biomolecules

MDPI

Article

Identification, Synthesis, Conformation and Activity of an Insulin-like Peptide from a Sea Anemone

Michela L. Mitchell [1,2,3,*], Mohammed Akhter Hossain [4,5], Feng Lin [4], Ernesto L. Pinheiro-Junior [6], Steve Peigneur [6], Dorothy C. C. Wai [1], Carlie Delaine [7], Andrew J. Blyth [7], Briony E. Forbes [7], Jan Tytgat [6], John D. Wade [4,5] and Raymond S. Norton [1,8,*]

1 Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, 381 Royal Parade, Parkville, VIC 3052, Australia; dorothy.wai@monash.edu
2 Sciences Department, Museum Victoria, G.P.O. Box 666, Melbourne, VIC 3001, Australia
3 Biodiversity and Geosciences, Queensland Museum, P.O. Box 3000, South Brisbane, QLD 4101, Australia
4 Florey Institute of Neuroscience and Mental Health, University of Melbourne, Parkville, VIC 3010, Australia; akhter.hossain@unimelb.edu.au (M.A.H.); feng.lin@florey.edu.au (F.L.); john.wade@florey.edu.au (J.D.W.)
5 School of Chemistry, University of Melbourne, Parkville, VIC 3010, Australia
6 Toxicology and Pharmacology, University of Leuven, O&N 2, Herestraat 49, P.O. Box 922, 3000 Leuven, Belgium; ernesto.lopes@kuleuven.be (E.L.P.-J.); steve.peigneur@kuleuven.be (S.P.); jan.tytgat@kuleuven.be (J.T.)
7 Flinders Health and Medical Research Institute, Flinders University, Bedford Park, SA 5042, Australia; carlie.delaine@flinders.edu.au (C.D.); andrew.blyth@flinders.edu.au (A.J.B.); briony.forbes@flinders.edu.au (B.E.F.)
8 ARC Centre for Fragment-Based Design, Monash University, Parkville, VIC 3052, Australia
* Correspondence: michela.mitchell@qm.qld.gov.au (M.L.M.); ray.norton@monash.edu (R.S.N.)

Abstract: The role of insulin and insulin-like peptides (ILPs) in vertebrate animals is well studied. Numerous ILPs are also found in invertebrates, although there is uncertainty as to the function and role of many of these peptides. We have identified transcripts with similarity to the insulin family in the tentacle transcriptomes of the sea anemone *Oulactis* sp. (Actiniaria: Actiniidae). The translated transcripts showed that these insulin-like peptides have highly conserved A- and B-chains among individuals of this species, as well as other Anthozoa. An *Oulactis* sp. ILP sequence (IlO1_i1) was synthesized using Fmoc solid-phase peptide synthesis of the individual chains, followed by regioselective disulfide bond formation of the intra-A and two interchain disulfide bonds. Bioactivity studies of IlO1_i1 were conducted on human insulin and insulin-like growth factor receptors, and on voltage-gated potassium, sodium, and calcium channels. IlO1_i1 did not bind to the insulin or insulin-like growth factor receptors, but showed weak activity against K_V1.2, 1.3, 3.1, and 11.1 (hERG) channels, as well as Na_V1.4 channels. Further functional studies are required to determine the role of this peptide in the sea anemone.

Keywords: *Oulactis*; cnidaria; peptide synthesis; ion channel; invertebrates; insulin

check for updates

Citation: Mitchell, M.L.; Hossain, M.A.; Lin, F.; Pinheiro-Junior, E.L.; Peigneur, S.; Wai, D.C.C.; Delaine, C.; Blyth, A.J.; Forbes, B.E.; Tytgat, J.; et al. Identification, Synthesis, Conformation and Activity of an Insulin-like Peptide from a Sea Anemone. *Biomolecules* **2021**, *11*, 1785. https://doi.org/10.3390/biom11121785

Academic Editors: Vladimir N. Uversky and Myron R. Szewczuk

Received: 27 September 2021
Accepted: 23 November 2021
Published: 29 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

The role of insulin as a key anabolic hormone that promotes the absorption of glucose from the blood into liver, fat, and skeletal muscle cells in vertebrates is well documented [1–3]. In contrast, the presence and role of insulin-like peptides (ILPs) in invertebrates, especially those found in the marine environment, are less well understood. The characterized invertebrate ILPs perform an array of functions, which vary depending upon their tissue of origin. Insect ILPs (e.g., from *Drosophila melanogaster*) are implicated in multiple functions, including regulating metabolism, growth, reproduction, and longevity [4–7]. For example, bombyxin A-1, a silk moth (*Bombyx mori*) ILP, is produced in the brain to activate the prothoracic glands to produce ecdysone (a moulting hormone) for growth development [8]. The nematode worm *Caenorhabditis elegans* ILPs act as agonists

and antagonists towards the insulin-like growth factor signaling pathway [9]. Marine ILPs have been documented in molluscs; in the Californian sea hare *Aplysia californica* ILPs have a role in metabolism [10]. In the Pacific oyster (*Crassostrea gigas*), ILPs are involved in sexual maturation [11]. ILPs have also been identified in the purple sea urchin (*Strongylocentrotus purpratus*) and the Eastern rock lobster (*Sagmariasus verreauxi*) [12,13].

More recently, marine invertebrate ILPs have been identified in the venom of two cone snails, *Conus geographus* and *C. tulipus* [14]. The snails release the ILPs into the water column amongst schools of fish. The fish succumb to a hypoglycemic stupor referred to as a "nirvana cabal", whereupon the snail extends its mouth like a fishing net to capture the fish prior to venom injection [14–16]. This intriguing discovery of an ILP in cone snail venom with a role in prey capture has stimulated the search for additional functions and biological applications of ILPs in invertebrates.

One motivation for investigating marine ILPs is their potential anti-diabetic properties which may be translatable into pharmaceuticals [16]. Several marine organisms have been documented to contain anti-diabetic compounds, including sea anemones [17]. Pascual et al. [18] demonstrated that aqueous crude extracts from the sea anemones *Bunodosoma granuliferum* and *Bartholomea annulata* inhibit porcine dipeptidyl peptidase IV (DPPIV). Inhibition of this enzyme results in a reduction of glucagon and blood glucose levels and may therefore be useful as an alternative treatment for diabetes. The near-explosive growth in genomic information has led to the identification of large numbers of insulin-like peptides (ILPs) in many species of mammals, insects and viruses [19–21]. As an example, *Caenorhabditis elegans* is predicted to have, remarkably, no less than 40 different insulins [22]. Many of these peptides share the characteristic two-chain (A and B) and canonical three-disulfide-bond structure of human insulin. However, some possess single-chain structures containing three disulfide bonds, and some do not include an A-chain intramolecular disulfide. Other ILPs contain an additional interchain disulfide bond [23]. Depending on their general structure and number of disulfide bonds, they have been classified as α-type, β-type or γ-type [22,23].

The rapid growth in 'omics' studies includes the discovery of numerous ILPs in the genomes and transcriptomes of sea anemones, including for the model sea anemone *Nematostella vectensis* [24,25]. However, sea anemone ILPs have not been characterized functionally. Unlike other animals, cnidarians lack a circulatory and central nervous system. Therefore, the processing and physiological role(s) of ILPs in cnidarians are unclear and their functions are still to be determined, along with their potential as a source of therapeutics to treat diabetes.

Several insulin-like sequences were identified in the tentacle transcriptomes of the Australian speckled anemone (*Oulactis* sp.) [26], which may represent a new class of ILP. In this study we describe the synthesis of a representative ILP from this sea anemone and the results of activity assays against human insulin and insulin-like growth factor receptors, as well as a range of ion channels.

2. Materials and Methods

2.1. Insulin-like Peptide Data and Bioinformatic Analysis

Translated amino acid sequences that share domain similarity within the insulin family (protein family PF00049) were identified during initial bioinformatic analysis of tentacle transcriptomes from the sea anemone *Oulactis* sp. (not yet formally described) [26]. These ILP sequences were deposited in the European Bioinformatics Institute—European Nucleotide Archive (EMBL-ENA) database under Project No. PRJEB34263 (Acc. No. LR700308-11). Following the initial identification, the search for ILPs was extended to two additional tentacle transcriptomes to determine whether the identified ILPs were present in other individuals of the species.

The search for ILPs was also extended to additional discrete tissue regions of *Oulactis* sp., including internal tissue (mesenterial filaments and gametic material) and external tissue (the frill and acrorhagi, located adjacent to the tentacles). Translated amino acid

sequences were selected to match the search criteria of similarity to the insulin protein family PF00049 domain with an *E*-value: >1 × 10^{-10}; full-length sequence i.e., including start and stop codons; and the presence of a signal peptide. The full-length translated amino acid sequence of insulin-like_Oulsp_1_i1 (simplified to IlO1_i1 for this study) (ENA-EMBL Acc. No. LR700308) formed the search basis in the NCBI non-redundant database (https://blast.ncbi.nlm.nih.gov/, accessed on May 2020) for other sequences with similarity. The NCBI-BLAST search was conducted online using the blastp algorithm (protein–protein BLAST).

Selected sequences of characterized ILPs were downloaded from Uniprot (https://www.uniprot.org/, accessed on May 2020) for sequence-function comparison. Sequences were aligned using Clustal OMEGA online [27–29], and pairwise sequence alignments were conducted via the EMBL-ENA EMBOSS needle tool using the Needleman–Wunsch algorithm [30]. SignalP (4.1 server) [31] was used to predict the signal peptide and cleavage position for mature protein sequences. B-, C-, and A-chain cleavage sites were determined as per documented cleavage positions for human insulin (Lys-Arg, Arg-Arg) [32].

2.2. ILP Synthesis

The sequence IlO1_i1 was selected for synthesis and subsequent functional characterization. For the purposes of the assembly, the N-terminal B-chain hexapeptide containing a pair of Cys residues was deleted, and the predicted canonical insulin Cys bonds were adopted (Figure 1). The deletion was undertaken for consistency with previously characterized insulin and ILP sequences, along with the regions of amino acid sequences known to be active on mammalian insulin receptors [16,33].

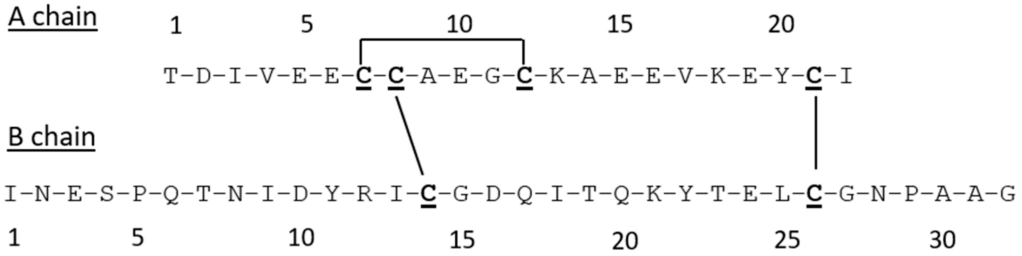


Figure 1. Truncated IlO1_i1 amino acid sequence synthesized using Fmoc solid-phase peptide synthesis chemistry. Cysteines are bolded and underlined, predicted disulfide bond connectivities are shown.

Each of the selectively S-protected A- and B-chains was assembled by Fmoc solid-phase peptide synthesis. For the A-chain peptide, S-protection was afforded by Trt (Cys7,12), Acm (Cys21), and *t*-Bu (Cys8). For the B-chain, Trt (Cys14) and Acm (Cys26) were used.

A-chain intramolecular disulfide oxidation: [Cys7,12 (S-thiol), Cys8(*t*-Bu), Cys21(Acm)] A-chain (66 mg, 25.5 µmol; Figure S4) was dissolved in 0.1 M Gly.NaOH (40.5 mL), and to this was added 1 mM 2-dipyridyl disulfide (DPDS) in MeOH (25.5 mL, 25.5 mmol). Oxidation was complete after 1 h, as monitored by analytical RP-HPLC. The peptide was isolated by preparative RP-HPLC and subsequent freeze drying to give 45.3 mg (44.0%) of purified [Cys21(Acm), Cys8(But)] A-chain (Figure S5).

[Cys8(Pyr), Cys21(Acm)] A-chain: Intramolecular disulfide-bonded [Cys8(But), Cys21(Acm)] A-chain (45.3 mg, 17.5 mmol) was converted to the Cys8 S-pyridinylsulfenyl (Pyr) form by treatment with DPDS in neat TFA (5.0 mL) containing anisole (0.5 mL) chilled to 0 °C, and then 5.0 mL TFMSA/TFA (1:5 *v*/*v*) was added and stirred for 20–30 min, maintaining the temperature at or below 0 °C. The peptide was then precipitated in ether and the pellet was suspended in 6 M GdnHCl for purification. The target peptide was isolated by preparative RP-HPLC to give 28.3 mg (61.1%) (Figure S6).

Combination of [Cys8(Pyr), Cys14(Acm)] A-chain with [Cys14(S-thiol), Cys26(Acm)] B-chain: A-chain peptide (25.9 mg, 9.6 mmol) was dissolved in 8 M GdnHCl (6 mL)

and added to purified B-chain (34.5 mg, 9.6 mmol; Figure S7) in the same buffer (6 mL). The mixture was stirred vigorously at room temperature or 37 °C for each buffer, respectively, and the reaction was monitored by analytical RP-HPLC. After 30 min (or 24 h if using the GdnHCl buffer), the reaction was terminated by the addition of glacial acetic acid, and the target product was isolated by preparative RP-HPLC to give 22.7 mg (38.7%) (Figure S8).

The [Cys21(Acm)] A-chain/[Cys26(Acm)] B-chain (22.7 mg, 3.7 mmol) was dissolved in glacial acetic acid (16.7 mL) and 80 mM HCl (2.3 mL), and to this was added dropwise 3.7 mL of 20 mM iodine/acetic acid (74 mmol). After 1 h [34], the reaction was stopped by the addition of 3.7 mL of 20 mM ascorbic acid. Preparative RP-HPLC, as described above, was then used to isolate and purify the product, giving 15.6 mg (70.4%) in an overall total yield of 27.2%; the expected mass of the final product was calculated at [MH+] 5972.62, and that observed was [M/4+H] = 1493.91, [M/5+H] = 1195.33 (Figure S9).

2.3. Spectroscopic Studies

2.3.1. Circular Dichroism Spectroscopy

Circular dichroism (CD) spectra were recorded on a Jasco J-1500 CD spectrometer from 300 to 180 nm with a 0.2 nm step size using a 1.0 s response time and 1.0 nm bandwidth in a quartz cuvette with a 0.2 cm path length. Insulin and IlO1_i1 peptide were resuspended in 10 mM sodium phosphate (pH 7.8) to a concentration of 0.2 mg/mL. To correct for background, the spectrum of buffer alone was subtracted from each sample spectrum. The machine units collected—θ in millidegrees, were converted to mean residue ellipticity (MRE), [θ] (degrees·cm^{2}dmol^{-1}residue^{-1}). Helical content was calculated using the CDSSTR algorithm [35] for deconvolution against the reference protein database set SMP180. This program is available on the DICROWEB website (http://dichroweb.cryst.bbk.ac.uk/html/home.shtml, accessed on May 2020).

2.3.2. Nuclear Magnetic Resonance (NMR) Spectroscopy

The NMR spectra of IlO1_i1 and Con-Ins G1 were acquired at 298 K on a Bruker Avance III spectrometer equipped with a TCI cryoprobe (Billerica, MA, USA). Quantities of 0.8 mg of IlO1_i1 and 1 mg of Con-Ins G1 were each dissolved in 500 µL H_2O, and the pH was measured without adjustment before the addition of D_2O (to a final concentration of 10%). ^{1}H chemical shifts were referenced using dioxane (3.75 ppm), added after acquisition of the initial spectrum. Spectra were processed using Topspin (version 3.6.2).

2.4. Functional Studies

2.4.1. Binding Assays

The ability to bind IlO1_i1 to the human insulin receptor (IR) and insulin-like growth factor receptor (IGF-1R) was measured in competition binding assays. BALB/c3T3 fibroblast cells overexpressing IGF-IR (P6 cells) [36] and IR-B cells (IGF-IR null mouse fibroblasts overexpressing the human insulin receptor isoform B) [37] were cultured in DMEM, 10% fetal calf serum, 1% penicillin/streptomycin, and G418 (250 µg/mL). IGF-IR and IR-B were solubilized from cells using lysis buffer (20 mM HEPES, 150 mM NaCl, 1.5 mM $MgCl_2$, 10% (*v*/*v*) glycerol, 1% (*v*/*v*) Triton X-100, 1 mM EGTA (pH 7.5)) for 1 h at 4 °C, and lysates were centrifuged for 10 min at 3500 rpm. Solubilized IGF-1R or IR-B (100 µL) was used to coat each well of a white Greiner Lumitrac 600 plate previously coated with 24–31 anti-human IGF-1R antibody or 83-7 anti-human IR antibody [38,39]. Europium-labelled IGF-I or insulin (~3,000,000 counts) was added to wells with increasing concentrations of competitive ligand (IGF-I, insulin, or IlO1_i1 peptide) and incubated for 16 h at 4 °C. Wells were washed three times with 20 mM Tris, 150 mM NaCl, 0.1% (*v*/*v*) Tween 20, and then DELFIA enhancement solution (100 µL) was added. Time-resolved fluorescence was measured with 340 nm excitation and 612 nm emission filters using a Victor X4, 2030 Multilabel Reader (Perkin Elmer). All assays were repeated at least 3 times with 3 technical replicates each, except for the IlO1_i1 peptide binding IGF-IR, where

2 assays were conducted. Mean IC_{50} values were calculated using the statistical software package Prism v9.0.0 (GraphPad Software) after curve fitting with a nonlinear regression (one-site) model.

2.4.2. Ion Channel Assays

For the expression of K_V channels (mammalian rK_V1.2, hK_V1.3, rK_V2.1, hK_V3.1, rK_V4.2, hK_V7.2/hK_V7.3, hK_V10.1, hERG and Shaker IR from the fruit fly *Drosophila melanogaster*), Na_V channels (mammalian rNa_V1.2, rNa_V1.3, rNa_V1.4, hNa_V1.5, rNa_V1.6, hNa_V1.7, BgNa_V from the cockroach *Blattella germanica,* as well as the auxiliary subunits rβ1, hβ1), and the hCa_V3.1 channel in *Xenopus* oocytes, the linearized plasmids were transcribed using the T7 or SP6 mMESSAGE-mMACHINE transcription kit (Ambion, Carlsbad, CA, USA). The harvesting of stage V–VI oocytes from an anaesthetized female *Xenopus laevis* frog was described previously [40] and was in compliance with the regulations of the European Union (EU) concerning the welfare of laboratory animals as declared in Directive 2010/63/EU. The use of *X. laevis* oocytes was approved by the Animal Ethics Committee of the KU Leuven with the licence number P186/2019. Oocytes were injected using a microinjector (Drummond Scientific, Broomall, PA, USA) with 4–50 nL of cRNA, depending on the channel subtype, and subsequently incubated in ND96 solution (96 mM NaCl, 2 mM KCl, 1.8 mM $CaCl_2$, 2 mM $MgCl_2$, and 5 mM HEPES, pH 7.4), supplemented with 50 mg/L gentamycin sulfate.

Two-electrode voltage-clamp recordings were performed at room temperature (18–22 °C) using a Geneclamp 500 amplifier (Molecular Devices, Downingtown, PA, USA) controlled by a pClamp data acquisition system (Axon Instruments, Union City, CA, USA). Whole cell currents from oocytes were recorded 1–7 days after injection. The bath and perfusion solutions were either the previously described ND96 (Na_V and K_V channels) or calcium-free ND96 supplemented with 10 mM $BaCl_2$ (Ca_V channels). Toxins were applied directly to the bath. The resistances of both electrodes were kept between 0.8 and 1.5 MΩ. Currents were sampled at 20 kHz (Na_V channels) and 2 kHz (K_V and Ca_V channels) and filtered using a four-pole low-pass Bessel filter, at 1 kHz for sodium and 500 MHz for potassium and calcium channels, except for hERG, in which the currents were filtered at 1 kHz. Leak subtraction was performed using a P/4 protocol. K_V1.x currents were evoked by 500 ms depolarizations to 0 mV followed by a 500 ms pulse to −50 mV, from a holding potential of −90 mV. K_V2.1, K_V3.1, and K_V4.2 currents were elicited by 500 ms pulses to +20 mV from a holding potential of −90 mV. Current traces of the hERG channel were elicited by applying a +40 mV pre-pulse for 2 s, followed by a step of −120 mV for 2 s. Current traces of K_V10.1 were elicited by 2 s depolarization to 0 mV, from a holding potential of −90 mV. Sodium current traces were evoked by a 100 ms depolarization to 0 mV. For Ca_V channels, current traces were elicited by 700 ms depolarizations to −20 mV from a holding potential of −90 mV.

3. Results

3.1. ILP Sequence Similarity

Several transcripts identified in the tentacle transcriptomes of three individuals of the speckled sea anemone (*Oulactis* sp.) [26] bore sequence similarity to the insulin family (Pfam PF00049) [41]. Analysis of individual transcripts showed that the A-chain is wholly conserved across individuals of the species (Supplementary Material, Figure S1). The B-chain is also highly conserved, varying at only one or two residues, including the penultimate residue (Supplementary Material, Figure S1).

The representative ILP full-length sequence IlO1_i1 (Figure 2a) from *Oulactis* sp. was selected for additional bioinformatic and functional studies. IlO1_i1 is composed of a signal peptide and B-, C-, and A-chains, with the full-length sequence having an *E*-value similarity of 1.7×10^{-6} to the insulin family. An NCBI-Blastp search of IlO1_i1 returned hits against sequences from six cnidarian genomes, including the sea anemones *Actinia tenebrosa* [42], *Exaiptasia pallida* [43] and *Nematostella vectensis* [44], along with three stony

corals, *Acropora millepora* [45], *Stylophora pistallata* [46] and *Pocillopora damicornis* [47]. None of the six sequences returned in the NCBI blastp search has been functionally characterized (Supplementary Material, Table S1 and Figure S2).

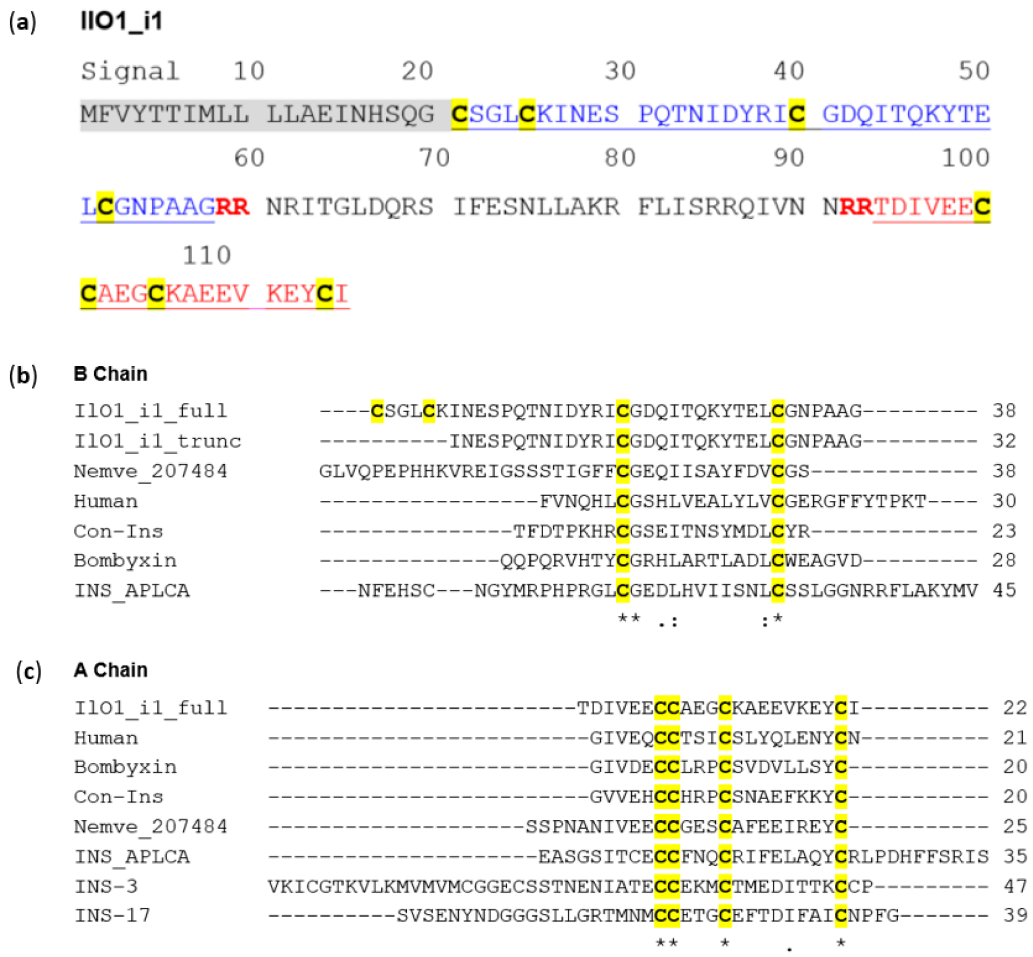


Figure 2. The full-length sequence of the sea anemone peptide ILP IlO1_i1 and alignment to characterized human and invertebrate insulin-like peptides (ILPs). (**a**) The full-length amino acid sequence of the sea anemone ILP. Signal peptide highlighted in grey, B-chain in blue, C-chain in black and A-chain in red. (**b**) Alignment of human and invertebrate B-chains to the predicted full-length B-chain of IlO1_i1 and the truncated sequence used for synthesis. (**c**) A-chain alignment of IlO1_i1, human and invertebrate ILPs. Sequences used in alignments with UniProt number supplied: Human insulin: P01308 [33]; Con-Ins: *Conus geographus*: A0A0B5AC95 [14]; Bombyxin: *Bombyx mori:* Q17192 [8]; Nemve_207484: *Nematostella vectensis*: A7S6C3 [44]; INS_APLCA: *Aplysia californica*: Q9NDE7 [10]; INS-3: *Caenorhabditis elegans*: Q09628 [9]: INS-17: *Caenorhabditis elegans*: G5EFH1 [9]. Cysteines highlighted in yellow and conserved residues bolded. Sequence alignments performed in Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/, accessed on May 2020) [27–29].

The aligned cnidarian sequences showed highly conserved residues in the A chain and more variability in the B chain. Four of the cnidarian sequences, including IlO1_i1, contained an additional cysteine pairing in the N terminus of the B chain (Supplementary Material Figure S2).

The predicted A- and B-chains of IlO1_i1 were aligned (Figure 2b,c) against human insulin and various characterized invertebrate ILPs, with the exception of the *Nematostella vectensis*, which is a predicted peptide. IlO1_i1 shares the same conserved cysteine arrangement reported previously in studies of ILPs [5]. The A- and B-chains share a highly conserved cysteine framework ($CCx_{[4]}Cx_{[8]}C$ and $Cx_{[11]}C$, respectively) with the exception

of an additional two cysteines at the N-terminus of the B-chain in IlO1_i1. In addition, the Gly after Cys_1 is wholly conserved in the B-chain.

The sequence similarity of human insulin and characterized invertebrate ILPs to IlO1_i1 was examined further using pairwise alignment (Table 1). As expected, IlO1_i1 has a much higher sequence similarity to cnidarian ILPs than to other invertebrate ILPs, with the lowest sequence similarity to *Caenorhabditis elegans* (INS-3 and INS-17). INS-3 and INS-17 are restricted in similarity only at the C-terminus of the A-chain, as opposed to INS_APLCA (Mollusc ILP), where the similarity occurs at the N-terminus of the A-chain (Table 1). Additionally, *Caenorhabditis elegans* sequences INS-3 and INS-17 do not possess B-chains and so were excluded from the subsequent analysis and alignment. Based on the percentage similarity of sequences alone, it is not possible to infer a function for IlO1_i1, as the most similar (characterized) ILPs are diverse in function, with roles in energy metabolism, inducement of a hypoglycemic coma, or growth development.

Table 1. Pairwise similarity of the A- and B-chains of the sea anemone insulin-like protein (IlO1_i1) to human insulin and select characterized invertebrate ILPs, and tissue expression. Uniprot code supplied in parentheses under the ILP name in addition to the WormBase ID for *Caenorhabditis elegans*.

Insulin-like Protein (ILP)	Phylum/Class Species	Function/ Tissue Expression	A-Chain% Similarity	B-Chain% Similarity	Ref.
Nemve_207484 (A7S6C3)	Cnidaria/Anthozoa *Nematostella vectensis*	predicted ILP (larval whole animal)	61.5	40.9	[25,44]
Insulin (P010308)	Chordata/Mammalia *Homo sapiens*	metabolize energy (pancreas)	54.5	24.4	[33]
Con-Ins 1A (A0A0B5AC95)	Mollusca/Gastropoda *Conus geographus*	induce insulin coma (venom)	50.0	34.2	[14]
Bombyxin (Q17192)	Arthropoda/Insecta *Bombyx mori*	growth development (brain)	50.0	32.5	[8]
INS_APLCA (Q9NDE7)	Mollusca/Gastropoda *Aplysia californica*	metabolize energy (ganglia and cellular clusters)	31.4	28.8	[10]
INS-3 (Q09628; WBGene00002086)	Nematoda *Caenorhabditis elegans*	antagonist of the IIS [1] pathway (coelomocyte; egg-laying apparatus; gonad; head muscle; and nervous system)	23.4	-	[9,48]
INS-17 (G5EFH1; WBGene0000210)	Nematoda *Caenorhabditis elegans*	agonist of the IIS [1] pathway (egg-laying apparatus; gonad; head muscle; neurons; and somatic nervous system.)	23.1	-	[9,48]

[1] insulin/insulin-like growth factor signaling.

To ascertain the distribution and isoform variability of ILPs in *Oulactis* sp. and possible tissue-specific biological functions, the bioinformatic search for ILPs was extended to additional *Oulactis* sp. transcriptomes from discrete morphological regions, including mesenterial filaments and gametes, and the frill and acrorhagi. These morphological regions perform different biological functions from those of the tentacles, which are used primarily in defense and prey capture by the sea anemone [49,50]. It has been shown that venom expression varies between tentacles and other discrete regions in sea anemones [51].

Amino acid sequences meeting the bioinformatic search parameters were identified, and data were extracted. The mesenterial filament and gametic tissue returned only partial sequences so were excluded from further analysis. The frill and acrorhagi contained two unique ILPs. The tentacle transcriptomes from an additional two individuals contained seven unique sequences. These sequences did not match the amino acid sequence of IlO1_i1, but they all shared the highly conserved residues in the A-chain, while the B- and C-chains had minimal similarity. Additional sequences (ENA-EMBL Acc. No. OU729069-77) and alignments against IlO1_i1 may be found in the Supplementary Material (Figure S3). Two sequences were found in two individuals' transcriptomes (tentacles, frill, and acrorhagi), indicating that there is some conservation of ILPs between tissues and individuals. Although no identical, translated ILP sequence was found in all five individuals, illustrating a high variability in the gene family among individuals.

3.2. Synthesis and Characterization

A truncated analogue of IlO1_i1 in which the additional cysteine pairing (Figure 2a) was omitted in order to mimic known active insulin and ILP sequences was synthesized by Fmoc solid-phase peptide synthesis [52] of the individual chains and subsequent regioselective disulfide bond formation [53,54]. Briefly, the two chains were separately chemically assembled using selective thiol protection on the Cys residues, and, after purification, sequential S-deprotection and simultaneous disulfide bond formation afforded the target peptide in an overall purified yield of 24%. Chemical characterization by RP-HPLC and MALDI-TOF MS confirmed the expected purity and composition of the peptide (Supplementary Materials, Figure S9).

3.3. Spectroscopic Studies

The conformation of the synthetic peptide was assessed by ^{1}H nuclear magnetic resonance (NMR) and circular dichroism (CD) spectroscopy. As shown in Figure 3, synthetic IlO1_i1 has a moderately well-dispersed ^{1}H NMR spectrum, consistent with a folded structure [55,56]. Indeed, the spectral dispersion and resonance linewidths are comparable to those of Con-Ins G1, a representative vertebrate insulin from the cone snail *Conus geographus* [14], which is active on the human insulin receptor but monomeric in aqueous solution [57].

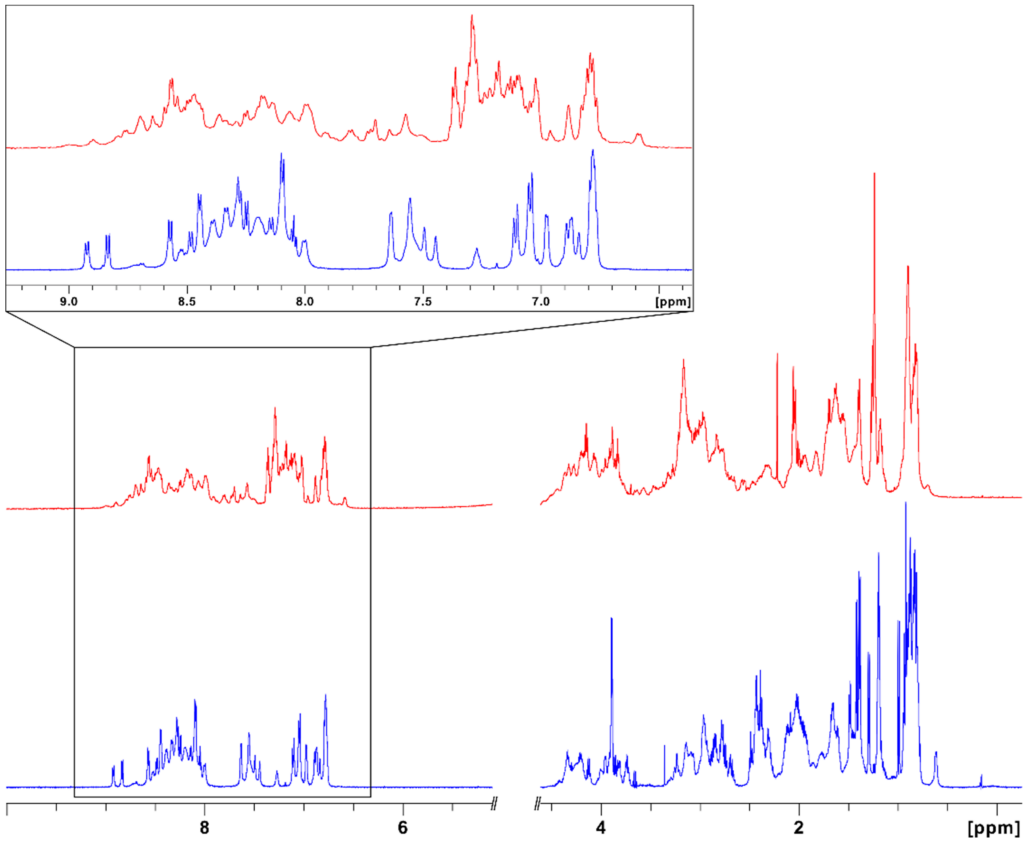
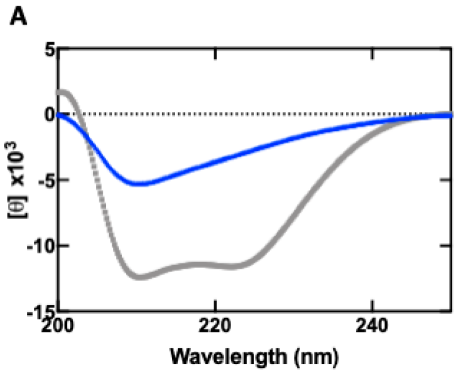


Figure 3. One-dimensional ^{1}H NMR spectrum of the synthetic *Oulactis* sp. insulin-like peptide IlO1_i1 (blue). The spectrum was acquired at 298 K and pH 3.6 in 90% H_2O/10% D_2O on a Bruker Avance III 600 MHz spectrometer. The spectrum in red is from synthetic Con-Ins G1 [14], acquired at 298 K and pH 3.5. Expanded views of the amide-aromatic regions are shown in the inset.

CD spectra (Figure 4), however, indicated that IlO1_i1 does not display the characteristic helical content of insulin (36% overall, Figure 4B) and is largely disordered, even though it has the canonical insulin disulfide connectivities.



B

Analogue	%Helix	%Strand	%Turns	%Unordered
Insulin	36	17	15	33
sINSL	7	41	12	40

Figure 4. (**A**) CD spectra of insulin (grey) and IlO1_i1 peptide (blue). The percentages of helical content shown in the table (**B**) were calculated using the CDSSTR algorithm for deconvolution against the reference protein database set SMP180.

3.4. Receptor Binding Assays

Since previously characterized ILPs have been shown to bind to the insulin and/or IGF-1 receptors [16,58,59], the ability of IlO1_i1 to bind human IR (isoform B, IR-B) and IGF-1R was measured in competition binding assays using europium-labelled insulin and IGF-1, respectively (Figure 5). Insulin bound with high affinity to IR-B (IC_{50} = 1 nM), as reported previously [37]. However, even at 10^{-5} M concentration, IlO1_i1 was unable to compete with europium-labelled insulin bound to IR-B (Figure 5b). Similarly, both IGF-1 and, to a lesser extent, insulin were able to effectively compete for binding, as expected (IC_{50} 0.22 nM and 3.7 nM, respectively) [37]. However, IlO1_i1 did not compete with europium-labelled IGF-I binding to the IGF-1R (Figure 5a).

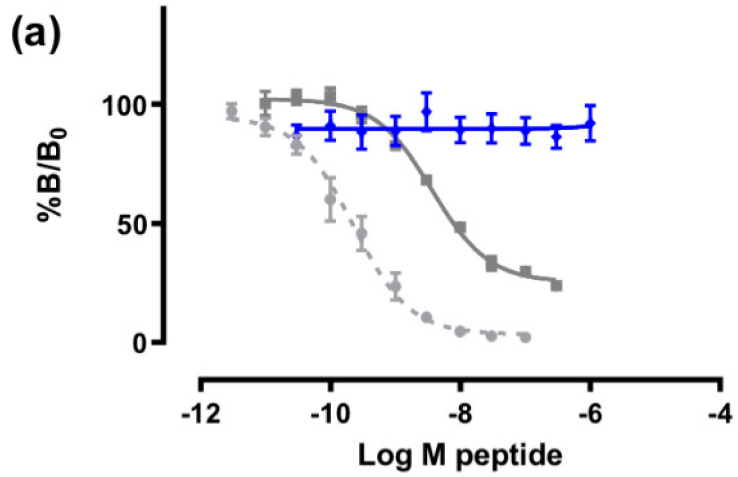


Figure 5. *Cont.*

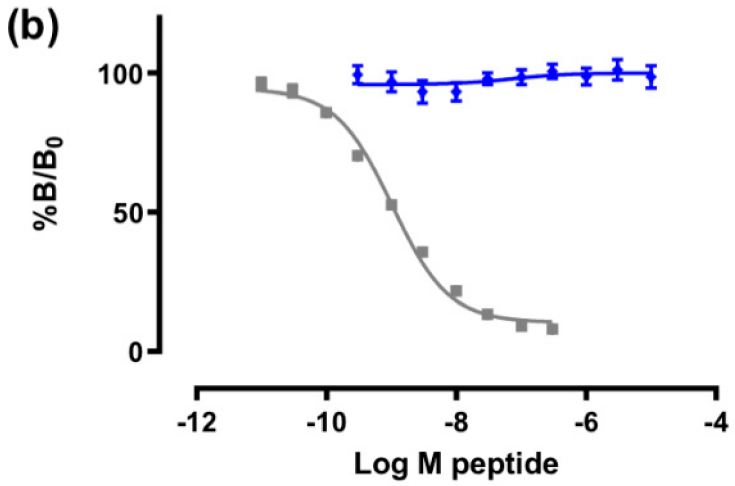


Figure 5. The ability of IlO1_i1 to bind to (**a**) IGF-1R and (**b**) IR-B was measured in competition binding assays using europium-labelled insulin for IR-B and europium-labelled IGF-1 for IGF-1R assays. Increasing concentrations of IlO1_i1 (blue), insulin (grey), and IGF-1 (grey dashed) were added. The results are expressed as a percentage of binding in the absence of competing ligand (%B/B_0). Data shown are the mean ± S.E., with error bars shown where greater than the size of the symbols. n ≥ 3 independent experiments were conducted, each with triplicate technical replicates, except for IlO1_i1 binding IGF-1R, where n = 2.

3.5. Potassium, Sodium, and Calcium Channel Assays

As many disulfide-rich peptides from sea anemones and other marine organisms target ion channels [60,61] IlO1_i1 was tested on a panel of different voltage-gated ion channels, comprising potassium (K_V), sodium (Na_V), and calcium (Ca_V) channels. At 8 µM, small inhibitory activity levels were observed on K_V1.2 (3.6 ± 1.9%), K_V1.3 (11.9 ± 0.9%), K_V3.1 (12.7 ± 2.6%), K_V11.1 (hERG) (30.5 ± 1.6%), and Na_V 1.4 (14.2 ± 2.2%) channels (Figure 6). Although only limited activity at a relatively high concentration was observed on these ion channels, insulin-like peptides may display different roles in fine-tuning specific physiologic functions governed by ion channel activity in certain target species. However, further studies would be needed to validate and better understand such hypotheses.

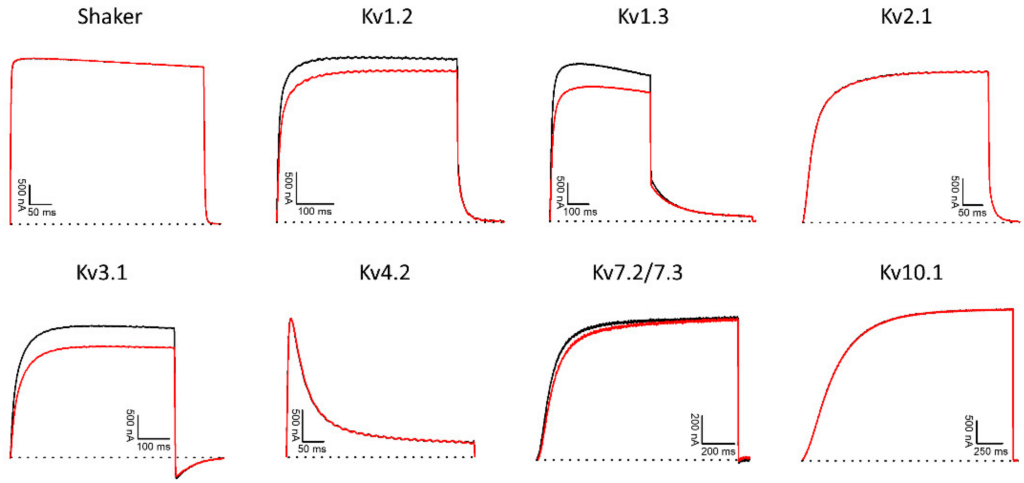


Figure 6. *Cont.*

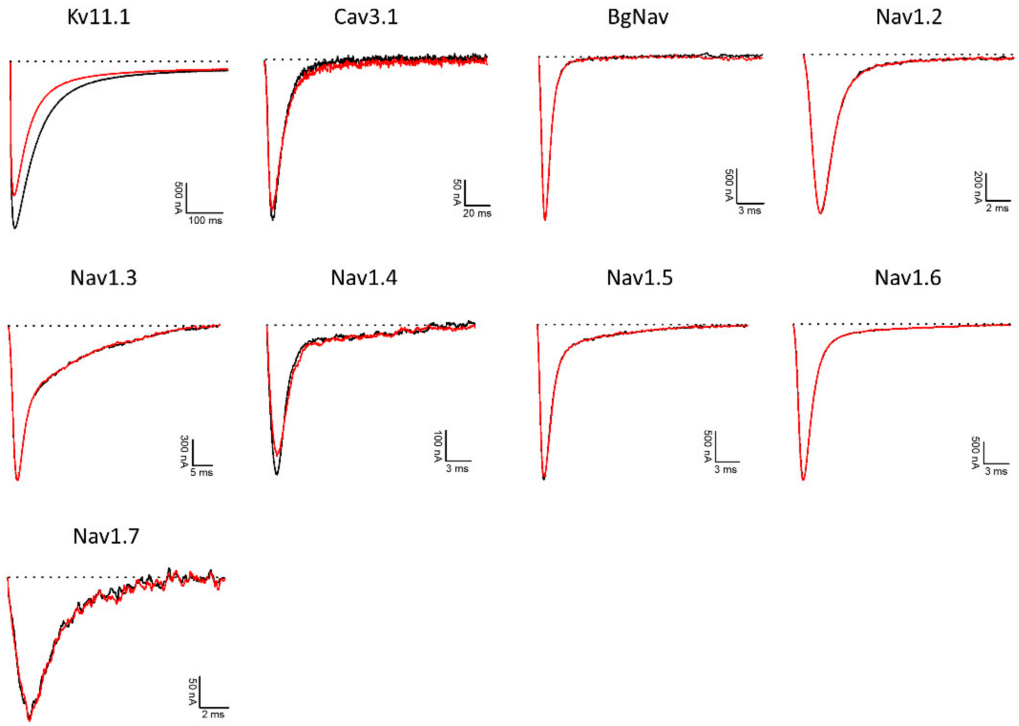


Figure 6. Electrophysiological screening of IlO1_i1 (8 μM) on potassium (K_V), sodium (Na_V), and calcium (Ca_V) channels. The black lines represent the controls, while the red lines indicate the current obtained after the addition of the peptide. Dotted lines represent zero current level. The graphs illustrate the effects obtained in a series of at least three independent experiments ($n \geq 3$).

4. Discussion

The conserved cysteine spacing in the A- and B-chains of IlO1_i1 found in this study is consistent with ILPs sequences found in other invertebrates, with the C peptide being highly variable [5]. Based on the sequence similarity to insulin and previously characterized ILPs, we were unable to predict the function of IlO1_i1. It is becoming increasingly evident that sequence similarity alone is not a sufficient indicator of peptide functionality [62–64]. As more invertebrate ILPs are characterized with a range of functions relating to tissue-specific regions, it may become possible to more accurately predict which sequences are likely to have a particular functionality.

To date, no native ILP has been identified that falls outside the three basic structural types of α-type, β-type, or γ-type. Consequently, the finding that the full-length gene sequence for *Oulactis* sp. ILP (IlO1_i1) contains a B-chain N-terminal extension that includes a putative intra-chain disulfide bond raised the question of whether it is an anomaly (see below). We made the reasonable assumption that it was unlikely to constitute a native ILP structure and therefore undertook to chemically assemble and assess the biological activity of the canonical insulin sequence of this peptide.

The solid-phase synthesis of IlO1_i1 was straightforward using our well-established procedures that have been utilized previously to prepare numerous other insulin-like peptides and their analogues [53,54]. Each chain was readily prepared, the stepwise formation of each disulfide bond was efficient, and the overall yield of the resulting, highly purified IlO1_i1 was comparable to those of other insulin-like peptide syntheses.

In light of the fact that we elected to truncate the additional cysteine pairing to mimic known active insulin and ILPs, in future chemical synthesis studies it may be worthwhile to prepare full-length IlO1_i1. Synthesis of the four-disulfide analogue would allow a comparison of its functional profile with that of the canonical ILP analogue investigated here, as well as determination of whether the extended B-chain identified through bioinformatics forms a new class of ILP structure that occurs in Cnidaria.

Whilst it was expected that the CD spectra of IlO1_i1 would reveal a peptide with a high helical content, as is seen with most insulin-like peptides (including the insulins from cone snail venom and *Drosophila* insulin-like peptide 5 [16,58]), the high number of sequence differences from other known insulin-like peptides makes it difficult to predict helical content. The IlO1_i1 peptide lacks the key FFY motif at the end of the B-chain, which in human insulin engages with the high-affinity binding site on the IR. Cone snail insulin Con-Ins G1 also lacks this motif and yet binds with reasonable affinity. Structural and analogue studies of Con Ins G1 revealed that two tyrosine residues at positions B15 and B20 substitute for the lack of the FFY motif [16,57]. Despite the IlO1_i1 peptide having a tyrosine at B15 it is unable to bind to the IR, suggesting that other sequence differences prevent binding.

Owing to the weak activity found on the range of ion channels tested in this study, conducting additional assays on IlO1_i1, including metabolism and growth studies, would be of interest. It appears unlikely from the assays conducted here that this sea anemone ILP has a role in prey capture, which is in contrast to some cone snails that use insulin as part of their venom arsenal [14]. Characterized sequences that show the closest similarity to IlO1_i1 play a role in metabolism (human insulin) and growth development in the brain (bombyxin). However, sea anemones lack the traditional circulatory system by which vertebrates, or even other invertebrates, process energy; they also lack a central nervous system and brain [50,65]. Nonetheless, as animals they must still metabolize energy from food.

Some sea anemone species contain symbiotic zooxanthellae, enabling them to acquire energy via photosynthesis from the algae [65,66]. Food is linked to the growth and size of sea anemones, which can change dependent upon food availability, i.e., they can reduce in size during times of shortage [50,67,68]. There are relatively few studies on the metabolic function of sea anemones and the genes involved; additional studies on ILPs identified in this study may progress our knowledge in this area. Conversely, conducting assays on other characterized invertebrate sequences and determining their ion channel activity may clarify whether the weak activities of IlO1_i1 on K_V1.2, 1.3, 3.1, and 11.1 channels, as well as Na_V1.4, are an isolated occurrence or specific to sea anemones. It would also be informative to assess the activity of IlO1_i1 on channels from invertebrate species, especially those from the marine environment such as small shrimp and crabs, which form part of the sea anemone diet.

The synthesis of ILPs sourced from discrete tissue regions other than the tentacles may reveal that sea anemone ILPs have varied functional roles dependent upon the tissue of origin. For example, the synthesis and assay of ILPs located in the gametic material may reveal functions and roles aligned to sexual reproduction. We cannot, however, predict that such ILPs will serve a similar function in other Anthozoa. Reproduction strategies can be highly variable among genera, i.e., sexual (broadcast spawners) vs. asexual reproduction (e.g., fission or pedal disc laceration), with some species even utilizing a combination of both strategies [69], illustrating the complex nature of these animals.

In this study we determined that novel ILPs are present in sea anemone tissues that perform specific biological functions (e.g., tentacles for defense), and that different ILPs can be localized to specific tissues. Additionally, ILPs within the Anthozoa have a highly conserved A-chain and a B-chain with an additional pair of cysteines that may prove to be a new structural class of ILPs. Further research needs to be conducted to establish the function of ILPs in sea anemones and their sister taxa, and to determine whether their activities correlate with specific biological functions in their tissue of origin.

Supplementary Materials: The following are available online at https://www.mdpi.com/article/10.3390/biom11121785/s1, Figure S1: Alignment of full-length translated insulin-like protein (ILPs) sequences from the tentacle transcriptomes of three individuals of the sea anemone *Oulactis* sp., Figure S2: Sequence alignment of top six species hits in NCBI-Blastp of Insulin-like_Oulsp_1_i1 (IlO1_i1) from *Oulactis* sp., Figure S3: Sequence alignment of ILPs from five individuals of *Oulactis* sp. from two discrete morphological tissue regions: the tentacles and the frill and acrorhagi,

Figure S4: Synthetic sea anemone insulin-like peptide [Cys7,12 (S-thiol), Cys8(But), Cys21(Acm)] A-chain, Figure S5: Synthetic sea anemone insulin-like peptide [Cys7,12 (S-S), Cys8(But), Cys21(Acm)] A-chain, Figure S6: Analytical RP-HPLC of sea anemone insulin-like peptide [Cys7,12 (S-S), Cys8(Pyr), Cys21(Acm)] A-chain, Figure S7: Synthetic sea anemone insulin-like peptide [Cys14 (S-thiol), Cys26(Acm)] B-chain, Figure S8: Synthetic sea anemone insulin-like peptide [Cys7,12 (S-S), Cys21(Acm)] A-chain/[Cys26(Acm)] B-chain, Figure S9: Sea anemone insulin-like peptide [Cys7,12 (S-S), Cys21(Acm)] A-chain/[Cys26(Acm)] B-chain, Table S1: Top 6 species in NCBI-Blastp hits for full-length insulin-like peptides (ILPs) against IlO1_i1.

Author Contributions: Conceptualization, M.L.M., R.S.N.; methodology, M.L.M., J.D.W., J.T., B.E.F., S.P.; formal analysis, M.L.M., F.L., E.L.P.-J., D.C.C.W., C.D., A.J.B.; investigation, M.L.M., F.L., E.L.P.-J., D.C.C.W., C.D., A.J.B.; resources, R.S.N., J.D.W., M.A.H., B.E.F., J.T.; data curation, M.L.M.; writing—original draft preparation, M.L.M.; writing—review and editing, all authors.; supervision, R.S.N., J.D.W., M.A.H., J.T., B.E.F.; funding acquisition, R.S.N., J.D.W., M.A.H., J.T., S.P., B.E.F. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded in part by an ARC Linkage grant LP150100621. M.L.M. acknowledges an Australian Government Research Training Program Scholarship, Monash Medicinal Chemistry Faculty Scholarship, and Monash University–Museums Victoria Scholarship top-up. J.D.W. is an NHMRC Principal Research Fellow (APP1117483). The studies undertaken in the laboratories of M.A.H. and J.D.W. were supported by an Australian National Health and Medical Research Council Project grant (APP1163310). Research at The Florey Institute of Neuroscience and Mental Health is supported by the Victorian Government Operational Infrastructure Support Program. The studies undertaken at the University of Leuven (KU Leuven) were supported by FAPESP (São Paulo Research Foundation, scholarship n. 2016/04761-4 to E.L.P.-J.), CAPES (Coordination for the Improvement of Higher Education Personnel; scholarship n. 88881.186830/2018-01 to E.L.P.-J.), FWO-Vlaanderen (grants GOA4919N, GOE7120N, GOC2319N to J.T., and 12W7822N to S.P.), and KU Leuven (grants CELSA 17/047 to J.T. and PDM/19/164 to S.P.).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Acknowledgments: M.L.M. would like to acknowledge the High-Performance Computing Facilities (MILTON) at the Walter and Eliza Hall Institute of Medical Research, on which bioinformatic analyses were conducted. We thank Chris MacRaild for running the NMR spectrum of Con-Ins G1.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Tokarz, V.L.; MacDonald, P.E.; Klip, A. The cell biology of systemic insulin function. *J. Cell Biol.* **2018**, *217*, 2273–2289. [CrossRef]
2. Ebberink, R.H.M.; Smit, A.B.; Van Minnen, J. The insulin family: Evolution of structure and function in vertebrates and invertebrates. *Biol. Bull.* **1989**, *177*, 176–182. [CrossRef]
3. Steiner, D.F.; Chan, S.J.; Welsh, J.M.; Kwok, S.C. Structure and evolution of the insulin gene. *Annu. Rev. Genet.* **1985**, *19*, 463–484. [CrossRef]
4. Grönke, S.; Clarke, D.-F.; Broughton, S.; Andrews, T.D.; Partridge, L. Molecular evolution and functional characterization of *Drosophila* insulin-like peptides. *PLoS Genet.* **2010**, *6*, e1000857. [CrossRef] [PubMed]
5. Wu, Q.; Brown, M.R. Signaling and function of insulin-like peptides in insects. *Annu. Rev. Entomol.* **2005**, *51*, 1–24. [CrossRef] [PubMed]
6. Chowański, S.; Walkowiak-Nowicka, K.; Winkiel, M.; Marciniak, P.; Urbański, A.; Pacholska-Bogalska, J. Insulin-like peptides and cross-talk with other factors in the regulation of insect metabolism. *Front. Physiol.* **2021**, *12*, 701203. [CrossRef]
7. Semaniuk, U.; Piskovatska, V.; Strilbytska, O.; Strutynska, T.; Burdyliuk, N.; Vaiserman, A.; Bubalo, V.; Storey, K.B.; Lushchak, O. *Drosophila* insulin-like peptides: From expression to functions—A review. *Entomol. Exp. Appl.* **2021**, *169*, 195–208. [CrossRef]
8. Iwami, M.; Kawakami, A.; Ishizaki, H.; Takahashi, S.Y.; Adachi, T.; Suzuki, Y.; Nagasawa, H.; Suzuki, A. Cloning of a gene encoding bombyxin, an insulin-like brain secretory peptide of the silkmoth *Bombyx mori* with prothoracicotropic activity. *Dev. Growth Differ.* **1989**, *31*, 31–37. [CrossRef]
9. Zheng, S.; Chiu, H.; Boudreau, J.; Papanicolaou, T.; Bendena, W.; Chin-Sang, I. A functional study of all 40 *Caenorhabditis elegans* insulin-like peptides. *J. Biol. Chem.* **2018**, *293*, 16912–16922. [CrossRef]
10. Floyd, P.D.; Li, L.; Rubakhin, S.S.; Sweedler, J.V.; Horn, C.C.; Kupfermann, I.; Alexeeva, V.Y.; Ellis, T.A.; Dembrow, N.C.; Weiss, K.R.; et al. Insulin prohormone processing, distribution, and relation to metabolism in *Aplysia californica*. *J. Neurosci.* **1999**, *19*, 7732–7741. [CrossRef]

11. Moon, J.-S.; Choi, Y.H. Role of the insulin-like growth factor system in gonad sexual maturation in Pacific oyster *Crassostrea gigas*. *Fish. Aquat. Sci.* **2020**, *23*, 3. [CrossRef]
12. Perillo, M.; Arnone, M.I. Characterization of insulin-like peptides (ILPs) in the sea urchin *Strongylocentrotus purpuratus*: Insights on the evolution of the insulin family. *Gen. Comp. Endocrinol.* **2014**, *205*, 68–79. [CrossRef]
13. Chandler, J.C.; Aizen, J.; Elizur, A.; Hollander-Cohen, L.; Battaglene, S.C.; Ventura, T. Discovery of a novel insulin-like peptide and insulin binding proteins in the Eastern rock lobster *Sagmariasus verreauxi*. *Gen. Comp. Endocrinol.* **2015**, *215*, 76–87. [CrossRef] [PubMed]
14. Safavi-Hemami, H.; Gajewiak, J.; Karanth, S.; Robinson, S.D.; Ueberheide, B.; Douglass, A.D.; Schlegel, A.; Imperial, J.S.; Watkins, M.; Bandyopadhyay, P.K.; et al. Specialized insulin is used for chemical warfare by fish-hunting cone snails. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 1743–1748. [CrossRef]
15. Robinson, S.D.; Safavi-Hemami, H. Insulin as a weapon. *Toxicon* **2016**, *123*, 56–61. [CrossRef] [PubMed]
16. Xiong, X.; Menting, J.G.; Disotuar, M.M.; Smith, N.A.; Delaine, C.A.; Ghabash, G.; Agrawal, R.; Wang, X.; He, X.; Fisher, S.J.; et al. A structurally minimized yet fully active insulin based on cone-snail venom insulin principles. *Nat. Struct. Mol. Biol.* **2020**, *27*, 615–624. [CrossRef]
17. Lauritano, C.; Ianora, A. Marine organisms with anti-diabetes properties. *Mar. Drugs* **2016**, *14*, 220. [CrossRef]
18. Pascual, I.; Lopéz, A.; Gómez, H.; Chappé, M.; Saroyán, A.; González, Y.; Cisneros, M.; Charli, J.L.; de los Angeles Cháveza, M. Screening of inhibitors of porcine dipeptidyl peptidase IV activity in aqueous extracts from marine organisms. *Enzyme Microb. Technol.* **2007**, *40*, 414–419. [CrossRef]
19. Antonova, Y.; Arik, A.J.; Moore, W.; Riehle, M.A.; Brown, M.R. Insulin-Like Peptides: Structure, signaling, and function. In *Insect Endocrinology*; Academic Press: San Diego, CA, USA, 2012; pp. 63–92. ISBN 978-0-12-384749-2.
20. Nässel, D.R.; Vanden Broeck, J. Insulin/IGF signaling in *Drosophila* and other insects: Factors that regulate production, release and post-release action of the insulin-like peptides. *Cell. Mol. Life Sci.* **2016**, *73*, 271–290. [CrossRef] [PubMed]
21. Altindis, E.; Cai, W.; Sakaguchi, M.; Zhang, F.; GuoXiao, W.; Liu, F.; De Meyts, P.; Gelfanov, V.; Pan, H.; DiMarchi, R.; et al. Viral insulin-like peptides activate human insulin and IGF-1 receptor signaling: A paradigm shift for host–microbe interactions. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 2461–2466. [CrossRef]
22. Matsunaga, Y.; Kawano, T. The *C. elegans* insulin-like peptides (ILPs). *AIMS Biophys.* **2018**, *5*, 217–230. [CrossRef]
23. Duret, L.; Guex, N.; Peitsch, M.C.; Bairoch, A. New insulin-like proteins with atypical disulfide bond pattern characterized in *Caenorhabditis elegans* by comparative sequence analysis and homology modeling. *Genome Res.* **1998**, *8*, 348–353. [CrossRef] [PubMed]
24. Steinmetz, P.R.H.; Aman, A.; Kraus, J.E.M.; Technau, U. Gut-like ectodermal tissue in a sea anemone challenges germ layer homology. *Nat. Ecol. Evol.* **2017**, *1*, 1535–1542. [CrossRef]
25. Anctil, M. Chemical transmission in the sea anemone *Nematostella vectensis*: A genomic perspective. *Comp. Biochem. Physiol. Part D Genom. Proteom.* **2009**, *4*, 268–289. [CrossRef] [PubMed]
26. Mitchell, M.L.; Tonkin-Hill, G.Q.; Morales, R.A.V.; Purcell, A.W.; Papenfuss, A.T.; Norton, R.S. Tentacle transcriptomes of the speckled anemone (Actiniaria: Actiniidae: *Oulactis* sp.): Venom-related components and their domain structure. *Mar. Biotechnol.* **2020**, *22*, 207–219. [CrossRef] [PubMed]
27. Sievers, F.; Wilm, A.; Dineen, D.; Gibson, T.J.; Karplus, K.; Li, W.; Lopez, R.; McWilliam, H.; Remmert, M.; Söding, J.; et al. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* **2011**, *7*, 539. [CrossRef]
28. Goujon, M.; McWilliam, H.; Li, W.; Valentin, F.; Squizzato, S.; Paern, J.; Lopez, R. A new bioinformatics analysis tools framework at EMBL-EBI. *Nucleic Acids Res.* **2010**, *38*, W695–W699. [CrossRef]
29. McWilliam, H.; Li, W.; Uludag, M.; Squizzato, S.; Park, Y.M.; Buso, N.; Cowley, A.P.; Lopez, R. Analysis tool web services from the EMBL-EBI. *Nucleic Acids Res.* **2013**, *41*, W597–W600. [CrossRef]
30. Madeira, F.; Park, Y.M.; Lee, J.; Buso, N.; Gur, T.; Madhusoodanan, N.; Basutkar, P.; Tivey, A.R.N.; Potter, S.C.; Finn, R.D.; et al. The EMBL-EBI search and sequence analysis tools APIs in 2019. *Nucleic Acids Res.* **2019**, *47*, W636–W641. [CrossRef]
31. Petersen, T.N.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 4.0: Discriminating signal peptides from transmembrane regions. *Nat. Methods* **2011**, *8*, 785–786. [CrossRef]
32. Steiner, D.F.; Park, S.-Y.; Støy, J.; Philipson, L.H.; Bell, G.I. A brief perspective on insulin production. *Diabetes Obes. Metab.* **2009**, *11*, 189–196. [CrossRef] [PubMed]
33. Nicol, D.S.H.W.; Smith, L.F. Amino-acid sequence of human insulin. *Nature* **1960**, *187*, 483–485. [CrossRef] [PubMed]
34. Hossain, M.A.; Okamoto, R.; Karas, J.A.; Praveen, P.; Liu, M.; Forbes, B.E.; Wade, J.D.; Kajihara, Y. Total chemical synthesis of a nonfibrillating human glycoinsulin. *J. Am. Chem. Soc.* **2020**, *142*, 1164–1169. [CrossRef]
35. Sreerama, N.; Woody, R.W. Estimation of protein secondary structure from circular dichroism spectra: Comparison of CONTIN, SELCON, and CDSSTR methods with an expanded reference set. *Anal. Biochem.* **2000**, *287*, 252–260. [CrossRef]
36. Pietrzkowski, Z.; Lammers, R.; Carpenter, G.; Soderquist, A.M.; Limardo, M.; Phillips, P.D.; Ullrich, A.; Baserga, R. Constitutive expression of insulin-like growth factor 1 and insulin-like growth factor 1 receptor abrogates all requirements for exogenous growth factors. *Cell Growth Differ.* **1992**, *3*, 199–205. [PubMed]

37. Denley, A.; Bonython, E.R.; Booker, G.W.; Cosgrove, L.J.; Forbes, B.E.; Ward, C.W.; Wallace, J.C. Structural determinants for high-affinity binding of insulin-like growth factor II to insulin receptor (IR)-A, the exon 11 minus isoform of the IR. *Mol. Endocrinol.* **2004**, *18*, 2502–2512. [CrossRef]
38. Soos, M.A.; Field, C.E.; Lammers, R.; Ullrich, A.; Zhang, B.; Roth, R.A.; Andersen, A.S.; Kjeldsen, T.; Siddle, K. A panel of monoclonal antibodies for the type I insulin-like growth factor receptor. Epitope mapping, effects on ligand binding, and biological activity. *J. Biol. Chem.* **1992**, *267*, 12955–12963. [CrossRef]
39. Soos, M.A.; Siddle, K. Immunological relationships between receptors for insulin and insulin-like growth factor I. Evidence for structural heterogeneity of insulin-like growth factor I receptors involving hybrids with insulin receptors. *Biochem. J.* **1989**, *263*, 553–563. [CrossRef]
40. Pinheiro-Junior, E.L.; Boldrini-França, J.; Takeda, A.A.S.; Costa, T.R.; Peigneur, S.; Cardoso, I.A.; de Oliveira, I.S.; Sampaio, S.V.; de Mattos Fontes, M.R.; Tytgat, J.; et al. Towards toxin PEGylation: The example of rCollinein-1, a snake venom thrombin-like enzyme, as a PEGylated biopharmaceutical prototype. *Int. J. Biol. Macromol.* **2021**, *190*, 564–573. [CrossRef]
41. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The protein families database. *Nucleic Acids Res.* **2014**, *42*, D222–D230. [CrossRef]
42. Surm, J.M.; Stewart, Z.K.; Papanicolaou, A.; Pavasovic, A.; Prentis, P.J. The draft genome of *Actinia tenebrosa* reveals insights into toxin evolution. *Ecol. Evol.* **2019**, *9*, 11314–11328. [CrossRef] [PubMed]
43. Baumgarten, S.; Simakov, O.; Esherick, L.Y.; Liew, Y.J.; Lehnert, E.M.; Michell, C.T.; Li, Y.; Hambleton, E.A.; Guse, A.; Oates, M.E.; et al. The genome of *Aiptasia*, a sea anemone model for coral symbiosis. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11893–11898. [CrossRef]
44. Putnam, N.H.; Srivastava, M.; Hellsten, U.; Dirks, B.; Chapman, J.; Salamov, A.; Terry, A.; Shapiro, H.; Lindquist, E.; Kapitonov, V.V.; et al. Sea anemone genome reveals ancestral Eumetazoan gene repertoire and genomic organization. *Science* **2007**, *317*, 86–94. [CrossRef]
45. Ying, H.; Hayward, D.C.; Cooke, I.; Wang, W.; Moya, A.; Siemering, K.R.; Sprungala, S.; Ball, E.E.; Forêt, S.; Miller, D.J. The whole-genome sequence of the coral *Acropora millepora*. *Genome Biol. Evol.* **2019**, *11*, 1374–1379. [CrossRef] [PubMed]
46. Chen, C.; Chiou, C.-Y.; Dai, C.-F.; Chen, C.A. Unique mitogenomic features in the scleractinian family Pocilloporidae (Scleractinia: Astrocoeniina). *Mar. Biotechnol.* **2008**, *10*, 538–553. [CrossRef]
47. Flot, J.-F.; Tillier, S. The mitochondrial genome of *Pocillopora* (Cnidaria: Scleractinia) contains two variable regions: The putative D-loop and a novel ORF of unknown function. *Gene* **2007**, *401*, 80–87. [CrossRef]
48. WormBase Release: W276. Available online: http://www.wormbase.org (accessed on 1 June 2020).
49. Kass-Simon, G.; Scappaticci, A.A., Jr. The behavioral and developmental physiology of nematocysts. *Can. J. Zool.* **2002**, *80*, 1772–1794. [CrossRef]
50. Shick, J.M. *A Functional Biology of Sea Anemones*; Chapman & Hall: London, UK, 1991.
51. Ashwood, L.M.; Mitchell, M.L.; Madio, B.; Hurwood, D.A.; King, G.F.; Undheim, E.A.B.; Norton, R.S.; Prentis, P.J. Tentacle morphological variation coincides with differential expression of toxins in sea snemones. *Toxins* **2021**, *13*, 452. [CrossRef] [PubMed]
52. Li, W.; O'Brien-Simpson, N.M.; Hossain, M.A.; Wade, J.D. The 9-Fluorenylmethoxycarbonyl (Fmoc) group in chemical peptide synthesis—Its past, present, and future. *Aust. J. Chem.* **2020**, *73*, 271–276. [CrossRef]
53. Hossain, M.A.; Wade, J.D. Novel methods for the chemical synthesis of insulin superfamily peptides and of analogues containing disulfide isosteres. *Acc. Chem. Res.* **2017**, *50*, 2116–2127. [CrossRef]
54. Lin, F.; Hossain, M.A.; Post, S.; Karashchuk, G.; Tatar, M.; De Meyts, P.; Wade, J.D. Total solid-phase synthesis of biologically active *Drosophila* insulin-like peptide 2 (DILP2). *Aust. J. Chem.* **2017**, *70*, 208–212. [CrossRef]
55. Norton, R.S. Peptide toxin structure and function by NMR. In *Modern Magnetic Resonance*; Webb, G.A., Ed.; Springer International Publishing: Cham, Switzerland, 2018; pp. 2081–2097. ISBN 9783319283883.
56. Tudor, J.E.; Pennington, M.W.; Norton, R.S. Ionisation behaviour and solution properties of the potassium-channel blocker ShK toxin. *Eur. J. Biochem.* **1998**, *251*, 133–141. [CrossRef]
57. Menting, J.G.; Gajewiak, J.; Macraild, C.A.; Chou, D.H.; Disotuar, M.M.; Smith, N.A.; Miller, C.; Erchegyi, J.; Rivier, J.E.; Olivera, B.M.; et al. A minimized human insulin-receptor-binding motif revealed in a *Conus geographus* venom insulin. *Nat. Struct. Mol. Biol.* **2016**, *23*, 916–920. [CrossRef]
58. Sajid, W.; Kulahin, N.; Schluckebier, G.; Ribel, U.; Henderson, H.R.; Tatar, M.; Hansen, B.F.; Svendsen, A.M.; Kiselyov, V.V.; Nørgaard, P.; et al. Structural and biological properties of the *Drosophila* insulin-like peptide 5 show evolutionary conservation. *J. Biol. Chem.* **2011**, *286*, 661–673. [CrossRef]
59. Ahorukomeye, P.; Disotuar, M.M.; Gajewiak, J.; Karanth, S.; Watkins, M.; Robinson, S.D.; Salcedo, P.F.; Smith, N.A.; Smith, B.J.; Schlege, A.; et al. Fish-hunting cone snail venoms are a rich source of minimized ligands of the vertebrate insulin receptor. *Elife* **2019**, *8*, e41574. [CrossRef] [PubMed]
60. Prentis, P.J.; Pavasovic, A.; Norton, R.S. Sea anemones: Quiet achievers in the field of peptide toxins. *Toxins* **2018**, *10*, 36. [CrossRef]
61. Robinson, S.D.; Norton, R.S. Conotoxin gene superfamilies. *Mar. Drugs* **2014**, *12*, 6058–6101. [CrossRef] [PubMed]
62. Krishnarjuna, B.; MacRaild, C.A.; Sunanda, P.; Morales, R.A.V.; Peigneur, S.; Macrander, J.; Yu, H.H.; Daly, M.; Raghothama, S.; Dhawan, V.; et al. Structure, folding and stability of a minimal homologue from *Anemonia sulcata* of the sea anemone potassium channel blocker ShK. *Peptides* **2018**, *99*, 169–178. [CrossRef]

63. Krishnarjuna, B.; Villegas-Moreno, J.; Mitchell, M.L.; Csoti, A.; Peigneur, S.; Amero, C.; Pennington, M.W.; Tytgat, J.; Panyi, G.; Norton, R.S. Synthesis, folding, structure and activity of a predicted peptide from the sea anemone *Oulactis* sp. with an ShKT fold. *Toxicon* **2018**, *150*, 50–59. [CrossRef]
64. Sunanda, P.; Krishnarjuna, B.; Peigneur, S.; Mitchell, M.L.; Estrada, R.; Villegas-Moreno, J.; Pennington, M.W.; Tytgat, J.; Norton, R.S. Identification, chemical synthesis, structure, and function of a new K_V1 channel blocking peptide from *Oulactis* sp. *Pept. Sci.* **2018**, *110*, e24073. [CrossRef]
65. Stephenson, T.A. *The British Sea Anemones*; The Ray Society: London, UK, 1928; Volume I.
66. Fitt, W.K. Photosynthesis, respiration, and contribution to community productivity of the symbiotic sea anemone *Anthopleura elegantissima* (Brandt, 1835). *J. Exp. Mar. Bio. Ecol.* **1982**, *61*, 213–232. [CrossRef]
67. Sebens, K. The energetics of asexual reproduction and colony formation in benthic marine invertebrates. *Am. Zool.* **1979**, *19*, 683–699. [CrossRef]
68. Chomsky, O.; Kamenir, Y.; Hyams, M.; Dubinsky, Z.; Chadwick-Furman, N.E. Effects of feeding regime on growth rate in the Mediterranean sea anemone *Actinia equina* (Linnaeus). *J. Exp. Mar. Bio. Ecol.* **2004**, *299*, 217–229. [CrossRef]
69. Fautin, D.G. Reproduction of Cnidaria. *Can. J. Zool.* **2002**, *80*, 1735–1754. [CrossRef]