majima, T.; Nishimura, S.; Lee, Y. C., Carbohydrate Analysis by a Phenol-Sulfuric Acid Method in Microplate Format. *Anal. Biochem.* **2005,** *339*, 69–72.

Acknowledgments:

We thank D. M. Rowell, I. S. Oliveira, C. Martin, and I. Schumann for help with specimen collection, S. Treffkorn for assistance with RNA extractions and I. S. Oliveira for photography of velvet worms. Collecting and export permits were kindly provided by the Office of Environment & Heritage (NSW National Parks & Wildlife Service), the Department of the Environment of Australia, and the National System of Conservation Areas (SINAC, MINAE) for *Eu. rowelli* and the Ministry of Environment and National Beautification, Green and Blue Economy of Barbados for *Ep. barbadensis*. This work was supported bythe German Research Foundation (DFG: grant MA 4147/7-1 and 4147/2 to G.M., and grants SCHM 2748/5-1 and SFB 120-project Z02 to S.S.), the Natural Sciences and Engineering Research Council of Canada (grant RGPIN-2018-05243 to M.J.H. and grant RGPIN-2018-06200 to I.M.), aCanada Research Chair award (CRC Tier grant 2 950-231953 to M.J.H.), the Fonds de Recherche du Quebec – Nature et Technologies (B2X doctoral research scholarship 275130 and international internship scholarship 293818 to A.P.), the National Institutes of Health (NIH)-funded R24 grant (R24GM137782 to P.A.), GlycoMIP, a National Science Foundation Materials Innovation Platform funded through Cooperative Agreement (DMR-1933525 to P.A.), a European Union's grant

agreement No 10100085000 (PANACEA), the National High Magnetic Field Laboratory is supported by the National Science Foundation Cooperative Agreement No. DMR-1644779 and the State of Florida and the MAS-DNP instrument and probe development is supported by the NIH GM122698.