Molecular and biochemical approaches to understanding conspecific cue mediated larval settlement induction in Pacific oysters *Crassostrea gigas*

同種によるマガキ（*Crassostrea gigas*）幼生の付着誘起機構の解明に関する分子的及び生化学的研究

Mary Grace Caliwan Sedanza

Graduate School of Fisheries and Environmental Sciences
Nagasaki University, Japan

September 2022

NAGASAKI UNIVERSITY

Graduate School of Fisheries and Environmental Sciences

A Dissertation Submitted in Fulfillment of

the Requirements for the Degree of

DOCTOR OF PHILOSOPHY

Molecular and biochemical approaches to understanding conspecific cue mediated larval settlement induction in Pacific oysters *Crassostrea gigas*

by

Mary Grace Caliwan Sedanza

Approved by the Review Committee:

Assoc. Prof. Dr. Hee-Jin Kim
Review Committee Member

Assoc. Prof. Dr. Asami Yoshida
Review Committee Member

Prof. Dr. Kiyoshi Osatomi
Review Committee Member

Prof. Dr. Kenichi Yamaguchi
Review Committee Member

Prof. Dr. Cyril Glenn Perez Satuito
Head Reviewer & Adviser

ACKNOWLEDGEMENTS

Deo Omnis Gloria! (All Glory to God)

To my Ph.D. Supervisor, Dr. Cyril Glenn Satuito. Thank you for giving me this opportunity to do my doctoral studies under your supervision and for providing me with many opportunities to work with other Professors in other laboratories as well as, widen my experience through conferences and various trainings. I have learned many new things about the larval ecology of marine invertebrates, especially oyster culture. Thank you also for all your support and for honing my research writing skills.

To my Ph.D. Dissertation Committee, Dr. Kiyoshi Osatomi, Dr. Kenichi Yamaguchi, Dr. Asami Yoshida, and Dr. Hee-Jin Kim. It has been an honor and a pleasant experience to have worked closely with all the professors in my Ph.D. panel. To Dr. Kiyoshi Osatomi, thank you for your generosity in allowing me to use all the facilities and materials from your laboratory. Thank you always for your kind words. To Dr. Kenichi Yamaguchi, I have gained a deeper appreciation for protein chemistry and proteomics. Thank you for your words of encouragement, patience, time, and effort in guiding me through my experiments. Thank you for always reminding me 'never give up' on my experiments even if they were challenging. To Dr. Asami Yoshida, I have learned many things about molecular biology and protein purification techniques. I am grateful for your patience, time, and effort in guiding me as well in my experiments. Thank you for your kindness and friendship. To Dr. Hee-Jin Kim, thank you for your eagerness to help and support me in my studies, for our constant conversations on many varied topics, and especially for your friendship.

To the Japanese Ministry of Education, Culture, Sports, Science, and Technology. Thank you for providing me with a scholarship that has helped me pursue my doctoral studies at Nagasaki University. I also wish to express my gratitude to the Japanese government and its people for granting me for the second time a scholarship for my studies in Japan.

To my friends and colleagues at the University of the Philippines Visayas. Thank you for your support and words of encouragement during my doctoral studies.

To Prof. Xerxes Seposo. Thank you for your patience and guidance in teaching me the R statistical programming which was very useful in my data analyses. Thanks to you and Aden for the friendship and wonderful experiences at Nagasaki. To Prof. Katsuya Hirasaka, thank you for your help and kindness in allowing me to use the equipments in your laboratory, and for all your advice in my experiments. To Prof. Tomoko Kiyota and members of the SONIC group, thank you for your genuine friendship and concern for the welfare of International students, especially to the Nagasaki University Student Organization Network for Intercultural Cooperation (NU-SONIC).

To the members of the Filipino Association of Nagasaki Students (NU-FANS; Dominic, Jean, Malou, Joy, Paul, Max, Yurika, Rhanee, Aden, Xerxes). I am blessed and truly grateful for

your friendship and support during my stay at Nagasaki University. Thank you for your words of encouragement and wonderful memories. To Prof. Joanna Orejola, thank you for your patience in guiding me through my Biochemistry-related questions and for all the friendly advice.

To Mr. Jalal Alshaweesh, Ms. Sugimaru, Mr. Chu, Ms. Gao Yi-Li, Ms. Jiang Yanrong, Mr. Yafei Zhang, Mr. Yuya Yamamoto. Thank you for all the collaborations and learnings in the different stages of my research. I have learned many things by working together with you. Thank you also to Han-san, friends at C.H.I.L.L. Saturdays (Anna, Cheechane, Adoma, Syahira, Rizki, to name a few) and MEXT 2019 batchmates (Fatema Khaladari and Nguyen Dieu Linh) for all your support and friendship. I also thank the various Staff of Nagasaki University for their kind assistance during my stay at Nagasaki University. I am sincerely grateful for all their help.

To my friends among the Nagasaki Locals, Nagasaki Seido School students, staff, and teachers who have been like a family to me. I am grateful for the warm hospitality and friendship that have been given to me. Thank you for helping me appreciate Japanese culture, its people, and its history. I am grateful for the many kind people I have met at Nagasaki.

To my family: Mama, Papa, Agnes with her husband, Jeof, and children, Jude, Jen, Jeoffy, and my other siblings Ann, Rose, and Jr. Thank you for your all-out support, your prayers, and inspiring words. I could have not accomplished this big feat without your help. Thank you for everything.

To Grace, Mikawa, Mihara, Ohara, Okuashiya, Shimogamo, Tawili Center families, and to the Prelate of Opus Dei, Msgr. Fernando Ocariz. Thank you for your support, for making me part of this beautiful family, and for inspiring me always to give my best, and to serve others with love for God.

I wish to dedicate this Ph.D. dissertation to all those who in one way or another, have helped and accompanied me on my Ph.D. journey.

Aim high, pray first, and stay focused. Never give up, either you find a way or you make one!

Mary Grace Sedanza
Nagasaki University
September 2022

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

The global decline of natural oyster populations emphasizes the need to improve our understanding of their biology. This present study elucidates the identity, characteristics, and mechanisms underlying the conspecific cue mediated induction of larval settlement in *Crassostrea gigas* using biochemical and molecular approaches. Understanding the molecular mechanisms governing oyster substrate selection may provide new insights into the molecular evolution of regulatory signal molecules, chemical sensing mechanisms, and their roles in determining the distribution and abundance of benthic organisms. A discussion is made on the key findings of the three phases of research conducted on the role of sugars in lectin-glycan interaction effects on conspecific cue mediated settlement induction and its localization on oyster larval tissues (Chapter II)*;* the identification and characterization of the biological cue responsible for gregarious settlement on conspecifics in *C. gigas* (Chapter III); and lastly, the transcriptome dynamics of the oyster larval response to the conspecific cue mediated settlement induction in *C. gigas* (Chapter IV).

In Chapter II, the involvement of lectin-glycan interaction and the important role played by sugar compounds in settlement site selection by *C. gigas* larvae is demonstrated. Sugar compounds were found to influence the ability of the larvae to identify and discriminate suitable substrata in response to chemical cues in either inhibiting or enhancing lectin-mediated processes involved in biorecognition. Sugars in the absence of a conspecific cue, CGSPPC, did not promote settlement whereas, in the presence of the cue showed varied effects, most of which were found inhibitory at different concentrations. Sugar treated larvae exposed for 2 h showed significant settlement inhibition in the presence of a conspecific cue. Sialic acid, as well as GlcNAc sugars, showed a similar interaction trend with wheat germ agglutinin (WGA) lectin. WGA-FITC conjugate showed positive binding on the foot, velum, and mantle when exposed to GlcNAc sugars. This study suggests that a WGA lectin-like receptor and its endogenous ligand are both found in the larval chemoreceptors and the conspecific cue, CGSPPC, that may complementarily work together to allow the oyster larva greater selectivity during site selection. The effect of sugars on settlement is significant in the study of chemical communication and could have substantial implications for oyster ecology.

In Chapter III, a novel role of several shell matrix proteins acting as a *Crassostrea gigas* Settlement Pheromone Protein Components (CGSPPC) is demonstrated as the biological cue responsible for gregarious settlement on conspecifics. A bioassay-guided protein separation approach aided by biochemical and molecular analyses reveals that Gigasin-6 isoform X1

and/or X2 isolated from the adult shells is the major inducing cue for larval settlement and may also play a role in postlarva-larva settlement interactions. Other isolated stains-all stainable acidic proteins may function as co-factor and a scaffold/structural framework for other matrix proteins to anchor within this assembly and provide protection. Notably, conspecific cue mediated larval settlement induction in *C. gigas* presents a complex system that requires an interplay of different glycans, disulfide bonds, amino acid groups, and phosphorylation crosstalk for recognition. These present findings introduce another novel concept in chemical ecology demonstrating several signaling molecules in CGSPPC that could induce larval settlement. This is in contrast to other reported settlement cues in marine invertebrates that come from a single signaling molecule with a less complex lectin-glycan settlement interaction. On the other hand, the receptor that could recognize CGSPPC may involve a WGA lectin-like larval receptor and a phosphoprotein. The molecular basis of the conspecific cue mediated larval settlement mechanism is also proposed.

In Chapter IV, a transcriptome profile of an oyster larval response to CGSPPC conspecific cue mediated settlement induction in *C. gigas* is reported for the first time. In this study, the transcriptomes of competent pediveligers (Pedi) and conspecific cue-induced postlarvae (PL) were described and compared. A total of 2,383 candidate transcripts were identified: 740 upregulated and 1643 downregulated transcripts, after settlement. Gene Ontology analysis revealed active chitin binding, calcium ion binding, and extracellular region processes in both stages. Results showed that the differential expression trend of six candidate transcripts were consistent between the quantitative real-time PCR and transcriptome data. The differential transcript expression related to shell formation showed closely linked dynamics with a gene regulatory network that may involve the interplay of various hormone receptors, neurotransmitters, and neuropeptide receptors working together in a concerted way in the Pedi and PL stages. Results also suggest the possible involvement of an ecdysone signal pathway that may be linked to a neuroendocrine-biomineralization crosstalk in *C. gigas* settlement. Based on these results, the molecular mechanism of the transcriptome dynamics of an oyster larval response to a conspecific cue mediated settlement induction in *C. gigas* is proposed.

In conclusion, oysters may seem simple, but this study provides insights into the complex chemical signaling processes and molecular mechanisms governing conspecific cue mediated larval settlement induction that may indicate their evolutionary success as benthic organisms. Continued study on CGSPPC and its corresponding oyster larval receptor will advance our understanding of their evolution, population dynamics, and chemical communication. These results may find application in the development of oyster aquaculture by using this extracted

compound as a surface-bound attractant to collect wild and hatchery-grown oyster larvae which could help recover declining marine species; as well as a target of anti-fouling agents on man-made structures.

要旨

天然マガキ（*Crassostrea gigas*）は、世界規模で資源量が減少しており、本種を生物学的観点から深く理解する重要性が指摘されている。本研究では、同種による本種幼生の付着誘起の根底にある（cue）キューの特定及び作用機序を分子的・生化学的手法を用いて検討した。本種の基盤選択を制御する分子機構を理解することは、底生生物の分布及び出現量とそれらを制御する調節シグナル分子の分子進化や化学センシング機構に対する新しい知見を得ることに繋がる。本研究で得られた成果について、同種キューを介した付着誘起におけるレクチン - グリカン相互作用に対する糖化合物の影響（第2章）、本種幼生の同種への付着（群居性）の根底にある生物学的キューの特定（第3章）及び同種キューを介した付着誘起に伴う幼生の応答におけるトランスクリプトーム変化（第4章）の3つの項目に分けて考察した。

第2章では、幼生の付着場所選択における糖化合物の役割及びレクチン - グリカン相互作用の関与を検討した。糖化合物は、付着場所探索の際に幼生の化学キュー認識に関与しレクチンを介したプロセスを促進あるいは阻害することが分かった。つまり、糖化合物は同種由来のキューであるCGSPPCの存在下で殆どが濃度依存的な阻害効果を示し、2時間暴露では幼生の付着を有意に阻害した。シアル酸及びN-アセチルグルコサミンGlcNAcは小麦胚芽凝集素（WGA）に対して同様の相互作用を示し、GlcNAcに暴露した幼生は足、面盤、外套膜がWGA-FITCに染色された。以上より、WGAレクチン様受容体及び内因性リガンドは幼生の化学受容体及びCGSPPCの両方に含まれていることが示唆され、両方の補完的作用により幼生の基盤選択性が向上すると考えられる。

第3章では、本種幼生の同種への付着誘起は貝殻マトリックスタンパク質（SMPsの高分子集合体がマガキ付着フェロモンタンパク質成分（CGSPPC）として生物学的キューの機能を担っているに起因することを証明した。つまり、付着実験を伴うタンパク質の分画とともに分子的・生化学的手法による解析の結果、成体貝殻から分離・特定したGigasin-6 X1/X2アイソフォームが幼生の主要な付着誘起キューであることを明らかにし、付着における幼生・稚貝の相互関係に関与する可能性も示された。他に分離したstains-allで染色された酸性タンパク質は補因子及び足場タンパク質として他のSMPsを集合体につなぎ止め保護する機能を担うと考えられた。とりわ

け本種では、同種による幼生の付着誘起は複雑なシステムによって認識され、幼生は複数のグリカン、ジスルフィド結合、アミノ酸、及びリン酸化クロストークの相互作用によって付着することが示唆された。本結果は、幼生の付着がCGSPPCの複数の伝達分子によって誘起されることを証明し化学生態学において新しい概念を提唱した。本新規概念は、これまで報告された海産無脊椎動物における付着キューが単一の伝達分子によるレクチン - グリカン相互作用によってもたらされるものと対照的である。一方、CGSPPCを認識する幼生の受容体にはWGAレクチン様受容体及びリン酸タンパク質が関与する可能性も示された。さらに、同種由来のキューを介した幼生付着機構の分子的機序を提言した。

第4章では、CGSPPCを介した付着誘起に伴う幼生の応答におけるトランスクリプトーム変化の解析を行った。すなわち、付着可能なペデイベリジャー幼生（Pedi）とCGSPPCによって付着した稚貝（PL）のトランスクリプトームを比較した結果、合計2,383の候補遺伝子群を特定し、幼生は付着後に遺伝子群のうち740が発現上昇し1,643が減少した。GO解析の結果、PediとPLの両ステージにおいてキチン結合の活性化、カルシウムイオン結合、細胞外領域での諸過程が確認された。また、ターゲットの6つの候補遺伝子群の発現変動が定量リアルタイムPCRでも一致していた。PediとPLでは、貝殻形成関連の遺伝子群発現変動は複数のホルモン受容体、神経伝達物質、神経ペプチド受容体の相互作用に関与する遺伝子調節ネットワークと連係した作用機序が示された。さらに、本種の付着は、エクジソン信号伝達経路の関与が示唆され、神経内分泌 - バイオミネラリゼーションクロストークとの関連も示唆された。

以上より、本種は一見シンプルな生物のように思われるが、同種由来のキューを介した本種幼生の付着は複雑な化学信号伝達過程及び分子機構によって支配されていることが示された。今後、CGSPPCの全構造及びその受容体を特定できれば、本種の進化・個体群動態・ケミカルコミュニケーションへの理解の向上につながる。本研究で得られた成果より、有用種の養殖のための新しい採苗器の開発や資源回復につながる技術開発等に応用が期待される。

Chapter I. General Introduction

The Pacific oyster *Crassostrea gigas* (Thunberg, 1973) is a benthic bivalve mollusk that is naturally distributed in the Northwest Pacific, ranging from China and Korea to Japan. It has been introduced worldwide, and has become one of the most well-studied organisms due to its biological, ecological, and aquaculture importance [1,2]. It has been chosen as model species for studying molecular genetics [2–5], developmental biology [6–11], biomineralization [12–14], and adaptation to coastal environments under climate change [15]. Pacific oysters, like other oyster species, act as foundation species by creating habitat for other estuarine species and providing key ecosystem services to human communities around the world [16–18]. As a food source, bivalves such as oysters are high in micronutrients and have great nutritional value. Bivalve demand has risen significantly over time, resulting in the expansion of the bivalve aquaculture industry in many countries [19]. Global aquaculture production of Pacific oysters has been estimated at 643,500 tonnes in 2018 [19] which continues to make it a leading species cultured among oysters, in particular, and 3rd to Japanese carpet shell, *Ruditapes philippinarum* among mollusks, in general [19]. Wild seed capture accounts for a large portion of the global supply of spat, which is done with a variety of settlement materials (cultch) suspended from longlines and rafts. When natural spats are scarce, some hatcheries cultivate larvae and induce metamorphosis using *C. gigas* oyster shells [20,21].

A recent study by Beck et al. (2011) has reported that oyster populations have significantly declined on a global scale, with an estimated 85% of oyster reefs gone, far outnumbering the projected loss of coral reefs [22]. The global decline of natural oyster populations emphasizes the need to improve our understanding of their biology. Understanding how these oysters, such as *C. gigas,* occupy appropriate substrata is crucial to learning about their evolution, population dynamics, and chemical communication. The biphasic life cycle of *C. gigas* is composed of a planktonic larval stage for dispersion and a sessile adult stage, with metamorphosis marking the transition between the two life stages in a relatively short time (generally less than 48 h) [4,7,23]. Unlike larval dispersion, settlement occurs on a small spatial scale and involves larval behavior in response to the physical and chemical characteristics of the substratum [24].

In this study, settlement is defined as the sequential transition of a competent pediveliger larva to a post-larva which includes the attachment to a substrate and then eventual metamorphosis. Oyster settlement behavior involves a swimming motion using its velum and with the foot extended forward, followed by a series of increasingly localized

crawling maneuvers until it finds a suitable substratum [6]. During the search phase, the oyster larva makes use of its external chemoreceptors and transduces the cues it encounters in the environment into internal processes with neural and/or hormonal elements [25]. This is followed by a set of behavioral changes such as permanent cementation to the substratum where it finally metamorphoses into a juvenile [6]. Metamorphosis is an irreversible physiological process that includes a change from a swimming larva to a sessile juvenile, including loss of the velum, eyespot, and foot; development of gills, and production of adult shell [10]. In addition, the metamorphic transition often involves morphological, physiological, structural, and functional changes that are governed by a gene regulatory network [26].

The induction of oyster larval settlement by chemical cues in the marine environment may come from a variety of sources. A wide array of waterborne or surface-bound cues associated with conspecific adults have been reported to induce larval settlement. These include shells in *C. gigas* [27,28], *C. virginica* [29,30], *C. ariakensis* [31]; soft tissue homogenates in *Ostrea edulis* [32]; water pre-conditioned by adults of *C. ariakensis* [31], *C. gigas* [33], *C. virginica* [34–36], and *O. puelchana* [37]; oyster shell liquor or extrapallial fluid of *C. virginica* [38]; and ammonia [39]. Another source of such cues are chemicals bound to or released from bacterial biofilms in the form of extracellular polymeric substances [6,40–42]. The common components of extracellular and cell surface polysaccharides in bacterial biofilms are neutral sugars (D-glucose, D-galactose, D-mannose, D-fucose, L-rhamnose), amino sugars (N-acetylglucosamine and N-galactosamine), and some uronic acids [43–45]. Despite the notable advances in this active field of study, much remains to be done in identifying the exact nature of these cues. Although several of these studies show that bacterial biofilm-derived cues may induce appropriate settlement sites. The study of conspecific-derived inducing cues provides another hypothesis explaining their possible role in the gregarious behavior exhibited by oyster larvae upon settlement. Previous studies by Vasquez et al. [27,28] have shown that a protein extract from shells of conspecifics could induce larval settlement in *C. gigas* competent pediveligers.

This present study elucidates the identity, characteristics, and mechanisms underlying the conspecific cue-mediated induction of larval settlement in *C. gigas* using bioassay-guided fractionation, biochemical and molecular approaches. A discussion is made on the role of sugars in lectin-glycan interaction effects on conspecific cue-mediated settlement induction and its localization on oyster larval tissues. The identification and characterization of the biological cue responsible for gregarious settlement on conspecifics in *C. gigas* were

investigated. Lastly, the transcriptome dynamics of the oyster larval response to the conspecific cue-mediated settlement induction in *C. gigas* were further explored.

Chapter II seeks to understand how larvae respond to the presence of carbohydrate moieties of chemical cues adsorbed in surfaces, as well as other waterborne cues that come in contact with larval external chemoreceptors during the search for suitable settlement sites. These carbohydrate moieties could bind to lectins associated with these chemoreceptors and mediate the selection process. Hence, to understand how the presence of varied chemical cues regulates settlement behavior, larvae were exposed to 12 types of sugars including WGA-specific sugars (*N*-acetylglucosamine and *N*-acetylneuraminic acid), and exposed to the conspecific cue settlement inducer. To test this hypothesis, the following questions were addressed: (1) Does the presence of dissolved low molecular weight sugar cues alone promote oyster larval settlement? (2 What types of dissolved sugars influence the ability of oyster larvae to discriminate suitable settlement sites in the presence of conspecific cues? (3) Does time exposure to sugar in the presence of a conspecific cue influence settlement? (4) Does blocking the carbohydrate moieties on the conspecific cue, i.e., adsorbed CgSE affect its settlement inducing activity? (5) Do lectins contained in the larval chemoreceptors of oysters bind dissolved sugars in the environment?

Chapter III seeks to highlight the chemical basis for the gregarious settlement behavior of Pacific oysters. The identification and characterization of the active cues that drive such induction of conspecific cue-mediated larval settlement phenomenon have not yet been clarified. Previously identified chemical cues in other marine invertebrate species were shown to be induced by a single protein [24,38]. However, a novel role of a macromolecular assembly of shell matrix proteins that are acting as an aggregating settlement cue on conspecifics is demonstrated for the first time in this study and hereon named *Crassostrea gigas* Settlement Pheromone Protein Complex (CGSPPC). A discussion is made on the properties and contribution of each protein component to settlement induction, the role of post-translational modifications (PTMs), and lastly, the molecular basis of the conspecific cue-mediated larval settlement mechanism.

Chapter IV seeks to explore the effect, at a molecular level, on the exposure of competent oyster larvae to a natural settlement-inducing cue from conspecifics, the CGSPPC. Hence, the aim of this study is to fill in some gaps in our knowledge on the underlying specific cellular and physiological processes involved in the development of marine invertebrates, specifically on *C. gigas*, with the aid of the 'omics' approach, such as transcriptomics. In this study, a DNBseq-G400RS high-throughput sequencing technology was applied to evaluate the

effect of settlement induction by a conspecific cue on *C. gigas* larvae, before and after its recognition, and the transcriptomes of the pediveliger and conspecific cue-induced postlarvae were compared. Several analyses were carried out such as gene function annotation, Gene Ontology enrichment, and quantitative real-time PCR verification. Moreover, key genes and molecular events involved in this process were identified and the possible involvement of a neuroendocrine-biomineralization crosstalk in *C. gigas* larval settlement is proposed.

Lastly, understanding the processes involved in the induction of conspecific cue-mediated larval settlement in *C. gigas* can help clarify the larval settlement mechanism of oysters, and these results may find application in the development of oyster aquaculture by using this extracted compound as a surface-bound attractant to collect wild and hatchery-grown oyster larvae which could help recover declining marine species; as well as a target of anti-fouling agents on man-made structures. This present work provides additional knowledge on chemical sensing mechanisms and the regulatory role of signal molecules that could influence the distribution and abundance of benthic organisms such as oysters.

Chapter II: Regulatory Role of Sugars on the Settlement Inducing Activity of a Conspecific Cue in Pacific Oyster *Crassostrea gigas*

II.1 Introduction

Larval settlement behavior and the choice of final substratum for attachment are important factors that influence survival, growth, and reproduction in benthic marine invertebrates [1–3]. The settlement behavior of larvae of the Pacific oyster *Crassostrea gigas* (Thunberg), like in other oyster species, often involves a swimming motion using its velum and with the foot extended forward, followed by a series of increasingly localized crawling maneuvers until it finds a suitable substratum [4–6]. During the search phase, the oyster larva makes use of its external chemoreceptors and transduces the cues it encounters in the environment into internal processes with neural and/or hormonal elements [3]. This is followed by a set of behavioral changes such as permanent cementation to the substratum where it finally metamorphoses into a juvenile [4]. Several studies have shown that chemical cues and surface properties of the substrate are important in settlement site selection by the larva [1,7].

The induction of oyster larval settlement by chemical cues in the marine environment may come from a variety of sources. A wide array of waterborne or surface-bound cues associated with conspecific adults have been reported to induce larval settlement. These include shells in *C. gigas* [8,9], *C. virginica* [10,11], *C. ariakensis* [12]; soft tissue homogenates in *Ostrea edulis* [13]; water pre-conditioned by adults of *C. ariakensis* [12], *C. gigas* [14], *C. virginica* [15–17], and *O. puelchana* [18]; oyster shell liquor or extrapallial fluid of *C. virginica* [19]; and ammonia [6]. Another source of such cues are chemicals bound to or released from bacterial biofilms in the form of extracellular polymeric substances [20–23]. The common components of extracellular and cell surface polysaccharides in bacterial biofilms are neutral sugars (D-glucose, D-galactose, D-mannose, D-fucose, L-rhamnose), amino sugars (*N*-acetylglucosamine and *N*-galactosamine), and some uronic acids [24–26]. However, despite the notable advances in this active field of study, much remains to be done in identifying the exact nature of these cues.

A previous report by Maki and Mitchell (1985) has demonstrated that lectins can be a useful tool in examining the role of sugar compounds in the larval settlement of marine organisms [27]. Lectins are sugar-binding proteins of non-immune origin that agglutinate cells and precipitate polysaccharides or glycoproteins [28]. Earlier works in other marine invertebrates, such as barnacles have shown that the presence of dissolved sugars influences

the exploration, identification of conspecific cues, and larval metamorphosis [29–31]. Also, Vasquez et al. (2014) have provided evidence that a WGA-binding sugar chain in shells of conspecifics may mediate the settlement of *C. gigas* larvae on conspecifics, thus, highlighting the role played by lectin-glycan in biorecognition [8,9]. However, the nature of the chemical cue from shells of *C. gigas* is not yet fully characterized and elucidated. Understanding the processes and cues that lead to oyster larval settlement can help us understand the underlying mechanisms that regulate population dynamics and chemical communication.

In this present study, the potential conspecific cue from adult shells of *C. gigas* was extracted using Ethylenediaminetetraacetic acid (EDTA) and its settlement inducing activity was evaluated. The suitable substrate for *C. gigas* shell EDTA extract (CgSE) for settlement assays was also determined. In nature, larvae also encounter different glycoproteins adsorbed to surfaces or exposed to dissolved cues. Hence, to understand how the presence of varied chemical cues regulate settlement behavior, larvae were exposed to 12 types of sugars including WGA-specific sugars (*N*-acetylglucosamine and *N*-acetylneuraminic acid) and exposed to conspecific crude shell extract (CgSE). I hypothesized that, during the search for suitable settlement sites, the carbohydrate moieties of chemical cues adsorbed in surfaces, as well as other waterborne cues come in contact with larval external chemoreceptors on the velum and foot. Therefore, the carbohydrate moieties could bind to lectins associated with these chemoreceptors and mediate the selection process. To test this hypothesis, we addressed the following questions: (1) Does the presence of dissolved low molecular weight sugar cues alone promote oyster larval settlement? (2) What types of dissolved sugars influence the ability of oyster larvae to discriminate suitable settlement sites in the presence of conspecific cues? (3) Does time exposure to sugar in the presence of a conspecific cue influence settlement? (4) Does blocking the carbohydrate moieties on the conspecific cue, i.e., adsorbed CgSE affect its settlement inducing activity? (5) Do lectins contained in the larval chemoreceptors of oysters bind dissolved sugars in the environment?

II.2 Materials and Methods

II.2.1 Spawning and Larval Culture of Crassostrea gigas

Adult Pacific oysters (*C. gigas*) were purchased from Konagai Fisheries Cooperative, Nagasaki, Japan. The spawning method and larval culture condition used in this experiment have been described in detail by Vasquez et al. [8]. Adult *C. gigas* broodstock were maintained in net cages, suspended from a raft of Nagasaki Prefecture Fisheries Station, Nagasaki, Japan (129°51′E; 32°43’N). They were brought to the laboratory for collection of gametes. Alternatively, the broodstocks were maintained in a 30-L aquarium inside the laboratory and were fed once daily with a combination of Chaetoceros gracilis and an artificial feed for bivalves (M1, Nosan Corp., Kanagawa, Japan). Adults were kept at 20 ± 1 °C in a 30-L aquarium during the spawning season to suppress the natural spawning of the broodstock. During the winter season, the adults were maintained at 25 ± 1 °C to allow gonad development and maturation. Every other day, 100% of the water in the aquarium was changed. We used at least four different broodstocks for spawning to obtain larvae for the different settlement assays.

Gametes were collected after stripping the adult oysters. Eggs and sperm were separately suspended in 2 L glass beakers containing GF/C (Whatman glass fiber filter; pore size: 1.2 mm) filtered seawater (FSW) adjusted to 27 °C. Eggs were washed several times with FSW through repeated decantation and were then fertilized with a small volume of sperm suspension. Thirty minutes after artificial fertilization, fertilized eggs were collected in a 20 mm net, washed four to five times with FSW, and re-suspended in 2 L glass beakers containing FSW. Fertilized eggs were kept at 27 ± 1 °C in an incubator for 24 h. Then, the swimming straight hinged larvae were collected in a 40 mm net, gently washed with FSW, stocked in 2 L glass beakers, at an initial density of 5 larvae mL^{-1}, and cultured in a water bath at 27 ± 1 °C.

C. gigas larvae were mass-reared in 10-L tanks and were fed with *Chaetoceros calcitrans* (10,000–50,000 cells/mL) from day 1 to day 5, fed a combination of *C. calcitrans* (25,000 cells/mL) and *C. gracilis* (25,000 cells/mL) from day 6 to day 10, and were then fed *C. gracilis* (50,000 cells/mL) from day 11 onward during the culture period. Cultures were kept in a dark environment and the water was renewed daily throughout the culture period. The salinity of seawater used was 32 psu. Larvae usually reached the pediveliger stage 17–18 days after fertilization. Pediveligers used in assays were between 20 days old and 28 days old after fertilization and ranged from 300–360 mm in shell length.

II.2.2 Shell Preparation and Matrix Extraction

Shells from freshly shucked oysters (*C. gigas*) were thoroughly cleaned and washed with tap water to remove adhering epibionts and traces of muscle tissues, and then dried. Shell chips (SC) were prepared as described by Vasquez et al. [8]. Dried shells were crushed with a hammer to pass through a 1.0 mm mesh screen then through a 0.5 mm mesh screen. Shell fragments that remained on the 0.5 mm mesh screen were collected and used as shell chips (SC).

Shell matrix extraction was performed following a modified method of Liu et al. [32]. Shell chips (150 g) were decalcified with 1 L of 0.8 M Ethylenediaminetetraacetic acid (EDTA, pH 8.0) for 60 h at 4 °C with continuous agitation. The supernatant was collected by centrifugation at 10,000 × g for 60 min at 4 °C. It was subsequently followed by filtration through a 0.22 µm membrane filter (MF-Millipore, Cork, Ireland), then suspended in 3.5 kDa molecular cutoff dialysis tubing (ø 28.6 mm; 1.25 nm pore size; Japan Medical Science, Osaka, Japan). To remove excess EDTA, the sample was extensively dialyzed against distilled water at 4 °C for 3 days, until the final pH of the extract was the same as the initial pH of the distilled water. The crude *C. gigas* shell EDTA extract (CgSE) was lyophilized, and the resultant powder was dissolved in as small a volume as possible of distilled water and stored at -20 °C until further use. The protein content of the extract was quantified by a BCA assay kit (Pierce, Thermo Fisher Scientific, Rockford, Illinois, USA).

II.2.3 Larval Substrate 'Choice' Settlement Assay

The crude shell EDTA extract (CgSE) was subjected to a larval assay at various protein extract amounts and substrate types: 0, 50 µg, and 100 µg; GF/C, polystyrene, and PVDF, respectively. Shell chips (SC) at 50 mg was used as the reference substrate. Detailed descriptions of the preparation of different substrates prior to use in assays have been reported [9], except for polystyrene, where CgSE was directly applied on the multi-well plates and dried at 37 °C. All settlement assays were carried out using polystyrene multiwell plates (6-wells; ø 34 mm x 17 mm height; Violamo, As One Corporation, Osaka, Japan). Ten larvae were released into each well plate filled with 10 mL filtered seawater (FSW) and was repeated twice using two different batches of larvae with three replicates at each trial (n = 6). Settlement was evaluated by the number of individuals that metamorphosed to post-larvae within 24 h. Post-larvae were confirmed under the microscope as individuals that secreted cement substances or those with post-larval shell growth. All succeeding larval settlement assays were conducted in a dark environment at 27 ± 1 °C in an incubator.

II.2.4 Treatment of Larvae with Sugars and Settlement Assay

Larvae were immersed in millipore-filtered seawater (0.22 μm) containing graded concentrations (10−10, 10−8, 10−6, 10−4 M) of dissolved sugars for 2 h [3] relative to those recorded in previous studies [3–6]. Monosaccharides (D-arabinose, D-fructose, D-galactose, D-glucose, D-mannose, D-xylose, N-acetylglucosamine (GlcNAc), N-acetylneuraminic acid (Sialic acid or Neu5Ac), α-methyl-D-mannoside) and Disaccharides (lactose, maltose, sucrose) were used (Nacalai Tesque, Kyoto, Japan except for D-galactose (Wako Pure Chemical Co., Osaka, Japan)). After 2 h immersion into the sugar solutions, the sugar treated larvae were washed 3 times in 1 L FSW. Ten sugar treated larvae were introduced into 6-well plates coated with CgSE and to 10 mL FSW. The CgSE-coated surfaces were prepared by inoculating the 6-well plates with CgSE at a protein concentration of 50 μg mL−1. Untreated larvae were also released into the well plates containing CgSE-coated surfaces and FSW as the control. The 24 h larval settlement assays were repeated twice using different batches of larvae with three replicates at each trial (n = 6).

II.2.5 Assay Protocol on the Effect of Exposure Time on Sugar and Conspecific Cues Induction of Settlement

The effect of exposure times on selected inhibitory sugars (GlcNAc, Neu5Ac, αMDM, lactose, maltose) at 10−6 and 10−4 M was examined. Exposure times were 0.25, 2, and 24 h. After a specific exposure time, the sugar-treated larvae were washed three times in 1 L FSW and were used in settlement assays, as described earlier.

II.2.6 Assay Protocol on WGA and Sugar (GlcNAc and Neu5Ac) Interaction

Wheat germ agglutinin lectin (WGA; Wako Pure Chemical Co., Osaka, Japan) is known to bind with N-acetylglucosamine (GlcNAc) and N-acetylneuraminic acid or Sialic acid (Neu5Ac) sugars. Lectin-sugar interaction between WGA-GlcNAc and WGA-Neu5Ac was investigated following the method of Vasquez et al. [9]. In brief, CgSE was inoculated on a GF/C filter paper (ø 47 mm, GF/C) and dried at 37 °C. Dried CgSE papers were immersed in solutions containing 50 μg mL−1 WGA lectin or in mixtures of 50 μg mL−1 WGA and different concentrations (10−8, 10−6, 10−4 M) of GlcNAc or Neu5Ac sugars for 2 h. After 2 h, treated CgSE papers were washed three times in 1 L FSW and then used in assays. CgSE papers that were not treated with the mixture of WGA and GlcNAc or Neu5Ac were also washed in the same manner and were used as the control. All GF/C filter papers were fastened

to the bottom of the well using double-sided adhesive tape. In a 6-well plate containing different coated and non-coated GF/C filter papers, ten larvae were released into each well filled with 10 mL FSW. A 24 h larval settlement assay was carried out using three replicates for each treatment (n = 3).

II.2.7 Treatment of Oyster Larvae with Fluorescein Isothiocyanate-Conjugated WGA

To examine the distribution of GlcNAc sugar bound on *C. gigas* larval tissues, the larvae were stained with FITC-conjugated WGA (Sigma-Aldrich, Rehovot, Israel). The larvae were immersed in FSW containing 10−10 and 10−4 M GlcNAc solutions for 2 h and subsequently rinsed three times in 1 L FSW, dyed for 5 min in a solution of FITC-WGA diluted in DW at 0.5 mg mL−1. After incubation for 5 min, the larvae were rinsed by filtered seawater; then immersed in 7.5% MgCl2 solution in a polystyrene 6-well plate and observed using an epifluorescence microscope (Excitation wavelength (437–460 nm (448 nm)); Emission wavelength (472 nm); Nikon Eclipse Ts2, Nikon Corporation, Tokyo, Japan). The use of MgCl2 has been reported as an effective anesthetic that renders an animal in a complete loss of muscle tone, thus, subjecting it to a relaxed state suitable for physiological observations [33]. The untreated larvae exposed to WGA-FITC were used as control. To check for autofluorescence, untreated larvae that were not dyed with WGA-FITC were also compared.

II.2.8 Statistical Analysis

The settlement percentage (%) was calculated based on the number of settled larvae over the total number of larvae in each well multiplied by 100. These percentages were presented as arithmetic means with standard deviation (SD). Data were analyzed using quasi-binomial generalized linear models (GLM). Wald test [34] was used for pairwise comparisons. All statistical tests were performed in RStudio (R-project.org, version 4.0.3).

II.3 Results

II.3.1 Settlement Inducing Activity of Crassostrea gigas Shell EDTA (CgSE) Extract and Larval Substrate Specificity

Settlement percentages of *C. gigas* larvae on different substrates coated with varying amounts of CgSE are shown in Figure 2.1. Among the variables, substrate type, and amount of extract, the amount of extract was only found significant (Chi-square = 38.9; p-value < 0.001). The presence of adsorbed CgSE across all substrates showed a significant amount of extract dependent settlement inducing activity ($p < 0.05$). No settlement was observed on all substrates in the absence of CgSE. When substrates were coated with 50 µg of CgSE, settlement ranged from 33 to 48%. Moreover, quasi-binomial glm analysis showed that the odds of settlement, if the amount is at 50 µg, is 57 times greater when compared to 0 µg CgSE (95% CI: 12.59 to 258.66; p-value < 0.001). At 100 µg, the mean settlement ranged from 52 to 70% in all CgSE-coated substrates. Hence, 50 µg CgSE was used as a suitable amount for the succeeding larval settlement assays. Shell chips at 50 mg were used as a reference substrate and settlement were at 35%.

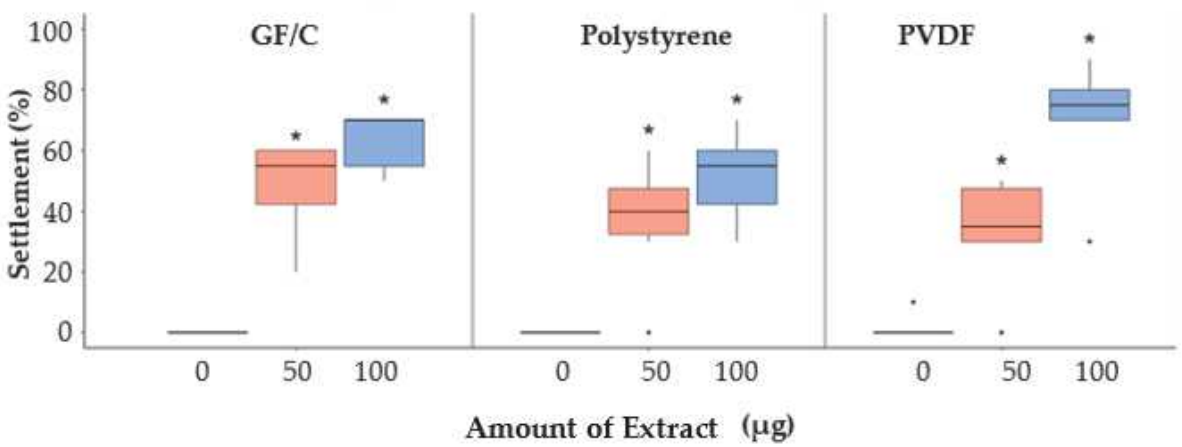

Figure 2.1. Settlement percentages of *C. gigas* larvae on different substrates coated with varying amounts, 0, 50, and 100 µg, of *Crassostrea gigas* shell EDTA-soluble extract (CgSE) after 24 h. Asterisks (*) denote significant differences in the amount coated on different substrates, using 0 µg as the baseline, determined via quasi-binomial glm ($p < 0.05$, $n = 6$, using different batches of larvae).

II.3.2 Sugars and CgSE Effects on the Settlement of Oyster Larvae

Twelve sugars, which represent a wide range of possible carbohydrate moieties, associated with settlement-inducing chemical cues, reported in marine invertebrate larvae, were screened to test their effect on *C. gigas* larvae, and the results are shown in Figure 2.2. No settlement was observed in untreated *C. gigas* larvae on the multi-well with FSW (C-FSW). Settlement percentages of sugar treated, and untreated, larvae when exposed to non-coated surfaces and CgSE-coated surfaces elicited different responses ($p < 0.05$). In non-coated surfaces (Figure 2.2A), treatment of larvae with sugars alone did not elicit settlement except for some sugars (αMDM, D-fructose, D-glucose, D-mannose, and Neu5Ac). However, the settlement behavioral pattern was found to be similar when each of these sugars was compared to untreated larvae exposed to FSW alone (C-FSW, $p > 0.05$). No mortality was recorded in all sugar treated larvae. Larvae that did not settle after 24 h were actively swimming.

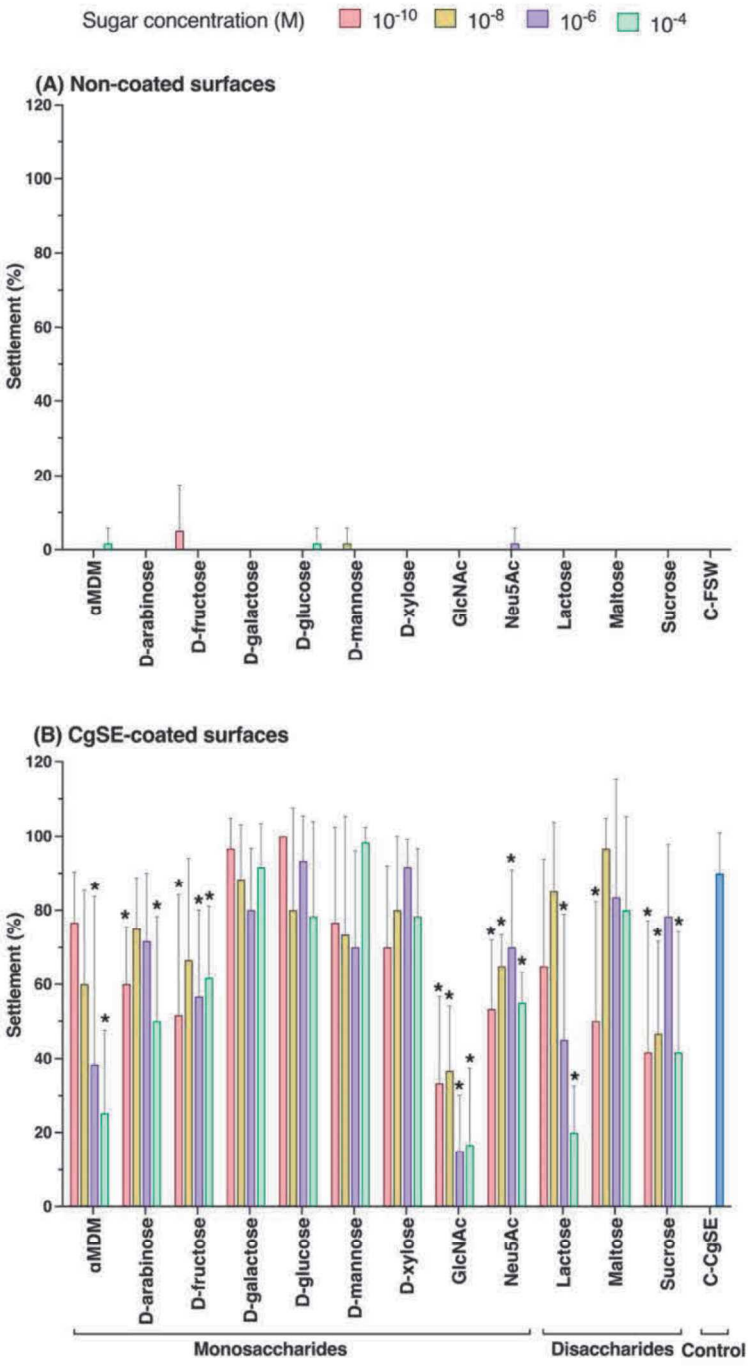

Figure 2.2. Settlement percentages of sugar treated *C. gigas* larvae exposed to different mono- and di-saccharides at different concentrations (10^{-10}, 10^{-8}, 10^{-6}, 10^{-4} M). Settlement percentages in **(A)** non-coated surfaces and **(B)** CgSE-coated surfaces. Untreated oysters in filtered seawater (C-FSW) and untreated oysters exposed to shell extract only (C-CgSE) served as control. CgSE (50 µg) was applied to all coated surfaces. A sugar-specific statistical analysis was performed. Asterisks (*) indicate significantly different to untreated oysters exposed to FSW alone (C-FSW, Figure 2A) and CgSE alone (C-CgSE, Figure 2B), determined via quasi-binomial glm ($p < 0.05$). Missing bars in the figure indicate no settlement. Data are the means (SD) of six replicates using different batches of larvae.

However, in the presence of a conspecific cue, CgSE-coated surfaces (Figure 2.2B), settlement responses were found similar to the control group (C-CgSE) when it was compared to larvae exposed to D-galactose, D-glucose, D-mannose, and D-xylose, irrespective of concentration differences ($p > 0.05$). On the contrary, an inhibitory effect on the settlement was exhibited by GlcNAc and Neu5Ac-treated larvae across all concentrations when compared to the control group (C-CgSE, $p < 0.05$). GlcNAc-treated larvae showed the highest inhibition pattern among all the sugars. Moreover, larvae treated with other sugars showed both inhibiting and no effect depending on the concentration. Sugars and concentration levels that inhibited larval settlement were: αMDM at 10^{-6} and 10^{-4} M; D-arabinose at 10^{-10} and 10^{-4} M; D-fructose at all concentrations except at 10^{-8} M; lactose at 10^{-6} and 10^{-4} M; Maltose at 10^{-10} M; Sucrose at all concentrations except at 10^{-6} M ($p < 0.05$). No mortality was recorded in all sugar treated larvae. Sugar treated larvae that did not settle after 24 h exhibited closed-shell behavior.

II.3.3 Effect of Exposure Time to Sugars and CgSE on Oyster Larval Settlement Response

Settlement percentages of larvae exposed to sugars at short-duration (0.25 h), medium-duration (2 h), and prolonged-duration (24 h) in the presence of CgSE are shown in Figure 2.3. Based on the previous results on the sugar screening assay, some sugars that showed a significant inhibiting effect in the presence of a conspecific cue were selected to assess whether exposure time to a specific sugar influenced settlement. A sugar-specific analysis was done to compare the effect on settlement when the larvae were exposed to different sugar concentrations and sugar exposure times on CgSE-coated surfaces. A high settlement percentage was recorded in untreated larvae exposed to CgSE alone (C-CgSE). All larvae that did not settle in this treatment were found actively swimming after 24 h. In sugar treated larvae, it is interesting to note that, while all showed similar low settlement patterns at 2 h exposure time, only GlcNAc-, Lactose- and Neu5Ac-treated larvae were inhibited to settle. Sugar treatments on the larvae that were found inhibiting to settlement were GlcNAc at 2 h exposure at 10^{-6} and 10^{-4} M concentrations, Lactose at 0.25 h exposure at 10^{-6} M as well as at 2 h exposure at both 10^{-6} and 10^{-4} M concentrations, and Neu5Ac at 2 h exposure at 10^{-4} M concentration ($p < 0.05$). In terms of behavioral observations, all sugar treated larvae irrespective of sugar type, exhibited closed-shell behavior upon initial contact with dissolved sugars and ceased any external movement. However, swimming activity eventually resumed after an hour of exposure. Moreover, all sugar treated larvae, irrespective of exposure time and concentration, that did not settle after the 24 h incubation period exhibited closed-shell

behavior. No mortality was recorded in all treatments. No settlement was observed on untreated *C. gigas* larvae on the multi-well with FSW (C-FSW, blank control).

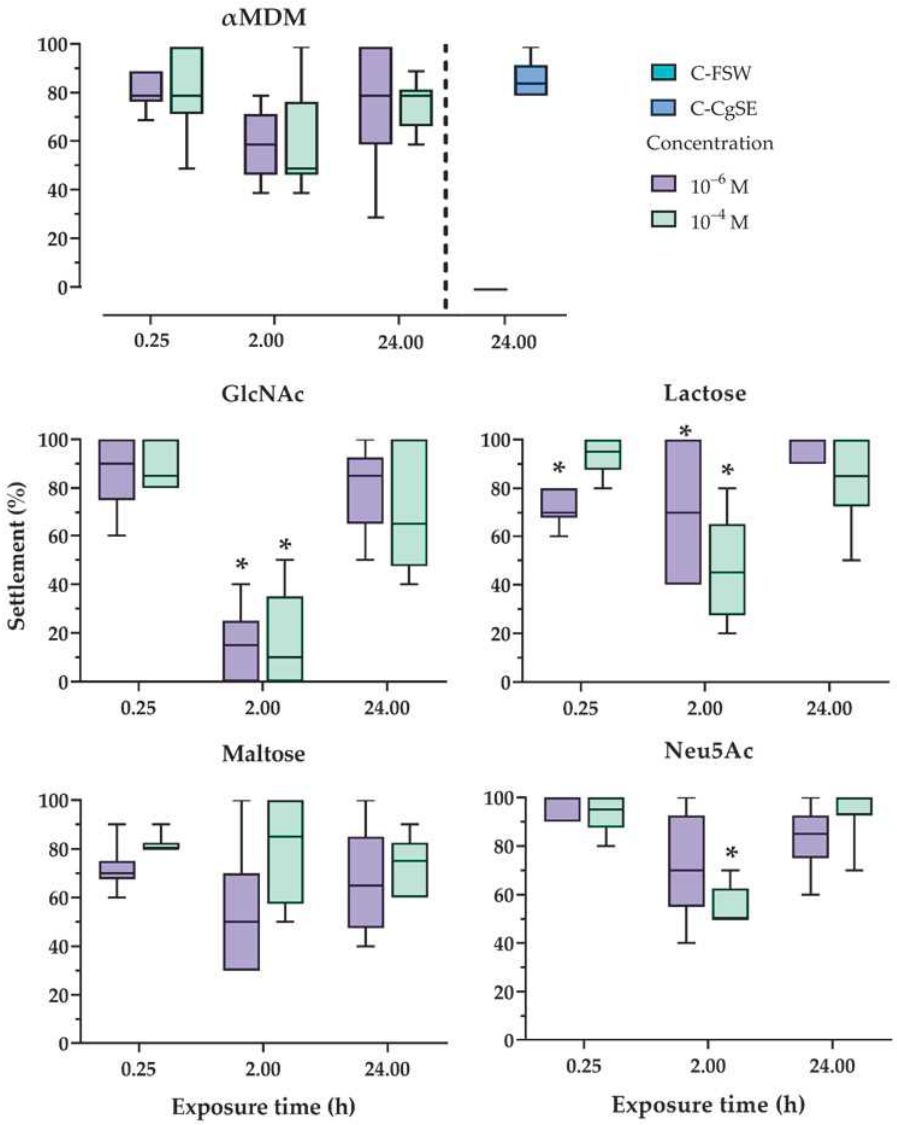

Figure 2.3. Settlement percentages of *C. gigas* larvae in response to different sugars and exposure times (0.25, 2, and 24 h) in the presence of CgSE. Following each exposure time, the larvae were thoroughly rinsed with filtered seawater and were incubated for 24 h in CgSE-coated wells. Larvae treated to a 24 h sugar exposure period was continuously immersed in CgSE-coated wells. Asterisks (*) indicate significantly inhibiting groups, determined via quasi-binomial glm ($p < 0.05$, $n = 6$, using different batches of larvae).

II.3.4 WGA Lectin Interaction with GlcNAc and Neu5Ac Binding: Effect on Oyster Larval Settlement

The settlement percentages of larvae on CgSE treated- GF/C papers in the presence, or absence, of wheat germ agglutinin (WGA) lectin, and its binding sugars GlcNAc and Neu5Ac, are shown in Figure 2.4. Co-exposure of CgSE papers and different concentrations of GlcNAc or Neu5Ac alone (Figure 2.4A and B, shaded bars) showed an inhibiting effect on the oyster larval settlement at 10^{-4} M when compared to the control (CgSE paper only, unshaded bars). In treatments with the mixture of WGA, and where increasing concentrations of GlcNAc or Neu5AC were added to CgSE papers, both showed an increasing trend of settlement with concentration (Figure 4A and B, striped bars). Quasi-binomial glm analysis showed that settlement, in the presence of WGA alone or its mixture with GlcNAc in CgSE papers was similar for all treatments ($p > 0.05$), while in WGA-Neu5Ac interaction, an inhibiting effect on the settlement was found at CgSE-Neu5Ac-WGA (50 µg: 10^{-8} M: 50 µg) mixture treatment ($p < 0.05$).

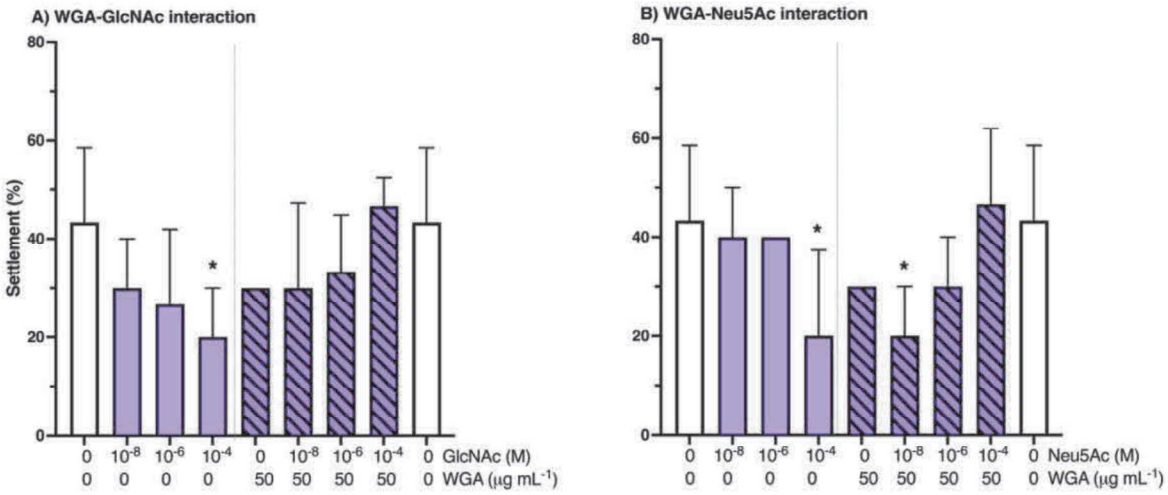

Figure 2.4. Settlement percentages in CgSE-treated GF/C filter papers when exposed to: WGA-binding sugars, (**A**) GlcNAc and (**B**) Neu5Ac, under varying concentrations of sugar treatment alone (shaded bars), and under WGA-GlcNAc or WGA-Neu5Ac mixture treatments (striped bars), for 2 h prior to assay. Adsorbed CgSE on GF/C filter papers alone served as a control (unshaded bars). Asterisks (*) denote a significant inhibiting effect on settlement, compared with other treatments determined via quasi-binomial glm ($p < 0.05$). Data are means (SD) of three replicates.

II.3.5 WGA-Binding Distribution on C. gigas Oyster Larval Tissues

The results on the distribution of WGA-FITC binding sites on *C. gigas* oyster larval tissues are shown in Figure 2.5. Under epifluorescence view, GlcNAc-treated (10^{-10} and 10^{-4} M) larvae showed decreasing intensity of WGA-FITC conjugated lectin staining (green fluorescence) on the velum, mantle, and foot tissues as the sugar concentration increases (Figure 2.5B and D). Surfaces of cilia present on the foot also showed WGA lectin binding (Figure D). This supports the hypothesis that GlcNAc sugar may have bound to these tissues. However, untreated larvae exposed to WGA-FITC also showed positive staining on the mantle and foot tissues (Figure 2.5F). This may indicate that, even without GlcNAc sugar exposure, the evidence points to a WGA-binding sugar present on these tissues as well. No fluorescence was detected on the tissues of untreated larvae without WGA-FITC exposure, which proves the absence of autofluorescence in *C. gigas* larvae.

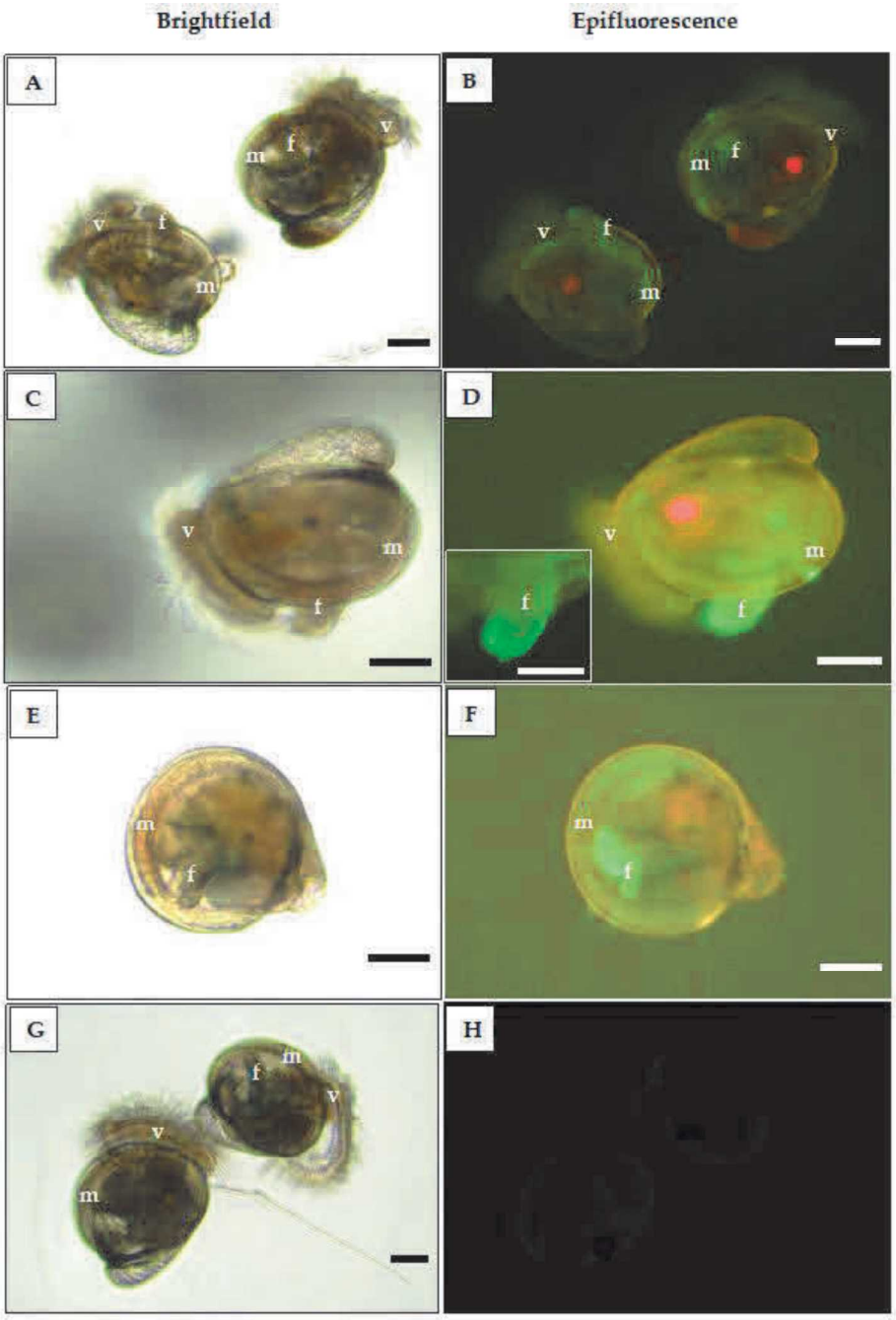

Figure 2.5. WGA binding to GlcNAc moieties on the mantle, foot, and velum tissues of *C. gigas* larvae stained green under epifluorescence view at 100× and 200× magnification (plate **D** inset only). **(A–B)** GlcNAc sugar treated larvae at 10^{-4} M show weakly stained mantle, foot, and velum tissues. (**C–D**) GlcNAc sugar treated oyster larva at 10^{-10} M shows an intense binding stain on the mantle, velum, and foot tissues, as well as its cilia (inset photo). (**E–F**) Untreated oyster larva (positive control) with WGA-FITC conjugated lectin shows binding stain on the mantle and foot tissues. (**G–H**) Untreated larvae (negative control) without WGA-FITC staining indicates that the larvae do not exhibit autofluorescence. Abbreviations: f = foot, m = mantle, v = velum. Scale bar = 100 µm.

II.4 Discussion

Although the active search for suitable substrate or settlement sites, in response to a wide range of variables and cues, has been well-documented in oyster larvae [1,2,8,9,11,16,23,35], the exact mechanism(s) of how larval settlements are induced is not yet fully understood. Understanding these processes and cues can help us understand the underlying mechanisms that shape population structures and chemical signaling. In the present study, a conspecific cue from the *C. gigas* shell-EDTA extract showed a high concentration-dependent settlement, inducing activity comparable to that of the hydrochloric acid (HCl) extract, as reported by Vasquez et al. (2013) [8]. Furthermore, the selection of a suitable substrate, i.e., inert, and non-toxic, for analyzing settlement-inducing compounds is important. In this experiment, no larval settlement was elicited from all the tested substrates in the absence of CgSE. No mortality was also observed, indicating that these substrates indeed were inert and non-toxic to the larvae. The adsorption of CgSE on GF/C, Polystyrene, and PVDF substrates resulted in similar larval settlement responses, which indicate that any of these substrates are suitable for use. Relatively similar results were also reported by Vazquez et al. (2014), where the highest number of post-larvae were recorded in GF/C and glass substrates [9]. Therefore, the direct application of the CgSE on polystyrene wells was used in the sugar screening and time exposure assays, while the GF/C filter papers were used in the lectin-sugar interaction assays. Fifty-microgram (50 μg) of CgSE was used for the succeeding larval settlement assays in this study.

Several studies have also shown that *C. gigas* larvae did not only settle through the mediation of conspecific cues but also settled in well-developed bacterial biofilms [20,23]. Bacterial biofilms are believed to possess chemicals bound to or released from in the form of extracellular polymeric substances [20–23]. The common components of extracellular and cell surface polysaccharides in bacterial biofilms are neutral sugars (D-glucose, D-galactose, D-mannose, D-fucose, L-rhamnose), amino sugars (*N*-acetylglucosamine and *N*-galactosamine), and some uronic acids [24–26]. The interactions between lectins and glycans have been implicated to play a critical role in biorecognition processes, such as in marine invertebrate larval settlement [9,36–38]. Herein, the results of this study demonstrate the influence of twelve types of dissolved sugars and a conspecific cue in *C. gigas* larval settlement. Larvae treated with different sugars in the absence of a conspecific cue, CgSE, did not settle. While untreated larvae exposed to CgSE alone resulted in a high settlement response. The results in this study suggest that the presence of dissolved sugars alone could be a chemotactic attractant to the larvae, but it does not provide the necessary metamorphic cue. Earlier studies have shown that

the larvae of *C. virginica* were attracted to bacterial metabolites, including glucose and amino acids, but were only induced to metamorphose in the presence of bacterial films [39]. Also, data in this present study indicates that induction of settlement in the larvae may not only require the presence of sugars alone, but also amino acids and other chemical groups, present at the site of chemical contact between the larvae and the substrate [40,41], as observed in the high settlement response on the control, CgSE alone (C-CgSE). This may also imply the need for multivalent interactions between multiple carbohydrate recognizing domains (CRDs) from lectins and multiple glycans that are often required to produce high-avidity binding interactions [41] for settlement induction to take effect. Hence, the high percentage of settlement on the conspecific cue, C-CgSE, could be in part, due to an increased avidity effect [41]. However, the precise mechanism of how lectins and glycans, in CgSE, enhance their avidity remains to be clarified. On the other hand, sugar treated larvae placed in CgSE-coated surfaces displayed varied settlement responses, most of which were found inhibitory (Figure 2.2B). One possible explanation for this could be that these low molecular weight sugars may have inhibited lectin-mediated processes as these sugars could have competed for the carbohydrate-binding sites on the CgSE or may have bound to the oyster larval lectin-like receptors [42]. In the present study, while several sugars were found inhibitory, their levels of inhibition were different, suggesting varying degrees of affinity of such sugars on the oyster larval chemoreceptors. Khandeparker et al. [31] and Neal and Yule [29], in separate experiments, inhibited cyprid larval settlement on surfaces with conspecific extract or the presence of barnacle species by exposing the larvae to different dissolved sugars. Their results indicated that sugars in solution can block the polar groups on the attachment disc through electrostatic adsorption of -OH groups. Higher sugar concentrations block more polar groups, nullifying their contribution to adhesion, thereby, lowering adhesion thresholds below those for cohesive failure [29]. Moreover, compounds that inhibit the reaction examined at the lowest concentration is considered to have the highest affinity for the lectin and to be most complementary to its combining site [28]. Among those sugars in this study that were significantly inhibitory were α-methyl-D-mannoside (αMDM), D-arabinose, D-fructose, *N*-acetylglucosamine (GlcNAc), *N*-acetylneuraminic acid (Neu5Ac), and all disaccharides (Lactose, Maltose, Sucrose). GlcNAc and Neu5Ac-treated larvae were significantly inhibited to settle on CgSE-coated surfaces at all concentrations suggesting that these sugars may have a very high affinity for the larval lectin-like receptor in the larvae. However, results in this study show that a lectin-like receptor in the larvae seems to bind with sugars from different specificity groups: hexose sugar (αMDM), pentose sugar (D-arabinose), ketose sugar (D-fructose), amino sugar (GlcNAc), sialic acid (Neu5Ac), reducing sugars

(lactose, maltose) and non-reducing sugar (sucrose). Whether these sugars bind to the lectin on the same combining site, or this lectin can combine simultaneously to different sugars, or exhibit ligand multivalency [28], remains to be clarified. On the other hand, D-galactose, D-glucose, D-mannose, and D-xylose-treated larvae were found to elicit similar responses compared to untreated larvae exposed to CgSE alone (C-CgSE). Also, in this study the influence of sugars –whether as monomers or dimers- and their concentrations have shown varying responses in oyster settlement. Many extracellular signal molecules act at very low concentrations (typically $\leq 10^{-8}$ M), and their receptors usually bind them with high affinity (dissociation constant Kd $\leq 10^{-8}$ M) [43]. The differential effects of sugar types on attachment to a substrate until settlement and metamorphosis suggests that the ratios of these sugars within a settlement inducing chemical cue could determine the maximum adhesion of a settling larvae on a substrata [44]. The effect of sugars on settlement are significant findings in the study of chemical communication and could have substantial implications for oyster ecology.

In this experiment, sugars that showed a concentration dependent and/or significant inhibiting effect in the presence of a conspecific cue were selected to assess whether exposure time to a specific sugar influence settlement. The results in this study demonstrate that the binding affinity of GlcNAc, Neu5Ac, and lactose sugars is influenced not only by concentration but also by exposure time. GlcNAc sugars showed the highest affinity, maximum inhibition occurring at 2 h exposure time at both $10^{-6,}$ and 10^{-4} M concentrations. Neu5Ac sugar show lesser affinity, maximum inhibition occurring at 2 h exposure time at 10^{-4} M concentration. While lactose showed the maximum inhibitory effect at 0.25 h exposure at 10^{-6} M concentration, as well as at 2 h exposure, at both 10^{-6} and 10^{-4} M concentrations. One possible factor that might be attributed to the low inhibitive effect at 0.25 h exposure time in almost all sugars could be the ability of oysters to close their shells in response to stress [45], thereby, nullifying their inhibitive effect on the ability of the larvae to discriminate the presence of a conspecific cue. Although, it is interesting to note that even in this short duration, lactose has shown an inhibitive effect in the settlement at a relatively lower concentration, suggesting yet an unreported role this sugar might play in settlement induction of *C. gigas*. At 2 h exposure time, while only GlcNAc, lactose, and Neu5Ac were significantly inhibitive, all the other sugars, i.e., αMDM and Maltose, showed a lower settlement trend in larvae in the presence of a conspecific cue, CgSE. It was observed that regardless of sugar type, all larvae upon contact with a sugar solution, resumed swimming activity after more than an hour of exhibiting closed-shell behavior. It is possible that through this period, these dissolved sugars could have bound to the larval external chemoreceptors, hence, interfering with lectin-mediated processes

involved in recognizing settlement inducing cues. The observation that larvae subjected to a 24 h co-exposure with sugar and conspecific cue were not significantly inhibited to settle in all sugar types could be attributable to the ability of oysters to produce high amounts of mucus under diverse forms of stresses [46]. Mucus contains mucoproteins associated with carbohydrates that are heavily glycosylated (up to 90% of carbohydrate) and present short carbohydrate chains whose charges are slightly negative [46]. Mucus matrices have been found to also contain repetitive highly sulfated polysaccharide and have been shown to play a role in cleaning body surfaces with unwanted substances [46]. In this study, an elevated mucus production could have been initiated as the larvae were exposed to dissolved sugars for a prolonged duration and might have competitively bound carbohydrate moieties on the surface of the larval external chemoreceptors. Moreover, the excessive dissolved sugars may have been subsequently washed out from its system. Hence, eliminating the possible interference of these sugars from negatively influencing the recognition of the settlement inducing cue from adult shell extract in *C. gigas*.

The results from the WGA-sugar binding experiments showed that blocking the carbohydrate moieties of CgSE-treated GF/C papers by dissolved GlcNAc sugars decreased larval settlement in *C. gigas* larvae in a concentration-dependent manner and was found to be inhibiting at 10^{-4} M, confirming the findings of Vazquez et al. (2014) on conspecific shell HCl extract [9]. They also reported in that study that this shell conspecific HCl extract showed increasing positive binding to WGA-FITC conjugate with increasing intensity as the concentration of the extract increased. Hence, one possible reason for this pattern could be that the increased sugar concentration on the GF/C paper could have competed for the carbohydrate-binding sites on the CgSE, such that it suppressed its settlement inducing effect on the larvae. An alternative explanation could be that the shell CgSE may also contain a WGA lectin-like receptor that could have specificity for the GlcNAc sugars. Foulon et al. (2019) have hypothesized that perlucin-like proteins containing C-type lectin domains, present in the periostracum of *C. gigas*, mediate linkage between the adhesive, secreted by the larvae during cementation, and the periostracum [47]. Whether this lectin, or any other lectin found in the shell matrix, may contain a WGA-sugar binding carbohydrate recognition domain, and is directly involved in larval settlement induction, remains to be investigated. Another possible factor is that the excessive amount of dissolved GlcNAc sugars may have bound to an oyster WGA lectin-like receptor, thereby, decreasing its ability to identify the signal from the conspecific cue in the extract. In a previous study by Vasquez et al. (2014), they postulated that the conspecific shell glycoprotein-induced settlement of *C. gigas* larvae are also mediated by

oyster WGA lectin-like receptors [9]. Vasquez et al. (2014) have demonstrated that treating oyster shell HCl extract-treated GF/C papers with WGA effectively inhibited larval settlement in a concentration dependent manner. However, its effect on the presumed settlement inducer in shell HCl extract was reversible, as shown by a significant increase in the larval settlement after co-treatment with increasing concentrations of GlcNAc sugars [9]. The interaction of lectins with sugars is reversible since their interaction does not result in the formation of covalent bonds [28]. A similar trend of lectin binding interaction on the larval settlement was also observed in this study. Also, this study reports for the first time the involvement of Neu5Ac sugars in inhibiting the settlement inducing activity of a conspecific cue from CgSE which also showed a similar concentration dependent inhibiting pattern as the GlcNAc sugar. Both GlcNAc and Neu5Ac residues are known to bind with Wheat germ agglutin (WGA) lectin [48,49] which could interact with carbohydrates from different monosaccharide specificity groups at the same combining site [28]. GlcNAc residues found in oligosaccharides are usually situated in an internal position while Neu5Ac is commonly found at the terminal end [28]. Another possible reason for the ability of WGA to identify Neu5Ac is due to the structural similarity of this monosaccharide to *N*-acetylglucosamine (GlcNAc) [28]. In this study, the addition of GlcNAc and Neu5Ac sugars competitively attenuated the inhibiting effect of WGA on the presumed settlement inducer on CgSE-treated GF/C papers. This evidence points to the possible involvement not only of GlcNAc sugars, but also that of Neu5Ac residues, present in the presumed larval settlement inducer from a conspecific cue in shells of *C. gigas*.

This study also shows evidence of dissolved GlcNAc sugar binding to WGA lectin-like receptors distributed on different *C. gigas* larval tissues-tagged by WGA-FITC conjugate. Observations under the epifluorescent microscope showed more intense binding of WGA-FITC conjugate on the foot, mantle, and velum tissues in the larvae treated with 10^{-10} M GlcNAc solution while at a higher concentration, 10^{-4} M, resulted in a weak binding effect. In contrast, Khandeparker et al. (2011) have demonstrated that as the sugar concentration increased, glucose-treated barnacles exhibited increasing glucose-binding fluorescence with FITC-conjugated lectin, Concanavalin A, on the third antennular segment [31]. Earlier studies have suggested that the lentil lectin-binding sugar chains (glucose and mannose) identified in the settlement inducing protein complex (SIPC) pheromone, and which serves as a cue for its gregarious settlement behavior, is produced by unicellular glands and is secreted onto the antennular discs [30,31,37,50]. In the present study, one possible factor that could be attributed to the weakly bound stain on the oyster larval tissues at a higher concentration, 10^{-4} M, could be the presence of an increased amount of mucus secretion. Although, no mortality was

recorded in the sugar exposure assay, an increased sugar concentration may have caused stress on the larvae, prompting it to produce more mucus, in order to remove excessive amounts of unwanted compounds in the body. The presence of mucus may have competitively bound to sites on the larval tissues recognized by the WGA-FITC conjugate. In a previous study of Espinosa et al. (2010), they have shown that pre-incubating microalgae with mucus from the pallial organs of *C. gigas* significantly reduced the binding effect of commercial FITC lectins to their surface ligands found in the microalgae [51]. Moreover, it is interesting to note that the non-sugar treated larvae tagged with WGA-FITC conjugate, in this study, showed positive binding on the foot and mantle tissues, suggesting the presence of endogenous GlcNAc, Neu5Ac, GalNAc residues [28] on the surfaces of these tissues. Also, no fluorescence was observed on the velum of the non-sugar treated larvae. A possible explanation for this could be that provided by Foulon et al. [52], where they demonstrated that a gland found on the *C. gigas* larval foot, similar to that reported on *Ostrea edulis* [53,54], showed acidic and neutral polysaccharide or glycoprotein content, and was also reported to play a role during the crawling phase. Additionally, the presence of endogenous WGA-binding sugar found on the larval foot and mantle, in the present study, further confirms one of the hypotheses of Vasquez et al. (2014). Their hypothesis explains that, during exposure of pediveliger larvae with WGA lectin treatment, this lectin may have to bound to shell matrix glycoproteins in the internal organs and mucopolysaccharides on the foot of the larvae [53,55–58], and thus, somehow interfered with *C. gigas* larval settlement behavior [9]. Shell matrix proteins, which are made by the mantle, interact with each other or with polysaccharides or chitin to produce the shell framework [59]. The settlement inducing protein from the conspecific shells may be found in the mantle and other multiple organs in *C. gigas* [3–5]. Although this warrants further investigation. Also, an earlier study demonstrated that the mantle specifically, its mantle margin, was actively involved during shell cementation in the settlement phase by adpressing the shell margin onto the substrate [61]. Moreover, the new findings in this present study suggest that a WGA lectin-like receptor and its endogenous ligand are both found in the larval chemoreceptors (foot, mantle, and velum). At the same time, the conspecific cue from the shell CgSE is also hypothesized to contain both, a settlement inducing endogenous ligand, and an endogenous lectin-like receptor that may complementarily work together to allow the larvae greater selectivity during site selection. The presence of bound endogenous lectin and its binding carbohydrate moieties in both the shell extract and the larvae could provide multiple binding sites and ensure higher affinity, selectivity, and complementarity [28], which could be advantageous for oyster larval growth and survival.

II.5 Conclusion

This study reports a conspecific cue from the *C. gigas* shell-EDTA extract, which exhibits a concentration-dependent settlement, thereby, inducing activities on all tested substrates, i.e., GF/C, polystyrene, and polyvinylidene difluoride. It also demonstrates the involvement of carbohydrate-lectin interaction and the important role played by sugar compounds in settlement site selection by *C. gigas* larvae. Sugar compounds can influence the ability of the larvae to identify and discriminate suitable substrata in response to chemical cues in either inhibiting or enhancing lectin-mediated processes involved in biorecognition. Sugars in the absence of a conspecific cue, *C. gigas* adult shell extract, did not promote settlement whereas, in the presence of the cue showed varied effects, most of which were found inhibitory at different concentrations. The binding affinity of selected inhibitory sugars such as GlcNAc, Neu5Ac, and lactose is influenced not only by its concentration but also by exposure time, with 2 h duration eliciting the maximum inhibitive effect on oyster larval settlement in the presence of a conspecific cue. Blocking of WGA-binding sites on the shell extract significantly reduced its settlement induction efficiency and the presence of GlcNAc sugar binding to WGA lectin-like receptors distributed on different *C. gigas* larval tissues-tagged by WGA-FITC conjugate was confirmed. New findings in this present study suggest that a WGA lectin-like receptor and its endogenous ligand are both found in the larval chemoreceptors (foot, mantle, and velum). At the same time, the conspecific cue from the shell CgSE has also been hypothesized to contain both a settlement-inducing endogenous ligand, as well as an endogenous lectin-like receptor that may complementarily work together to allow the larvae greater selectivity during site selection. This present work provides new insights into how carbohydrate-lectin interaction could play an important role in the natural environment when larvae encounter different glycoproteins adsorbed to surfaces and are also exposed to dissolved cues. It opens new possibilities for further analysis to help us understand the underlying mechanisms that shape population structures and chemical signaling.

II.6 References

1. Hadfield, M.; Paul, V. Natural Chemical Cues for Settlement and Metamorphosis of Marine Invertebrate Larvae. In: Marine Chemical Ecology; 2001; 431–461, ISBN 9781420036602.
2. Prendergast, G.S. Settlement and Behaviour of Marine Fouling Organisms. In Biofouling; Wiley-Blackwell: Oxford, UK, 2010; 30–59, ISBN 9781444315462, doi:10.1002/9781444315462.ch3.
3. Hadfield, M.G. Biofilms and Marine Invertebrate Larvae: What Bacteria Produce That Larvae Use to Choose Settlement Sites. Ann. Rev. Mar. Sci. 2011, 3, 453–470, doi:10.1146/annurev-marine-120709-142753.
4. Cranfield, H.J. Observations on the Behaviour of the Pediveliger of *Ostrea edulis* during Attachment and Cementing. Mar. Biol. 1973, 22, 203–209, doi:10.1007/BF00389174.
5. Coon, S.L.; Bonar, D.B.; Weiner, R.M. Induction of Settlement and Metamorphosis of the Pacific Oyster, *Crassostrea gigas* (Thunberg), by L-DOPA and Catecholamines. J. Exp. Mar. Bio. Ecol. 1985, 94, doi:10.1016/0022-0981(85)90059-0.
6. Coon, S.L.; Walch, M.; Fitt, W.K.; Weiner, R.M.; Bonar, D.B. Ammonia Induces Settlement Behavior in Oyster Larvae. Biol. Bull. 1990, 179, 297–303, doi:10.2307/1542321.
7. Joyce, A.; Vogeler, S. Molluscan Bivalve Settlement and Metamorphosis: Neuroendocrine Inducers and Morphogenetic Responses. Aquaculture 2018, 487, 64–82, doi: 10.1016/j.aquaculture.2018.01.002.
8. Vasquez, H.E.; Hashimoto, K.; Yoshida, A.; Hara, K.; Imai, C.C.; Kitamura, H.; Satuito, C.G. A Glycoprotein in Shells of Conspecifics Induces Larval Settlement of the Pacific Oyster *Crassostrea gigas*. PLoS One 2013, 8, e82358, doi: 10.1371/journal.pone.0082358.
9. Vasquez, H.E.; Hashimoto, K.; Kitamura, H.; Satuito, C.G. Wheat Germ Agglutinin-Binding Glycoprotein Extract from Shells of Conspecifics Induces Settlement of Larvae of the Pacific Oyster *Crassostrea gigas* (Thunberg). J. Shellfish Res. 2014, 33, 415–423, doi:10.2983/035.033.0210.
10. Crisp, D.J. Chemical Factors Inducing Settlement in *Crassostrea virginica* (Gmelin) J. Anim. Ecol. 1967, 36, 329–335, doi: 10.2307/2916.
11. Poirier, L.A.; Clements, J.C.; Davidson, J.D.P.; Miron, G.; Davidson, J.; Comeau, L.A. Sink before You Settle: Settlement Behaviour of Eastern Oyster (*Crassostrea virginica*)

Larvae on Artificial Spat Collectors and Natural Substrate. Aquac. Reports 2019, 13, doi:10.1016/j.aqrep.2019.100181.

12. Tamburri, M.N.; Luckenbach, M.W.; Breitburg, D.L.; Bonniwell, S.M. Settlement of *Crassostrea ariakensis* Larvae: Effects of Substrate, Biofilms, Sediment and Adult Chemical Cues. J. Shellfish Res. 2008, 27, 601–608, doi:10.2983/0730-8000(2008)27[601:SOCALE]2.0.CO;2.
13. Bayne, B.L. The Gregarious Behaviour of the Larvae of *Ostrea edulis* L. at Settlement. J. Mar. Biol. Assoc. United Kingdom 1969, 49, 327–356, doi:10.1017/S0025315400035943.
14. Hirata, Y.; Tamura, Y.; Nagasawa, K. Influence of Presoaking Conditions of Spat Collectors in Seawater Containing Adult Pacific Oysters (*Crassostrea gigas*) on Their Larval Settlement. Nippon Suisan Gakkaishi 2008, 74, 1017–1023, doi:10.2331/suisan.74.1017.
15. Zimmer-Faust, R.K.; Tamburri, M.N. Chemical Identity and Ecological Implications of a Waterborne, Larval Settlement Cue. Limnol. Oceanogr. 1994, 39, 1075‐1087, doi:10.4319/lo.1994.39.5.1075.
16. Turner, E.J.; Zimmer‐Faust, R.K.; Palmer, M.A.; Luckenbach, M.; Pentchef, N.D. Settlement of Oyster (*Crassostrea virginica*) Larvae: Effects of Water Flow and a Water‐soluble Chemical Cue. Limnol. Oceanogr. 1994, 39, 1579‐1593, doi:10.4319/lo.1994.39.7.1579.
17. Carroll, J.M.; Riddle, K.; Woods, K.E.; Finelli, C.M. Recruitment of the Eastern Oyster, *Crassostrea virginica*, in Response to Settlement Cues and Predation in North Carolina. J. Exp. Mar. Bio. Ecol. 2015, 463, 1–7, doi:10.1016/j.jembe.2014.10.024.
18. Pascual, M.S.; Zampatti, E.A. Evidence of a Chemically Mediated Adult-Larval Interaction Triggering Settlement in *Ostrea puelchana*: Applications in Hatchery Production. Aquaculture 1995, 133, 33–44, doi:10.1016/0044-8486(95)00005-M.
19. Veitch, F.P.; Hidu, H. Gregarious Setting in the American Oyster *Crassostrea virginica* Gmelin: I. Properties of a Partially Purified “Setting Factor.” Chesap. Sci. 1971, 12, 173, doi:10.2307/1350777.
20. Fitt, W.K.; Labare, M.P.; Fuqua, W.C.; Walch, M.; Coon, S.L.; Bonar, D.B.; Colwell, R.R.; Weiner, R.M. Factors Influencing Bacterial Production of Inducers of Settlement Behavior of Larvae of the Oyster *Crassostrea gigas*. Microb. Ecol. 1989, 17, 287–298, doi:10.1007/BF02012841.

21. Fitt, W.K.; Coon, S.L.; Walch, M.; Weiner, R.M.; Colwell, R.R.; Bonar, D.B. Settlement Behavior and Metamorphosis of Oyster Larvae (*Crassostrea gigas*) in Response to Bacterial Supernatants. Mar. Biol. 1990, 106, 389–394, doi:10.1007/BF01344317.
22. Tamburri, M.N.; Zimmer-Faust, R.K.; Tamplin, M.L. Natural Sources and Properties of Chemical Inducers Mediating Settlement of Oyster Larvae: A Re-Examination. Biol. Bull. 1992, 183, 327–338, doi:10.2307/1542218.
23. Hashimoto, K.; Vasquez, H.E.; Kitamura, H.; Satuito, C.G. Larval Settlement of the Pacific Oyster *Crassostrea gigas* in Response to Marine Microbial Films and Monospecies Bacterial Films. Sess. Org. 2017, 34, doi:10.4282/sosj.34.21.
24. Christensen, B.E. The Role of Extracellular Polysaccharides in Biofilms. J. Biotechnol. 1989, 10, 181–202, doi:10.1016/0168-1656(89)90064-3.
25. Neu, T.R.; Marshall, K.C. Bacterial Polymers: Physicochemical Aspects of Their Interactions at Interfaces. J. Biomater. Appl. 1990, 5, 107–133, doi:10.1177/088532829000500203.
26. Weiner, R.; Langille, S.; Quintero, E. Structure, Function and Immunochemistry of Bacterial Exopolysaccharides. J. Ind. Microbiol. 1995, 15, 339–346, doi:10.1007/BF01569989.
27. Maki, J.S.; Mitchell, R. Involvement of Lectins in the Settlement and Metamorphosis of Marine Invertebrate Larvae. Bull. Mar. Sci. 1985, 37, 675–683.
28. Sharon, N.; Lis, H. Lectins. In Encyclopedia of Biological Chemistry; Elsevier, 2013, 701–705 ISBN 9780123786319.
29. Neal, A.L.; Yule, A.B. The Link between Cypris Temporary Adhesion and Settlement of *Balanus balanoides* (L.). Biofouling 1992, 6, 33–38, doi:10.1080/08927019209386207.
30. Khandeparker, L.; Anil, A.C.; Raghukumar, S. Exploration and Metamorphosis in *Balanus amphitrite* Darwin (Cirripedia; Thoracica) Cyprids: Significance of Sugars and Adult Extract. J. Exp. Mar. Bio. Ecol. 2002, 281, 77–88, doi:10.1016/S0022-0981(02)00414-8.
31. Khandeparker, L.; Anil, A.C. Role of Conspecific Cues and Sugars in the Settlement of Cyprids of the Barnacle, *Balanus amphitrite*. J. Zool. 2011, 284, 206–214, doi:10.1111/j.1469-7998.2011.00794.x.
32. Liu, C.; Li, S.; Kong, J.; Liu, Y.; Wang, T.; Xie, L.; Zhang, R. In-Depth Proteomic Analysis of Shell Matrix Proteins of *Pinctada fucata*. Sci. Rep. 2015, 5, 1–14, doi:10.1038/srep17269.
33. Messenger, J.B.; Nixon, M.; Ryan, K.P. Magnesium Chloride as an Anaesthetic for

Cephalopods. Comp. Biochem. Physiol. Part C, Comp. 1985, 82, 203–205, doi:10.1016/0742-8413(85)90230-0.

34. Draper, N.R.; Smith, H. Applied Regression Analysis, 3rd ed.; Wiley: Hoboken, NJ, USA, 1998; ISBN 9780471170822.
35. Finelli, C.M.; Wethey, D.S. Behavior of Oyster (*Crassostrea virginica*) Larvae in Flume Boundary Layer Flows. Mar. Biol. 2003, 143, 703–711, doi:10.1007/s00227-003-1110-z.
36. Kirchman, D.; Graham, S.; Reish, D.; Mitchell, R. Lectins May Mediate in the Settlement and Metamorphosis of *Janua (Dexiospira) brasiliensis grube* (Polychaeta: Spirorbidae). Mar. Biol. Lett. 1982, 3, 131.
37. Matsumura, K.; Nagano, M.; Fusetani, N. Purification of a Larval Settlement-Inducing Protein Complex (SIPC) of the Barnacle, *Balanus amphitrite*. J. Exp. Zool. 1998, 281, doi:10.1002/(SICI)1097-010X(19980501)281:1<12::AID-JEZ3>3.0.CO;2-F.
38. Matsumura, K.; Mori, S.; Nagano, M.; Fusetani, N. Lentil Lectin Inhibits Adult Extract-Induced Settlement of the Barnacle, *Balanus amphitrite*. J. Exp. Zool. 1998, 280, 213–219, doi:10.1002/(SICI)1097-010X(19980215)280:3<213::AID-JEZ2>3.0.CO;2-Q.
39. Mitchell, R. Colonization by Higher Organisms. In Microbial Adhesion and Aggregation; Marshall, K.C., Ed.; Springer: Berlin, Germany, 1984; 189–200.
40. Tommasone, S.; Allabush, F.; Tagger, Y.K.; Norman, J.; Köpf, M.; Tucker, J.H.R.; Mendes, P.M. The Challenges of Glycan Recognition with Natural and Artificial Receptors. Chem. Soc. Rev. 2019, 48, 5488–5505, doi:10.1039/C8CS00768C.
41. Taylor, M.E.; Drickamer, K.; Schnaar, R.L.; Etzler, M.E.; Varki, A. Discovery and Classification of Glycan-Binding Proteins; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2015; ISBN 9780879697709.
42. Lis, H. Lectins as Molecules and as Tools. Annu. Rev. Biochem. 1986, 55, 35–67, doi:10.1146/annurev.biochem.55.1.35.
43. Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. Molecular Biology of the Cell; Wilson, J., Hunt, T., Eds.; 6th ed.; W.W. Norton & Company, 2017; ISBN 9781315735368.
44. Neal, A.L.; Yule, A.B. The Effects of Dissolved Sugars Upon the Temporary Adhesion of Barnacle Cyprids. J. Mar. Biol. Assoc. United Kingdom 1996, 76, 649–655, doi:10.1017/S0025315400031350.
45. Charifi, M.; Sow, M.; Ciret, P.; Benomar, S.; Massabuau, J.-C. The Sense of Hearing in the Pacific Oyster, *Magallana gigas.* PLoS One 2017, 12, e0185353,

doi:10.1371/journal.pone.0185353.

46. Pales Espinosa, E.; Koller, A.; Allam, B. Proteomic Characterization of Mucosal Secretions in the Eastern Oyster, *Crassostrea virginica*. J. Proteomics 2016, 132, 63–76, doi:10.1016/j.jprot.2015.11.018.
47. Foulon, V.; Boudry, P.; Artigaud, S.; Guérard, F.; Hellio, C. In Silico Analysis of Pacific Oyster (*Crassostrea gigas*) Transcriptome over Developmental Stages Reveals Candidate Genes for Larval Settlement. Int. J. Mol. Sci. 2019, 20, 197, doi:10.3390/ijms20010197.
48. Monsigny, M.; Sene, C.; Obrenovitch, A.; Roche, A. - C; Delmotte, F.; Boschetti, E. Properties of Succinylated Wheat - Germ Agglutinin. Eur. J. Biochem. 1979, 98, 39-45, doi:10.1111/j.1432-1033.1979.tb13157.x.
49. Petraccioli, A.; Maio, N.; Guarino, F.M.; Scillitani, G. Seasonal Variation in Glycoconjugates of the Pedal Glandular System of the Rayed Mediterranean Limpet, *Patella caerulea* (Gastropoda: Patellidae). Zoology 2013, 116, 186–196, doi:10.1016/j.zool.2012.10.006.
50. Matsumura, K.; Nagano, M.; Kato-Yoshinaga, Y.; Yamazaki, M.; Clare, A.S.; Fusetani, N. Immunological Studies on the Settlement-Inducing Protein Complex (SIPC) of the Barnacle *Balanus amphitrite* and Its Possible Involvement in Larva-Larva Interactions. Proc. R. Soc. B Biol. Sci. 1998, 265, 1825–1830, doi:10.1098/rspb.1998.0508.
51. Espinosa, E.P.; Perrigault, M.; Ward, J.E.; Shumway, S.E.; Allam, B. Microalgal Cell Surface Carbohydrates as Recognition Sites for Particle Sorting in Suspension-Feeding Bivalves. Biol. Bull. 2010, 218, 75–86, doi:10.1086/BBLv218n1p75.
52. Foulon, V.; Artigaud, S.; Buscaglia, M.; Bernay, B.; Fabioux, C.; Petton, B.; Elies, P.; Boukerma, K.; Hellio, C.; Guérard, F.; et al. Proteinaceous Secretion of Bioadhesive Produced during Crawling and Settlement of *Crassostrea gigas* Larvae. Sci. Rep. 2018, 8, 1–14, doi:10.1038/s41598-018-33720-4.
53. Cranfield, H.J. A Study of the Morphology, Ultrastructure, and Histochemistry of the Foot of the Pediveliger of *Ostrea edulis*. Mar. Biol. 1973, 22, 187–202, doi:10.1007/BF00389173.
54. Cranfield, H.J. Observations on the Function of the Glands of the Foot of the Pediveliger of *Ostrea edulis* during Settlement. Mar. Biol. 1973, 22, 211–223, doi:10.1007/BF00389175.
55. Johnstone, M.B.; Ellis, S.; Mount, A.S. Visualization of Shell Matrix Proteins in

Hemocytes and Tissues of the Eastern Oyster, *Crassostrea virginica*. J. Exp. Zool. Part B Mol. Dev. Evol. 2008, 310B, 227–239, doi:10.1002/jez.b.21206.

56. Johnstone, M.B.; Wheeler, A.P.; Falwell, E.P.; Staton, M.E.; Saski, C.A.; Mount, A.S. Folian-Cv1 Is a Member of a Highly Acidic Phosphoprotein Class Derived From the Foliated Layer of the Eastern Oyster (*Crassostrea virginica*) Shell and Identified in Hemocytes and Mantle. Front. Mar. Sci. 2019, 6, 1–21, doi:10.3389/fmars.2019.00366.
57. Davies, M.S.; Hawkins, S.J. Mucus from Marine Molluscs. Adv. Mar. Biol. 1998; 34, 1–71 ISBN 0120261340.
58. Jing, X.; Espinosa, E.P.; Perrigault, M.; Allam, B. Identification, Molecular Characterization and Expression Analysis of a Mucosal C-Type Lectin in the Eastern Oyster, *Crassostrea Virginica*. Fish Shellfish Immunol. 2011, 30, 851–858, doi:10.1016/j.fsi.2011.01.007.
59. Wang, X.; Li, L.; Zhu, Y.; Du, Y.; Song, X.; Chen, Y.; Huang, R.; Que, H.; Fang, X.; Zhang, G. Oyster Shell Proteins Originate from Multiple Organs and Their Probable Transport Pathway to the Shell Formation Front. PLoS One 2013, 8, e66522, doi:10.1371/journal.pone.0066522.
60. Cranfield, H.J. Observations on the Morphology of the Mantle Folds of the Pediveliger of *Ostrea edulis* L. and their Function during Settlement. J. Mar. Biol. Assoc. UK 1974, 54, 1–12.
61. Yamaguchi, K. Shell Structure and Behaviour Related to Cementation in Oysters. Mar. Biol. 1994, 118, 89–100, doi:10.1007/BF00699222.

Chapter III: Identification and Characterization of the Larval Settlement Pheromone Protein Components in Adult Shells of *Crassostrea gigas*: A Novel Function of Shell Matrix Proteins

III.1 Introduction

The Pacific oyster *Crassostrea gigas* (Thunberg, 1973) is a benthic bivalve mollusk that has been introduced worldwide and has become one of the most well-studied organisms due to its biological, ecological, and aquaculture importance [1,2]. Pacific oysters, like other oyster species, act as foundation species by creating habitats for other estuarine species and providing key ecosystem services to human communities around the world [3–5]. Pacific oysters have been widely used as a model species for studying molecular genetics [1,2,6], developmental biology [7–10], biomineralization [11,12], and adaptation to coastal environments under climate change [13]. Beck et al. [14] have reported that oyster populations have significantly declined on a global scale, with an estimated 85% of oyster reefs gone, far outnumbering the projected loss of coral reefs. Understanding how oysters, such as *C. gigas,* occupy appropriate substrata is crucial to learning about their evolution, population dynamics, and chemical communication.

Population dynamics are controlled by recruitment and post-settlement survival in marine invertebrates such as oysters. Finding a suitable settlement substrate is crucial for their survival, growth, and reproduction [15–17]. Oysters are also known to form reef-like structures characterized by monospecific aggregations. The benefits of such gregarious settlement behavior seem to prove beneficial in most sessile invertebrate species as this facilitates increased post-metamorphic survival [18,19] and synchronized fertilization [19,20]. Oyster substrate selection behavior is influenced by a variety of chemical cues that may come from bacterial biofilms, prey or predator, conspecifics, and macroalgal hosts [15,21].

Chemical signaling by a wide array of water-soluble or surface-bound cues such as pheromones associated with conspecific adults has been reported to influence gregarious behavior in oysters and induce larval settlement. Cole and Knight-Jones (1939) were the first to report on gregarious behavior in oysters specifically for *Ostrea edulis* and their work was later confirmed by the study of Bayne [22]. Bayne demonstrated that larval settlement could be induced by the application of *O. edulis* tissue extracts to a surface [22]. This eventually led to more studies investigating the effect of chemical cues from conspecifics and their

involvement in gregarious behavior for other oyster species. Such studies include shells in *C. gigas* [23–26], *C. virginica* [27,28], *C. ariakensis* [29]; soft tissue homogenates in *Ostrea edulis* [22]; water pre-conditioned by adults of *C. ariakensis* [29], *C. gigas* [30–32], *C. virginica* [33–35], and *O. puelchana* [36]; oyster shell liquor or extrapallial fluid of *C. virginica* [37]; and ammonia in *C. gigas* [38]. The chemical basis for such gregarious settlement behavior of oysters is not yet fully clarified and the identification and characterization of the active cues that drive such induction of conspecific cue-mediated larval settlement phenomenon have not yet been achieved. In this study, settlement is defined as the sequential transition of a competent pediveliger larva to a postlarva, which includes the attachment to a substrate and then eventual metamorphosis.

Diederich (2005) demonstrated through field experiments that larval recruitment was higher on its conspecifics, whether on *C. gigas* as a living substrate or on its dead shells [25]. In addition, Vasquez et al. [23] have reported an unidentified 55 kDa HCl-soluble organic matrix extract from *C. gigas* shells that may mediate larval settlement induction among its conspecifics, and which could only be visualized under the stains-all stain method. They have also demonstrated in another study the involvement of a WGA-binding sugar chain from the shell extract that could mediate conspecific cue settlement induction [24]. I have, therefore, tried to pursue the investigation of the nature and properties of this chemical cue by using a different chemical solution that has been widely used in extracting molluscan shell matrix proteins, i.e., Ethylenediaminetetraacetic acid (EDTA) [39,40]. My previous study has demonstrated that a chemical cue from adult shells (*C. gigas* EDTA-soluble shell extract, CgSE) was isolated that could induce larval settlement in a positive concentration-dependent manner and that GlcNAc and Sialic acid sugars from the chemical cue were found to interact with a WGA lectin-like receptor from the oyster larva [26].

Shells of *C. gigas* are made up of 99% calcium carbonate and around 0.5% occluded organic matrix [11]. This organic matrix is composed of a combination of proteins, glycoproteins, and polysaccharides that self-assemble and govern the shell's calcium carbonate polymorph (calcite, aragonite), crystallite size, shape, and texture [11]. Several molluscan shell matrix proteins (SMPs) have been recently identified in shells, including that of *C. gigas*, but much of the focus has been on their biomineralization roles [41,42]. One of the main challenges to studying SMPs is their inherent difficulty to isolate, characterize and obtain enough amount for functional analyses due to their close association with the mineral phase, highly charged nature, and presence of post-translational modifications (PTMs) [43,44]. To overcome this, a larval settlement-guided approach was developed to locate the signal molecule from *C. gigas*

conspecific shells, aided by various biochemical analyses using different gel staining methods and enzyme treatments, mass spectrometry analyses, bioinformatics analyses of protein sequences, and gene expression analysis of a selected gene between the *C. gigas* pediveliger and postlarvae. Previously identified chemical cues in other marine invertebrate species, i.e., barnacle, *Balanus amphitrite,* were shown to be induced by some polypeptides processed from a single precursor protein encoded in a single gene, which was designated to be the Settlement-Inducing Protein Complex (SIPC) [45,46]. However, a novel role of several shell matrix proteins that are acting as an aggregating settlement cue is demonstrated for the first time in this study and hereon named *Crassostrea gigas* Settlement Pheromone Protein Components (CGSPPC) that cooperatively function by at least two or more unrelated gene products. This study elucidates the identity, characteristics, and mechanisms underlying the conspecific cue-mediated induction of *C. gigas* larval settlement. A discussion is made on the properties and contribution of each protein component to settlement induction, the role of PTMs, and lastly, the molecular basis of the conspecific cue-mediated larval settlement mechanism.

III.2 Materials and Methods

III.2.1 Spawning and Larval Culture

The adult Pacific oysters (*Crassostrea gigas*) used in this experiment were purchased from Konagai Fisheries Cooperative, Nagasaki, Japan. The method for artificial fertilization and larval culture was based on Sedanza et al. [26,47]. Two to four male and five to eight female adult oysters were used as broodstocks for gamete stripping. The *C. gigas* larvae were cultured in 10-L filtered sea water-filled tanks under dark conditions with constant aeration. The seawater was renewed daily, and the larvae were fed with *Chaetoceros calcitrans* (10,000–50,000 cells/mL/day) from day 1 to day 5, a combination of *C. calcitrans* (25,000 cells/mL/day) and *Chaetoceros gracilis* (25,000 cells/mL/day) from day 6 to day 10, and then *C. gracilis* (50,000 cells/mL/day) from day 11 onward during the culture period. The water quality parameters during the larval culture period were salinity (30-32 psu), pH (8.02-8.08), and temperature (25-27 °C). The seawater was renewed daily prior to feeding. At least ten batches of different broodstocks were used for spawning to obtain larvae for the different settlement assays. Competent pediveligers used in settlement assays were when they reached an average shell height of 322 μm [26].

III.2.2 Extraction of Shell Matrices

Field observations on oyster reef structures indicate that oysters tend to attach to any part of the adult conspecific shells, whether it be on the outer surface or the underside. Hence, whole adult shell samples were crushed into smaller pieces and used for shell matrix extraction. The preparation of the *C. gigas* crude shell EDTA (CgSE) extract was done according to Sedanza et al. [26]. Shells from freshly shucked oysters (*C. gigas*) were thoroughly cleaned and washed with tap water to remove adhering epibionts and traces of muscle tissues, and then dried. Cleaned shells were then crushed with a hammer until the shell chips of about 0.5 to 1.0 mm in size were collected. These shell chips (150 g) were decalcified with 1 L of 0.8 M Ethylenediaminetetraacetic acid (EDTA, 99% purity, Nacalai Tesque, Inc., Kyoto, Japan, pH 8.0) for 60 h at 4 °C with continuous agitation. The supernatant was collected by centrifugation at 10,000 × g for 60 min at 4 °C. Then, it was subsequently filtered and dialyzed against distilled water at 4 °C for 3 days. The crude *C. gigas* shell EDTA-soluble extract (CgSE) was lyophilized, and the resultant powder was dissolved in as small a volume as possible of distilled water. This extract form is herein called Freeze-dried CgSE. Alternatively, another batch of dialyzed supernatant extract inside a dialysis tube was air-dried at a cool room temperature

until the volume was reduced and the supernatant concentrated. This extract form is herein called Air-dried CgSE. The Freeze-dried (FD)-CgSE extract was used in all larval settlement assays while the Air-dried (AD)-CgSE extract was used to compare with the FD-CgSE extract to check for protein stability during the extraction process. These two forms of CgSE extracts were stored at -20 °C until further use. The protein content of the extract was quantified by a BCA assay kit (Pierce, Thermo Fisher Scientific, Rockford, IL, USA) according to the manufacturer's protocol. A gram of shell chips could yield 500 μg of CgSE extract.

III.2.3 First Protein Separation Approach

III.2.3.1 Fractionation by Ultrafiltration

Ultrafiltration: Freeze-dried CgSE was fractionated into three molecular range fractions: >100K, >50K, and <50K molecular weight cut-off (MWCO) CgSE. This was done by separating CgSE into >100K and <100K MWCO fractions using an Amicon ultra-15 ultrafiltration tube (100,000 MWCO, Merck Co., Ltd, Darmstadt, Germany). Then the <100K fraction was applied to a Vivaspin 20 ultrafiltration tube (50,000 MWCO PES, Sartorius, Gloucestershire, United Kingdom) to separate it into >50K and <50K MWCO fractions. All the fractionated samples were centrifuged at 2500 x g for 40 minutes.

Larval Settlement Assay**:** Fractionated samples were subjected to a larval settlement assay at various extract amounts of 1 and 100 μg that were coated separately on the base of 6-well plates. Larval settlement assay was performed according to Sedanza et al. [26]. In brief, ten larvae were released into each pre-coated well plate filled with 10 mL filtered seawater (FSW). The larval settlement was confirmed by the number of individuals that metamorphosed to postlarvae within 24 h. Postlarva was confirmed under the microscope as individuals that secreted cement substances or those with post-larval shell growth. All succeeding larval settlement assays were conducted in a dark environment at 27 ± 1 °C in an incubator. Another batch of larvae was also released into the well plates containing filtered seawater (FSW) as the control. The larval settlement assays were repeated twice using different batches of larvae with three replicates at each trial (n = 6). All data were presented as the mean ± standard deviation (SD).

III.2.3.2 Gel filtration of >50K CgSE

Gel filtration chromatography: The >50K CgSE active fraction (6 mg) was applied to a Superdex 200 10/300 GL column (GE Healthcare, Tokyo, Japan) equilibrated with 0.15 M

NaCl using a 500 μL loop and eluted with the same buffer at a rate of 0.25 mL min^{-1}, using an AKTA explore 10S- FPLC system (GE Healthcare, Tokyo, Japan). A total of 35 tubes of 0.5 mL were collected and pooled into six molecular range fractions: F1, F2, F3, F4, F5, and F6.

Larval Settlement Assay: The Gel filtration eluted fractions were subjected to a larval settlement assay at various extract amounts of 1 μg (n = 3), 10 μg (n = 6), and 50 μg (n = 5). Larval assay replications for each extract amount varied due to the limited protein yield from the eluted fractions. Several 6-well plates were pre-coated with the test extracts and the larval assay was carried out as described earlier. All data were presented as the mean ± standard deviation (SD).

SDS-PAGE Analysis: For SDS-PAGE analysis on the active fractions >50K CgSE, F2, and F3, 10 μg of each sample under reducing conditions were loaded in each lane on a 10% acrylamide gel for 1 h at 200 volts [48]. A Pre-stained Protein Standard marker (Broad range, Bio-Rad, Hercules, California, USA) was also loaded on the gel. Following SDS-PAGE, all the protein bands were visualized by the Stains-all staining method following the manufacturer's instruction manual.

III.2.3.3 Separation of F2 proteins

Protein separation by SDS-PAGE: Sixty μg of the active fraction, F2, from gel filtration was applied per lane and resolved in duplicate on 10% SDS-PAGE gels under non-reducing and reducing (Laemmli) conditions at 200 V for 1 h [48]. One gel was stained with Stains-all (Sigma-Aldrich, Buchs, Switzerland), following the manufacturer's protocol, to reveal the positions of acidic and phosphorylated proteins. The unstained duplicate gel was equilibrated in a transfer buffer (containing 10% methanol, 5 mM Tris, 38 mM Glycine, pH 8.3) and shaken for 15 mins. A PVDF membrane (Sequi-Blot, Bio-Rad, Hercules, California USA), with the same size as the duplicate gel, was immersed in 100% methanol until it became translucent. Subsequently, the membrane was equilibrated in a transfer buffer following the same procedure as earlier described. Next, the equilibrated duplicate gel was transferred electrophoretically onto the PVDF membrane for 150 mins at 0.8 mA/cm3. Following the transfer, the membrane was immersed in a container containing ultrapure water and shaken for 15 mins. This step was repeated six times. Then, the six major bands of F2 under non-reducing conditions (35, 38, 43, 48, 60 kDa) and three major bands under reducing conditions (45, 48, 53 kDa) were cut out using the stained marker (Pre-stained SDS-PAGE Standards Broad range, Bio-Rad, Hercules, California, USA). The Stains-all stained gel was used as a guide. This process was repeated

several times until enough PVDF strips containing the isolated polypeptides were enough for larval settlement assay. Isolated polypeptide bands that showed settlement inducing activity were further subjected to Edman sequencing for identification.

Larval settlement assay: Three PVDF strips containing an isolated polypeptide were fastened side-by-side to the base of each well on a 6-well plate. This step was repeated for all polypeptides. Each well was then filled with 10 mL FSW followed by the immersion of ten larvae. The succeeding steps for the larval assay were carried out as described earlier. These settlement assays were repeated twice using different batches of larvae with three replicates at each trial (n = 6). All data were presented as the mean ± standard deviation (SD).

III.2.4 Second Protein Separation Approach

III.2.4.1 Wheat Germ Agglutinin (WGA) Agarose bound Lectin Affinity Chromatography

Previous protein purification steps in this study showed that the settlement inducing effect of the target proteins decreased as they were further separated into each polypeptide. Hence, a change of strategy was performed by subjecting crude CgSE to another fractionation by ultrafiltration, this time, using a 50K MWCO centrifugation tube to separate it into <50K and >50K fractions. The >50K CgSE fraction was applied to lectin affinity chromatography on WGA-agarose bound resin. WGA-agarose bound lectin with a protein concentration of 7 mg/mL gel and binding capacity of >8 mg GlcNAc/mL gel was obtained from Vector Laboratories (Burlingame, CA, USA). A 1.0 M NaCl buffer system was used for WGA lectin column binding. A modified method of Yang and Hancock was adopted in the preparation of lectin affinity columns [49]. WGA affinity column was prepared by adding 0.5 mL of the corresponding agarose-bound lectin to an empty 0.8 mL Micro Bio-Spin chromatography column (Bio-Rad, Hercules, California, USA). The WGA agarose bound spin column was equilibrated with 10 column volumes of binding buffer containing 0.05 M Tris-HCl, 1.0 M NaCl, pH 8.0. Subsequently, 3 mg of >50K CgSE was loaded on the column. After 10 mins. of the sample application, the unbound proteins were eluted with 6 column volumes of the binding buffer and collected and pooled as the Unbound fraction (UF). Then the captured glycoproteins, which bound to the immobilized WGA lectin, were eluted using 10 column volumes of a buffer containing 0.5 M GlcNAc, 0.05 M Tris-HCl, 1.0 M NaCl, pH 8.0. The eluted fraction was collected as the pooled, Bound Fraction (BF) component. Washing of the column was done by a repetition of loading binding buffer in the spin column and draining it out by centrifugation at 5000 rpm for 30 s for 10 times. The process of the sample application and elution was repeated twice to gather enough samples for larval settlement assays and

biochemical analyses. The UF and BF components were dialyzed overnight against 2-3 times buffer change containing 20 mM Tris-HCl (pH 8.0). The dialyzed UF and BF inside the dialysis tubes were subsequently concentrated by air-drying at cool room temperature until the original volume was reduced to a small volume. These were then stored at -20 °C until further use.

CgSE, >50K CgSE, UF, and BF protein samples were subjected to a larval settlement assay at various extract amounts of 0, 0.5, 5, and 50 μg that were coated separately on the base of 6-well plates. Each treatment was done in triplicates. The succeeding steps for the larval assay were carried out as described earlier.

III.2.4.2 SDS-PAGE Analysis of WGA Affinity Chromatography Eluted Fractions

The CgSE, >50K CgSE, UF, and BF protein samples were analyzed on a 12% TGX FastCast polyacrylamide gel (1.0 mm x 10 well) (Bio-Rad, Hercules, California, USA) with a loading amount of 10 μg for each sample. Pre-stained molecular weight markers (Precision Plus Protein Dual Xtra, Bio-Rad, Hercules, California, USA) were loaded alongside the samples. Samples were suspended in 2x Laemmli sample buffer (125 mM Tris-HCl (pH 8.5), 20% glycerol, 4% SDS, 10% 2-mercaptoethanol, and 0.0025% bromophenol blue) under reducing conditions in the presence of 2-mercaptoethanol [48]. SDS-PAGE was run using a Mini-Protean Tetra System (Bio-Rad, Hercules, California, USA) at 200 volts for 40 mins. Gels were stained with QC-Colloidal Coomassie stain (Bio-Rad, Hercules, California, USA), Silver stain (Silver Stain MS Kit, Fujifilm Wako Pure Chemical Corporation, Osaka, Japan), and Stains-all for putative calcium-binding and highly acidic, phosphorylated proteins [50,51]. Gel staining was performed according to the manufacturer's protocol. Gel images for figure presentation and band molecular weight determination were captured using a GELSCAN® laser scanner (iMeasure, Nagano, Japan). For purposes of figure presentation, another set of colored gel images was captured using a digital camera. The estimated molecular weight of each protein band in a sample was calculated using Gel-Pro Analyzer 4.0 software.

III.2.5 Glycoprotein Detection

Ten micrograms of the eluted WGA affinity chromatography components, i.e., UF and BF samples, were deglycosylated using 1 μL PNGase F (0.5 mU/μL, Takara Bio Inc., Kyoto, Japan) under denaturing conditions. Another batch of samples was treated with a mixture of 2 μL *O*-glycosidase (40,000 U/μL), 2 μL PNGase F (0.5 mU/μL), and 2 μL Neuraminidase (50 U/μL) (New England BioLabs Inc., Massachusetts, USA). All enzyme reactions were done

with an overnight incubation at 37 °C for 20 h following the manufacturer's protocol. Subsequently, the samples (1) >50K CgSE, (2) untreated, and (3-4) deglycosylated WGA affinity chromatography eluted fractions (UF and BF), were subjected to Cys-β-propionamidation following the method of Unno et al. (2018) [52] for the mass spectrometry analysis preparation. In brief, 5 μL of each protein (10 μg) sample was added to 5 μL of 2x SDS-sample buffer for Cys alkylation (125 mM Tris-HCl (pH 8.5), 20% glycerol, 4% SDS, 10% 2-mercaptoethanol, and 0.0025% bromophenol blue), and the protein solution was incubated at 95 °C for 10 min. Five microliters of 7 M acrylamide were added to the reduced protein solution and held for 1 h at room temperature. The resulting propionamidated protein sample solutions (15 μL/well) were applied to a 12% TGX FastCast polyacrylamide gel and performed SDS-PAGE [48] following the same conditions as described earlier. The gel was stained with Pro-Q Emerald 488 glycoprotein gel stain kit (Invitrogen, Carlsbad, USA) following the manufacturer's protocol. The stained gel was viewed and photographed using a CCD camera from FluoroPhorestar 3000 (Anatech, Tokyo, Japan). For protein identification through mass spectrometry analysis, the gel was double stained with QC-Colloidal Coomassie stain following the manufacturer's protocol.

For the glycoprotein detection, enzymatic deglycosylation with PNGase F only and with a mixture of *O*-glycosidase, PNGase F, and Neuraminidase was performed and resolved on a 12% SDS-PAGE gel. After which, the gel was stained with Pro-Q Emerald for detection of glycoproteins and double-stained with QC-CC stain for further mass spectrometry analysis.

III.2.6 Phosphoprotein Detection

CgSE in its freeze-dried (non-reducing and reducing conditions) and air-dried (reducing condition) forms, >50K CgSE, UF, and BF component samples were analyzed for phosphoprotein detection. Ten micrograms of each sample were applied to SDS-PAGE under the same conditions as described earlier (Methods III.2.5), except for the omission of the Cys-β-propionamidation step and with some slight modifications. In brief, after SDS-PAGE, the gel was stained by Pro-Q Diamond Phosphoprotein gel stain (Invitrogen, Carlsbad, CA, USA) according to the product's manual. The gel was viewed and photographed using a CCD camera from FluoroPhorestar 3000 (Anatech, Tokyo, Japan).

III.2.7 Protease and TFMS Treatment

To characterize further the stains-all stainable acidic protein components through biochemical analysis, CgSE was treated with various proteases and Trifluoromethanesulfonic acid (TFMS) for chemical deglycosylation. Two hundred micrograms of CgSE were treated separately by three proteases and performed according to the manufacturer's protocol. Twenty microliters of Trypsin (Trypsin singles, Proteomics grade, Sigma-Aldrich, Missouri, USA) were added to CgSE dissolved in 10% ACN, 25 mM NH_4HCO_3, to constitute a 1:100 enzyme to protein ratio. Twenty microliters of Chymotrypsin (Roche, Mannheim, Germany) were added to CgSE dissolved in 100 mM Tris-HCl, and 10 mM $CaCl_2$, pH 7.8. The ratio of Chymotrypsin enzyme to protein was at 1:100. Twenty-five microliters of Asp-N (Roche, Mannheim, Germany) were added to CgSE dissolved in 50 mM sodium phosphate buffer, pH 8.0. The ratio of Asp-N enzyme to protein was at 1:200. All the protease digests were incubated at 37°C for 16 h. For the TFMS treatment assay (Wako Pure Chemicals Co., Osaka, Japan), equal amounts of CgSE (0.5 mL) corresponding to 40 µg of protein and TFMS (0.5 mL) were mixed and incubated for 1 h on ice, following the manufacturer's instruction manual. The TFMS digest was dialyzed against distilled water to remove excess TFMS in the mixture as recommended by Vasquez et al. [23]. Forty micrograms from each protease and TFMS digests were loaded on a 12% polyacrylamide gel and analyzed for SDS-PAGE under reducing conditions [48] following the steps described earlier. All the protein bands were visualized under the Stains-all stain. These protein bands were further subjected to mass spectrometry analysis.

Endoproteinase, Glu-C (*S. aureus* V8 protease, Mass spectrometry grade, 20 µg/mL, Pierce, Illinois, USA) was added to 10 µg of >50K CgSE dissolved in 25 mM NH_4HCO_3 under different enzyme to protein ratios of 1:100, 1:50, 1:10. The Glu-C digests were incubated at 37°C for 16 h, loaded on a 12% polyacrylamide gel, and analyzed for SDS-PAGE under reducing conditions [48] following the steps described earlier. All the protein bands were visualized under the Stains-all stain.

For SDS-PAGE analysis on >50K CgSE, 10 µg of each sample under reducing conditions were loaded in each lane on a 10% acrylamide gel, for 1 h at 200 volts [48]. A Pre-stained Protein Standard marker (Broad range, Bio-Rad, Hercules, California, USA) was also loaded on the gel. Following SDS-PAGE, all the protein bands were visualized under the Stains-all stain.

III.2.8 Mass Spectrometry Analyses

CgSE extract and its varied protein-separated forms that were selected for mass spectrometry analyses were subjected to Cys-β-propionamidation before SDS-PAGE following the procedures described earlier (Methods III.2.5). After SDS-PAGE, the method reported by Yamaguchi [53] was followed for the in-gel enzymatic digestion, peptide extraction from the gel, and sample loading onto a matrix-assisted laser desorption/ionization (MALDI) target plate. MS and MS/MS spectra were obtained using a matrix-assisted laser desorption/ionization (MALDI) quadruple-ion trap (QIT) time-of-flight (TOF) mass spectrometer (AXIMA Resonance, Shimadzu, Kyoto, Japan) with 2, 5-dihydroxybenzoic acid (DHBA, Shimadzu, Kyoto, Japan) as the matrix in positive mode. MALDI-QIT-TOF mass spectra were externally calibrated using human angiotensin II (m/z 1046.54) and human ACTH fragment 18–39 (m/z 2465.20) in a ProteoMass Peptide and Protein MALDI-MS Calibration Kit (Sigma-Aldrich, St. Louis, MI, USA). Modification of the N-terminal peptide was analyzed by PMF and an MS/MS ion search using MASCOT ver. 2.3 (Matrix Science, London, UK) with an original database from (*C. gigas* 65,722 protein sequences; https://www.ncbi.nlm.nih.gov/Taxonomy/Browser/wwwtax.cgi?id=29159) accessed on 8 April 2022 in our MASCOT server. Search parameters used for PMF were the following: enzyme, Trypsin; fixed modifications, propionamide (C); variable modifications, oxidation (HW and M); mass values, monoisotopic; peptide mass tolerance, +/−0.5 Da; peptide charge state, 1+; and a maximum number of missed cleavage sites, 1. Search parameters used for the MS/MS ions search were as follows: enzyme, Trypsin; fixed modifications, propionamide (C); variable modifications, oxidation (HW and M); mass values, monoisotopic; peptide mass tolerance, +/−0.3 Da; fragment tolerance, +/−0.2 Da; max missed cleavages, 1. A protein score (PMF) > 51 and individual ion score (MS/MS ions search) >14 were considered significant ($p < 0.05$).

III.2.9 Localization of N-glycosylation sites by Mass Spectrometry Analysis

Mass spectrometric localization of N-glycan sites of Gigasin-6 isoform X1 and/or X2 912 (66 kDa) was done by comparing the MS and MS/MS spectra profiles of the deglycosylated 913 (DG2/64 kDa) and glycosylated (B2/66 kDa) bands of interest, i.e., with and without 914 PNGase F treatment, respectively. The method for the determination of actual N- 915 glycosylation sites was adopted from León et al. [54] with slight modifications. In brief, before SDS-PAGE, UF and BF components were treated with PNGase F, reduced, alkylated, and

further processed as described earlier (Methods III.2.5). Then, the gels were stained with QC-Colloidal Coomassie stain and subjected to in-gel enzymatic digestion and mass spectrometry analysis. The full range MS spectra of both glycosylated and deglycosylated bands were compared and searched for a newly appeared signal on the m/z region of the deglycosylated digest. This newly appeared signal corresponds to a positive mass shift of approximately 0.98 da for every previously glycosylated asparagine that was deaminated to aspartic acid. All newly appeared signals were further subjected to MS/MS and Collision-Induced Dissociation (CID, 250), and the resulting analyzed peptide fragments were analyzed against an error-tolerant search to confirm their identity with matched protein sequences reported in Gigasin-6 isoform X1 and/or X2 and those found in our MASCOT server. Search parameters were as follows: enzyme, Trypsin; fixed modifications, propionamide (C); variable modifications, oxidation (HW and M); mass values, monoisotopic; peptide mass tolerance, +/−0.5 Da; peptide charge state, 1+; and a maximum number of missed cleavage sites, 1. Alternatively, newly appeared signals from the deglycosylated digest that did not get an error-tolerant search hit were manually annotated by matching observed MS/MS ion fragments with the Protein Prospector MS-Product program (https://prospector.ucsf.edu/prospector/cgi-bin/msform.cgi?form=msproduct, accessed on 31 March 2022). Manual validation of identified unique peptides took into account the assignment of the major peaks, the occurrence of uninterrupted y- or b- ion series of at least four 4 consecutive amino acids, the possible presence of a2/b2 ion pairs and immonium ions, and mass accuracy [55].

III.2.10 Characterization of Protein Sequences

Putative sequences identified through a mass spectrometry analysis were characterized using different bioinformatic tools. Signal peptide prediction was conducted with SignalP 6.0 [56]. Functional domains were searched using the InterProscan and ScanProsite platforms [57,58]. SMART (simple modular architecture research tool) was used to allow rapid identification and annotation of signaling domain sequences [59,60]. Transmembrane helices were predicted and proteins containing transmembrane helices lower than 18 or found within the signal peptide region were filtered out according to the Transmembrane Helices Hidden Markov Model (TMHMM) [61]. Phosphorylation sites in the sequences were predicted using NetPhos 3.1 [62]. *N*- and *O*-glycan sites were predicted using Net*N*Glyc-1.0 and Net*O*Glyc-4.0, respectively [63,64]. Prediction of intrinsically unstructured proteins and detection of low complexity regions (LCRs) were done using IUPred3 software [65]. Both identified and unidentified groups of protein sequences in the mass spectrometry analysis were also Blastp-

searched against nr in NCBI using default settings (https://blast.ncbi.nlm.nih.gov/Blast.cgi/, accessed on 31 March 2022). After mass spectrometry analysis, a homology search was performed on the unidentified stains-all stainable acidic proteins with a putative homolog in *C. gigas* to Folian-cv1 of *C. virginica* [40]. Multiple sequence alignments were done on various putative homolog groups using Clustal Omega (https://www.ebi.ac.uk/Tools/msa/clustalo/, accessed on 31 March 2022) [66].

III.2.11 Molecular Characterization of Gigasin-6 isoform X1

To examine the mRNA expression pattern of Gigasin-6 isoform X1 gene in the larvae before (pediveliger) and after conspecific cue induction (postlarvae), a larval settlement assay was conducted using the unfractionated CgSE sample and a quantitative real-time PCR (qRT-PCR) assay was performed.

Larval sample preparation was done following the method reported by Sedanza et al. [47] by using two batches of larvae: 1) Control, pediveliger larvae before conspecific cue settlement induction (which means no CgSE exposure), and 2) Postlarvae, conspecific cue-induced larvae (which means with CgSE exposure). Larvae were at 24 days post-fertilization (or 24 days of culture, DOC) when approximately 70–80% have developed an eyespot and started to crawl. To ensure that all the larvae were at a competent stage, larvae from multiple 10-L FSW-filled tanks were sieved over a 300 μm mesh-covered tip nestled on a glass bowl filled with FSW. They were carefully washed in Millipore FSW (0.22 μm) and transferred to several 3-L beaker containers newly filled with FSW. After filtration, more than 95% of the larvae were verified under the microscope to have reached the pediveliger stage. The larval samples were then divided into two batches. Beakers containing the control batch of competent pediveliger larvae were siphoned over a 300 μm mesh-covered tip nestled on a glass bowl filled with FSW, carefully washed with distilled water, and collected as a pooled Pedi sample. Approximately 100,000 individuals of pooled pediveliger larvae (183 mg wet weight) in a microcentrifuge tube were treated immediately with RNAlater Stabilization Solution (Invitrogen, Vilnius, Lithuania) for 24 h at 4 ◦C following the manufacturer's protocol. After which, RNAlater was removed and the pooled Pedi batch was stored at −80 ◦C prior to the subsequent experiments. For the PL batch, the method previously described by Sedanza et al. [26] was adopted for the settlement induction of oyster larvae with CgSE. Briefly, 6-well plates were pre-coated with 50 μg CgSE and subsequently filled with 10 mL of filtered seawater per well. Twenty larvae were released in each well and induced to settle on the same day as the collection of the Pedi larvae, i.e., at 24 DOC. In this study, we define settlement as the sequential transition of a

competent pediveliger larva to a postlarva, which includes the attachment to a substrate and then eventual metamorphosis.

Postlarvae were confirmed under the microscope as individuals that secreted cement substances or those with postlarval shell growth. Approximately 80% settlement success was observed after induction. The remaining larvae that were not induced to settle were found actively swimming after 24 h. No mortality was observed. All postlarvae that metamorphosed after 24 h were carefully removed from the 6-well plate bases. After removal of the postlarvae from each well, they were pipetted out to a 300 μm mesh-covered tip nestled on a glass bowl filled with FSW, washed gently several times with distilled water to remove any traces of the protein extract, and were transferred to a clean microcentrifuge tube. After the removal of the excess liquid on the tube, the postlarvae were immediately treated with RNAlater. This whole process was repeated several times until all the samples were collected. After 24 h, a similar procedure to that of the Pedi batch was carried out, wherein all the postlarvae were pooled and RNAlater was removed. Approximately 3000 individuals of postlarvae, equivalent to about 100 mg of pooled sample, were stored at −80 ◦C until further use.

Quantitative Real-Time PCR (qRT-PCR) analysis on Gigasin-6 isoform X1 (G6iX1) was determined and compared in both oyster larval stages. PCR primers for Gigasin-6 isoform X1 were designed using the NCBI primer blast tool (https://www.ncbi.nlm.nih.gov/tools/primer-blast/, accessed on 25 October 2021) and the oligoanalyzer tool of Integrated DNA Technologies (IDT) (https://sg.idtdna.com/calc/analyzer, accessed on 25 October 2021). A forward primer (5'– AGGCTCTCGGGGTCAAGGACA – 3') corresponding to the 233-253 bp region of G6iX1 and a reverse primer (5'- GAGTAGAATCTGGGTTGTTTG – 3') corresponding to the 362-381 bp region were purchased from Life Technologies Japan Ltd. The expected length of the target G6iX1 product was 150 bp. The total RNA from both larval stage samples was reverse transcribed following the method of Gao et al. [67]. In brief, all RNA samples were measured for quantity and quality by a NanoDrop ND-1000 spectrometer (Thermo Fisher Scientific, Wilmington, DE, USA). cDNA was synthesized by treating the total RNA sample (500 ng) with oligo-dT primer CDS-BR (5-GGCCACGCGTCGACTAGTAC(T)16–3'), random primer (5-NNNNNNNNN-3'), and M-MLV reverse transcriptase (Promega, Madison, WI, USA). The qRT-PCR assay was conducted using Power SYBR Green PCR Master Mix and a QuantStudio 3 Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific, Rockford, IL, USA) following the manufacturer's instructions. Thermocycling was performed as follows: enzyme activation for 10 min at 95 ◦C, denature step for 15 s at 95 ◦C, and annealing and extension for 1 min at 60

◦C for 40 cycles. Expressions for G6iX1 were normalized to that of the reference gene *elf* (elongation factor 1) for each larval stage sample collection. The reference gene primer used was as reported by Huan et al. [68]. The mRNA expression levels were quantified by the standard curve method, and the relative mRNA expression levels were calculated by the ratio of G6iX1 to the *elf* reference gene. All data were presented as the mean ± standard deviation (SD). This experiment was conducted in triplicate.

III.2.12 Statistical Analyses and Visualization

The settlement percentage (%) was calculated based on the number of settled larvae over the total number of larvae in each well multiplied by 100. These percentages were presented as arithmetic means with standard deviation (SD). Data were analyzed using quasi-binomial generalized linear models (GLM). All larval settlement statistical tests were performed in RStudio (R-project.org, version 4.1.3) [69].

The comparison of the relative Gigasin-6 isoform X1 mRNA expression levels between pediveliger and postlarvae samples were analyzed using a student t-test and was determined using the GraphPad 9 software program (GraphPad Prism, Software Inc., San Diego, CA, USA). All figures in this study have been created using BioRender.com and GraphPad 9 software program.

III.3 Results

III.3.1 Protein separation and identification of the settlement inducing cues from Crassostrea gigas EDTA-soluble shell matrix proteins

III.3.1.1 Protein separation

To locate the signal molecule from the *C. gigas* adult shell EDTA-soluble extract (CgSE), two bioassay-guided protein separation approaches were performed. In this first approach (Appendix text 3A, Appendix Figures A3.1-A3.4), the identification and characterization of the protein extract that could only be visualized by the Stains-all stain method reported by Vasquez et al. [23,24], were further pursued. The chemical solution for extraction was changed from HCl to EDTA since it was considered that the former might have a strong effect on the extracted shell protein components that could make it challenging for identification and downstream analyses.

To determine whether there could be other settlement inducing factors in CgSE, the crude form of CgSE was subjected to another round of protein separation by ultrafiltration but this time, separating it directly into <50K and >50K fractions. The resultant >50K CgSE fraction may include high molecular weight settlement factors that might have been previously missed. To examine whether this >50K CgSE obtained from the second approach could bind to WGA lectin, a WGA-agarose bound lectin affinity chromatography was performed (Figure 3.1 and Appendix Figure A3.5). By using QC-CC and Silver staining methods, 198 (UB1), 66 (UB2), and 17 (UB3) kDa bands were detected in the unbound fraction (UF) component, while the 198 (B1) and 66 (B2) kDa bands were found in the bound fraction (BF) component (Figure 3.1 A, B). On the other hand, under the Stains-all staining method, 48 (S1) and 43 (S2) kDa dominant bands were detected in the UF component (Figure 3.1C).

In Figure 3.1D, no settlement response was observed in multi-wells in the absence of settlement inducing cues, i.e., CgSE, UF, and BF components. Notably, both the UF and BF components showed similar concentration-dependent settlement inducing patterns in all amounts of extract exposure and a 90% settlement response at 50 µg ($p < 0.05$). This result indicates that the signaling molecule has been found in the BF component that may have GlcNAc and Sialic acid-containing sugar moieties which could bind with WGA lectin and have been hypothesized to mediate settlement in *C. gigas* [26]. On the other hand, although the

protein components in the unbound fraction may not have bound to the WGA lectin, it indicates that it could still induce settlement. No mortality was recorded in all bioassay experiments.

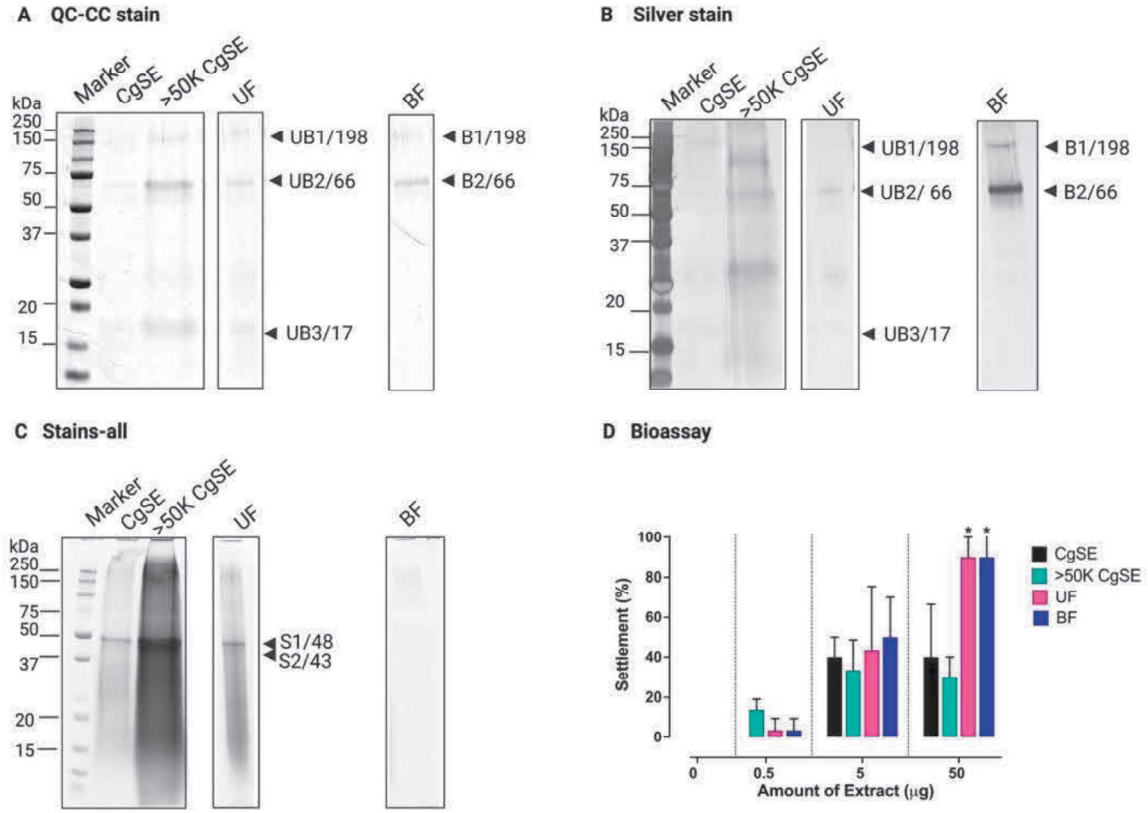

Figure 3.1. SDS-PAGE analyses and larval settlement responses to *Crassostrea gigas* EDTA-soluble shell matrix proteins and its fractions after elution from WGA lectin affinity chromatography. (**A-C**) SDS-PAGE profile of fractions eluted from WGA lectin affinity chromatography under different staining methods. (**D**) Settlement percentages of *C. gigas* larvae response with varying amounts of different forms of CgSE and fractions eluted after WGA affinity chromatography. All larvae were exposed to the settlement inducing cue for 24 h. Asterisks " * " indicate significant differences among varied forms of CgSE determined via quasi-binomial glm ($p < 0.05$). Missing bars in the figure indicate no settlement. Data are the means (SD) of three replicates. Abbreviations: UF, Unbound fraction; BF, Bound fraction; UB1-UB3, B1, B2, S1, and S2 bands correspond to the identified proteins in the mass spectrometry analyses and are summarized in Table 3.1. This figure was created using BioRender.com and GraphPad 9 software program.

III.3.1.2 Identification by Mass Spectrometry

To identify the dominant matrix proteins of CgSE, both the UF and BF components were subjected to Cys-β-propionamidation prior to SDS-PAGE and further analyzed by peptide mass fingerprinting (PMF) following in-gel tryptic digestion. As summarized in Table 3.1 and shown in Figure 3.1, three protein groups were identified. Annotated MS spectra are provided in Appendix Table A3.1. Availability of the gene sequences was evaluated with *C. gigas* CDS

and NCBInr databases. The disparity of molecular masses between the observed and theoretical mass may be due to the additional presence of post-translational modifications [44].

Table 3.1. Identification of the EDTA-soluble shell matrix proteins of *Crassostrea gigas* by Peptide Mass Fingerprinting Analysis in the unbound and bound fractions eluted after WGA affinity chromatography.

Band No.	Protein name [a]	Accession Number [a]	Theoretical mass (kDa)	Observed mass (kDa)	PS/ NMP [b]	SC [c] (%)
		Unbound fraction components				
UB1	Gigasin-6 isoform X2	XP_011449648.2	63	198	79/15	20
	Gigasin-6 isoform X1	XP_011449647.2	64	198	78/15	19
UB2	Gigasin-6 isoform X2	XP_011449648.2	63	66	85/15	21
	Gigasin-6 isoform X1	XP_011449647.2	64	66	84/15	20
UB3	Surface protein P12p-like	XP_034321529.1	15	17	80/8	37
	Surface protein P12p-like	XP_034319527.1	15	17	80/8	37
S1	Unidentified Stains-all stainable acidic protein	N/A	N/A	48	N/A	N/A
S2	Unidentified stains-all stainable acidic protein	N/a	N/A	43	N/A	N/A
		Bound fraction components				
B1	Gigasin-6 isoform X1	XP_011449647.2	64	198	97/17	21
	Gigasin-6 isoform X2	XP_011449648.2	63	198	96/17	22
B2	Gigasin-6 isoform X1	XP_011449647.2	64	66	63/13	18
	Gigasin-6 isoform X2	XP_011449648.2	63	66	63/13	18

a: Protein name and accession number according to NCBI database, b: Protein score (PS) and the number of matched peptides (NMP) obtained from Mascot search, c: Percentage of sequence coverage of identified peptides related to the corresponding sequence in the database.

The major band UB2 (66 kDa) in the unbound fraction (UF) and band B2 in the WGA column-bound fraction (BF) were identified as Gigasin-6 isoform X1 and/or X2. The minor

band UB1 (198 kDa) in the unbound fraction (UF) and band B1 in the WGA column-bound fraction (BF) band were also identified as the same Gigasin-6 isoform(s). This 198 kDa band could be an SDS-resistant trimer of Gigasin-6 (66 kDa) as reported in certain other proteins [70–72]. The 17 kDa band (UB3) was identified as (putative) mature proteins from two gene product(s) of Surface protein P12p-like that were indistinguishable from each other. Furthermore, both QC-CC and silver stain profiles of Gigasin-6 isoform X1 and/or X2 and its trimer showed that a more intense band pattern of this protein was visualized in the bound fraction component than in the unbound fraction, indicating that majority of this protein bound to the WGA lectin. The low amount of Gigasin-6 isoform X1 and/or X2 and its trimer that remained in the unbound fraction may be due to the lack of available binding sites on the immobilized WGA lectin resin or as a result of the oversaturated presence of the same protein. Those that could not bind to the lectin might have flowed through together with other protein components in the unbound fraction. On the other hand, 43 and 48 kDa bands which could only be visualized with Stains-all stain were not successfully identified due to a lack of tryptic cleavage sites and also due to stain interference [44].

III.3.2. Biochemical characterization of the settlement inducing cues from C. gigas EDTA - soluble shell matrix proteins

III.3.2.1 Posttranslational modifications (PTMs)

To examine the presence of PTMs in the WGA affinity chromatography eluted UF and BF components, glycoprotein and phosphoprotein detection were performed as shown in Figures 3.2 and 3.3, respectively. Bands that showed fluorescence staining were confirmed according to each staining method. Mass spectrometry analyses of identified bands are shown in Tables 3.1 and 3.2.

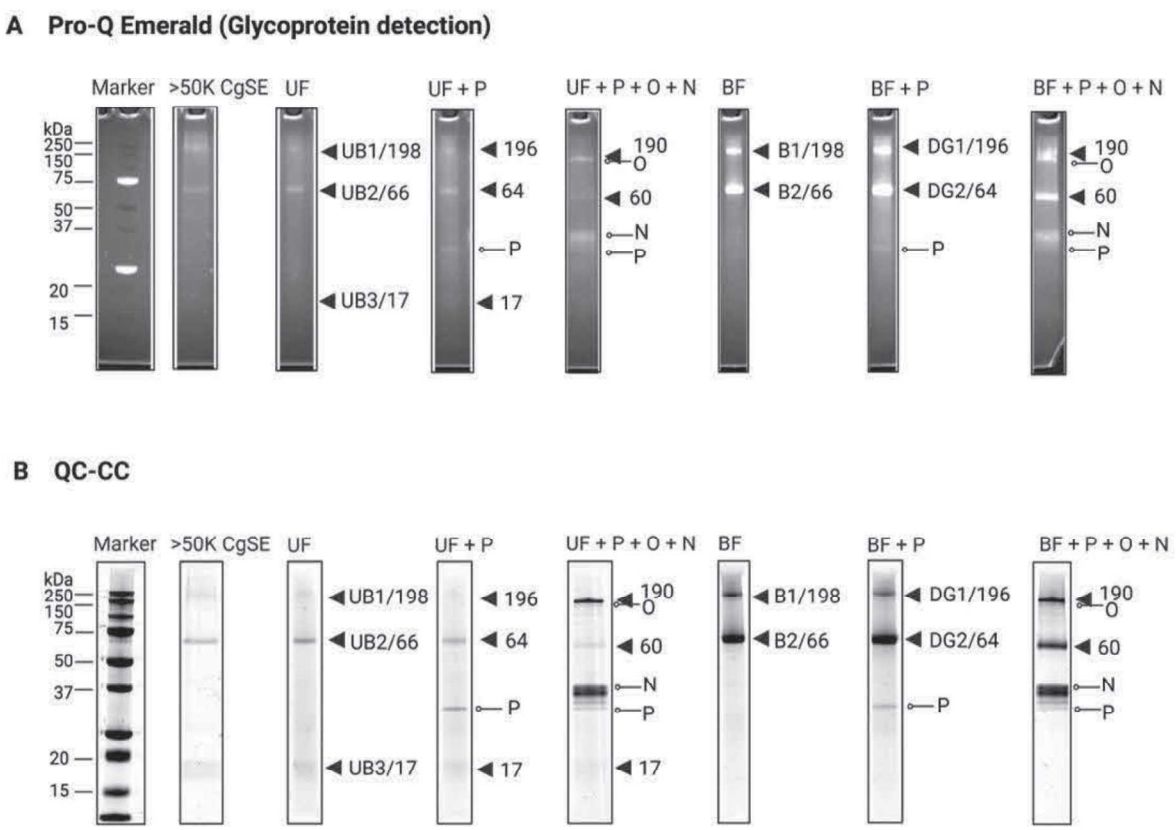

Figure 3.2. Enzymatic deglycosylation and glycoprotein staining profiles. (**A**) Unbound and Bound fractions eluted from WGA affinity chromatography were enzymatically deglycosylated with PNGase F (P), *O*-glycosidase (O), and Neuraminidase (N). Ten µg of each sample was loaded on a 12% SDS-PAGE and the gel was stained by Pro-Q emerald for Glycoprotein detection. (**B**) The gel was subsequently double-stained with QC-Colloidal Coomassie (QC-CC) Stain. Abbreviations: UF, Unbound Fraction; BF, Bound Fraction; UB1-UB3, B1, B2, DG1, DG2, S1, and S2 bands correspond to the identified proteins in the mass spectrometry analyses and are summarized in Tables 3.1 and 3.2. This figure was created using BioRender.com (accessed on 25 July 2022).

Enzymatic deglycosylation and Pro-Q Emerald staining revealed that Gigasin-6 isoform X1 and/or X2 (66 kDa) and its trimer (198 kDa) is a highly glycosylated protein (Figure 3.2). Gigasin-6 isoform X1 and/or X2 (66 kDa) and its trimer (198 kDa) found in the Unbound fraction (UF) component showed a weak band fluorescence compared to its Bound fraction (BF) component. After PNGase F treatment alone, the Gigasin-6 isoform X1 and/or X2 trimer (198 kDa) showed a slight band shift from 198 (UB1/B1) to 196 (DG1) kDa while, in the native form of this protein (66 kDa), a slight band shift from 66 (UB2/B2) to 64 kDa (DG2) was also observed on both the UF and BF components. Simultaneous removal of *N*-glycans and *O*-glycans on the polypeptide backbone [73] in both the UF and BF components showed a significant band shift on Gigasin-6 isoform X1 and/or X2 and its trimer. As a result, the 198 kDa band trimer shifted to a 190 kDa band, while the native form of Gigasin-6 kDa (66 kDa)

shifted to 60 kDa. The fluorescence intensity pattern within each of the UF and BF components of the deglycosylated Gigasin-6 isoform X1 and/or X2 and its trimer became weak compared to their glycosylated forms. These results demonstrate that Gigasin-6 isoform X1 and/or X2 and its trimer contain several types of *N*-glycans containing GlcNAc and Sialic acid residues (for PNGase F treatment) [23,74]; *O*-glycans containing non-reducing terminals of *N*-acetylgalactosamine Core 1 and Core 3 *O*-linked disaccharides (for *O*-glycosidase treatment) [73], α2-3, α2-6, and α2-8 linked *N*-acetyl-neuraminic acid residues (for Neuraminidase treatment) [73]. Notably, both the 190 kDa and 60 kDa deglycosylated bands of Gigasin-6 isoform X1 and/or X2 continued to show a positive fluorescence stain albeit weaker after the sequential treatment of a mixture of deglycosylation enzymes. This suggests that Gigasin-6 isoform X1 and/or X2 may possess other types of uncleaved *O*-linked glycans. A similar finding was also reported by Vasquez et al. [23] regarding the presence of GlcNAc residues on the settlement inducing cue. On the other hand, Surface Protein P12p-like was shown to have a faint fluorescence band pattern in the UF component (Figure 3.2 A) indicating that it is lightly glycosylated. Whereas 43 (S2) and 48 (S1) kDa Stains-all stainable acidic protein bands showed no Pro-Q Emerald fluorescence. This may indicate that the stains-all stainable acidic protein bands may have no or few glycosylation sites. Also, no band shift was observed on these respective protein bands after enzymatic deglycosylation treatments.

To confirm the identity of Gigasin-6 isoform X1 and/or X2 (66 kDa) and its trimer (198 kDa in the WGA column-bound fraction (DG 1 and 2 bands) after treatment with PNGase F, mass spectrometry analysis was performed. The WGA column-bound fraction (BF) proteins were subjected to Cys-β-propionamidation prior to SDS-PAGE and further analyzed by MS/MS ions search following in-gel tryptic digestion. Annotated MS/MS spectra are provided in Appendix Table A3.2. These same MS spectra were used to further determine *N*-glycosylation sites of Gigasin-6 isoform X1 and/or X2 in the WGA column-bound fraction (BF) after treatment with PNGase F.

Table 3.2. MS/MS-based identification and characterization of the protein band containing the determined *N*-glycosylation sites of Gigasin-6 isoform X1 and/or X2 in the WGA column-bound fraction (BF) after treatment with PNGase F.

Band No.	Protein name [a]	Accession Number [a]	Theoretical mass (kDa)	Observed mass (kDa)	PS/ NMP [b]	SC [c] (%)
Bound fraction components						
		Deglycosylated				
DG1	Gigasin-6 isoform X1	XP_011449647.2	64	196	73/14	21
	Gigasin-6 isoform X2	XP_011449648.2	63	196	72/14	22
DG2	Gigasin-6 isoform X1	XP_011449647.2	64	64	91/16	21
	Gigasin-6 isoform X2	XP_011449648.2	63	64	90/16	22

a: Protein name and accession number according to NCBI database, b: Protein score (PS) and the number of matched peptides (NMP) obtained from Mascot search, c: Percentage of sequence coverage of identified peptides related to the corresponding sequence in the database.

For the phosphoprotein detection, a separate set of samples from UF and BF components together with other forms of CgSE were subjected to SDS-PAGE and visualized by the Pro-Q Diamond staining method as shown in Figure 3.3. Results revealed high fluorescence intensity on the 43 (S2) and 48 (S1), and 50 kDa Stains-all stainable acidic bands as well as Gigasin-6 isoform X1 and/or X2 (66 kDa) and its trimer (198 kDa). This indicates that these bands are highly phosphorylated. However, the 38 kDa stains-all stainable acidic protein band and Surface protein P12p-like (UB3/17 kDa) showed fluorescence but with a weaker signal compared to other stained proteins and may contain few phosphorylated sites. This finding suggests that all these dominant proteins in CgSE are phosphorylated.

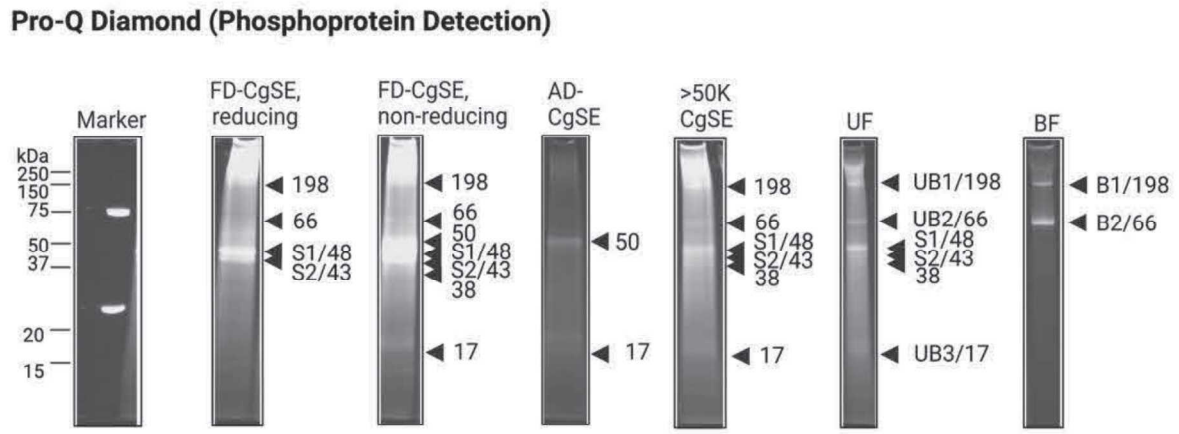

Figure 3.3. Phosphoprotein detection. Ten μg of the sample was loaded on each lane and stained with Pro-Q Diamond for phosphoproteins. Abbreviations: FD, Freeze-dried CgSE, AD, Air-dried CgSE; UF, Unbound Fraction; BF, Bound Fraction; UB1-UB3, B1, B2, S1, and S2 bands correspond to the identified proteins in the mass spectrometry analyses and are summarized in Table 3.1. This figure was created using BioRender.com (accessed on 25 July 2022).

III.3.2.2 Characterization of Stains-all stainable acidic proteins by protease and chemical deglycosylation treatments

To further determine the characteristics of the protein bands that were consistently visualized by the Stains-all stain method, CgSE was treated separately with Trypsin, Chymotrypsin, Asp-N, and Trifluoromethanesulfonic acid (TFMS) for chemical deglycosylation. Figure 3.4 A shows that without any enzyme treatment, CgSE contains 38, 43, and 48 kDa, bands. After enzyme treatment, a 10 kDa band shift was observed on CgSE treated with trypsin and chymotrypsin. Notably, the band shift in both treatments showed similar patterns wherein 24, 30, 36, and 38 kDa newly appeared. However, these target bands were not completely cleaved and showed some resistance to these enzymes. The cleavage specificity of trypsin is on the carboxyl side of arginine and lysine residues while chymotrypsin acts on the carboxyl side of tyrosine, phenylalanine, tryptophan, and leucine residues. Interestingly, after Asp-N treatment, all bands disappeared owing to its fragmentation into very low molecular weight size undetectable under SDS-PAGE. Asp-N specifically cleaves the amino side of aspartic acid residues. This indicates that CgSE, which is predominantly composed of stains-all stainable acidic proteins, is rich in Aspartic acid. This is typical for proteins found in shells [40,51] and which may exhibit properties common to intrinsically disordered proteins [51]. Moreover, TFMS-treated CgSE showed no distinguishable bands, instead, a diffused and scaling smeared pattern from 20 kDa to >250 kDa appeared. Chemical deglycosylation with TFMS allows the removal of monosaccharides without damaging the

peptide [75]. However, in this experiment, results indicate that peptide bond hydrolysis may have occurred thereby, creating such a band pattern.

To examine the presence of glutamic acid residues in the stains-all stainable acidic proteins, >50K CgSE was treated with an endoproteinase, Glu-C, at different enzyme to protein ratios. Figure 3.4 B consistently showed an uncleaved prominent 48 kDa band after varied Glu-C concentrations. This indicates that this matrix protein is Glu-C resistant and may be due in part to the presence of Glu-C proteinase resistant Glu-X bonds, such as Glu-Asp, Glu-Asn, and Glu-Gly, where X is, respectively, linked to Pro or Gln residues which could yield a missed cleavage [76].

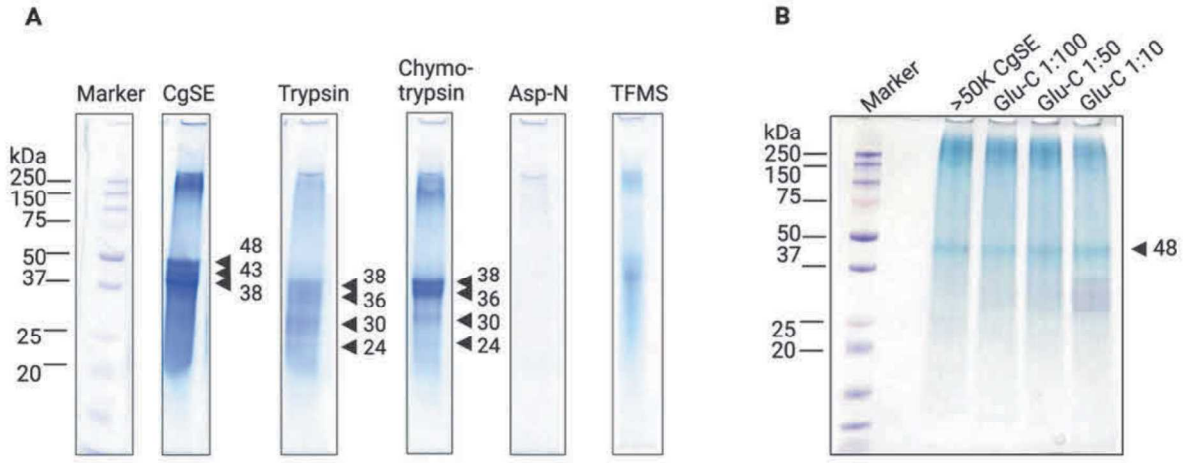

Figure 3.4. SDS-PAGE profiles of CgSE and >50K CgSE-protease and chemically deglycosylated treated samples. (**A**) Forty µg of CgSE was treated with various endoproteinase enzymes and stained with Stains-all. (**B**) Ten µg of >50K CgSE was treated with Glu-C Endoproteinase under different enzyme-to-protein ratios and stained with Stains-all. Abbreviation: CgSE, *Crassostrea gigas* shell EDTA extract; TFMS, Trifluoromethanesulfonic acid. This figure was created using BioRender.com (accessed on 25 July 2022).

In addition, in the first protein separation approach, the stains-all stainable acidic proteins showed that under reducing conditions, no settlement inducing activity was detected from the isolated polypeptides in active fraction F2, i.e., with the addition of 2-mercaptoethanol (Appendix text A.3, Figure A3.4). Conversely, under non-reducing conditions, significant settlement inducing activities were observed at the 35, 38, and 48 kDa polypeptide bands (Appendix text A.3, Figure A3.4). This suggests that disulfide bonds might be necessary for the folding and stabilization of the settlement inducing factors in this cue [44].

The overall SDS-PAGE profiles in Figure 3.4 suggest that aspartic acid-rich proteins are the major components of the EDTA-soluble organic matrix in this study as is reported in other bivalves [40,77]. This implies that these proteins may also bind to calcium [77] and that they

are intrinsically disordered proteins. Moreover, as the amount of extract loaded on each lane increased, the intensity of the smeared background along each lane and the appearance of aggregated proteins close to the loading lane was noticeable. This also confirms the presence of a high level of post-translational modifications found in this extract component [39,50].

III.3.3 Molecular characterization of the settlement inducing cues from C. gigas EDTA -soluble shell matrix proteins

To examine further the characteristics of the dominant protein bands in the UF and BF components following WGA affinity chromatography, bioinformatic characterization of all protein sequences, mass spectrometry, and quantitative real time-PCR (qRT-PCR) (for Gigasin-6 isoform X1 only) analyses were performed.

III.3.3.1 Gigasin-6 isoform X1 and/or X2

Peptide Mass Fingerprinting analysis revealed that the genes encoding both 66 kDa and 198 kDa bands were identified as Gigasin-6 isoform X1 and/or X2 (Table 3.1). Mass spectrometry profiles of this gene showed identical mass to charge (m/z) ratio signals that made it difficult to differentiate both isoforms (Appendix Table A3.1). Multiple sequence alignment of the precursor Gigasin-6 and its two isoforms confirmed that these isoforms contain identical sets of amino acid sequences except that Gigasin-6 isoform X1 possesses an additional 9 residues (DSNNPDSTL) at positions 115 to 123 (Appendix Figure A3.6).

A graphical representation of the predicted and experimentally observed properties of Gigasin-6 isoform X1 is shown in Figure 3.5 A. Bioinformatic analyses on Gigasin-6 isoform X1 protein sequence suggest 5 *N*-glycan, 12 *O*-glycan, and 70 putative phosphorylation sites; 1 putative RGD cell attachment sequence at positions 177-179; and 10 putative *N*-myristoylation sites (Figure 3.5 A, Appendix text B3.1, Appendix Figure A3.6). If the *N*-myristoylation sites are functional, each site may determine the final location of this protein on the shell layer [77].

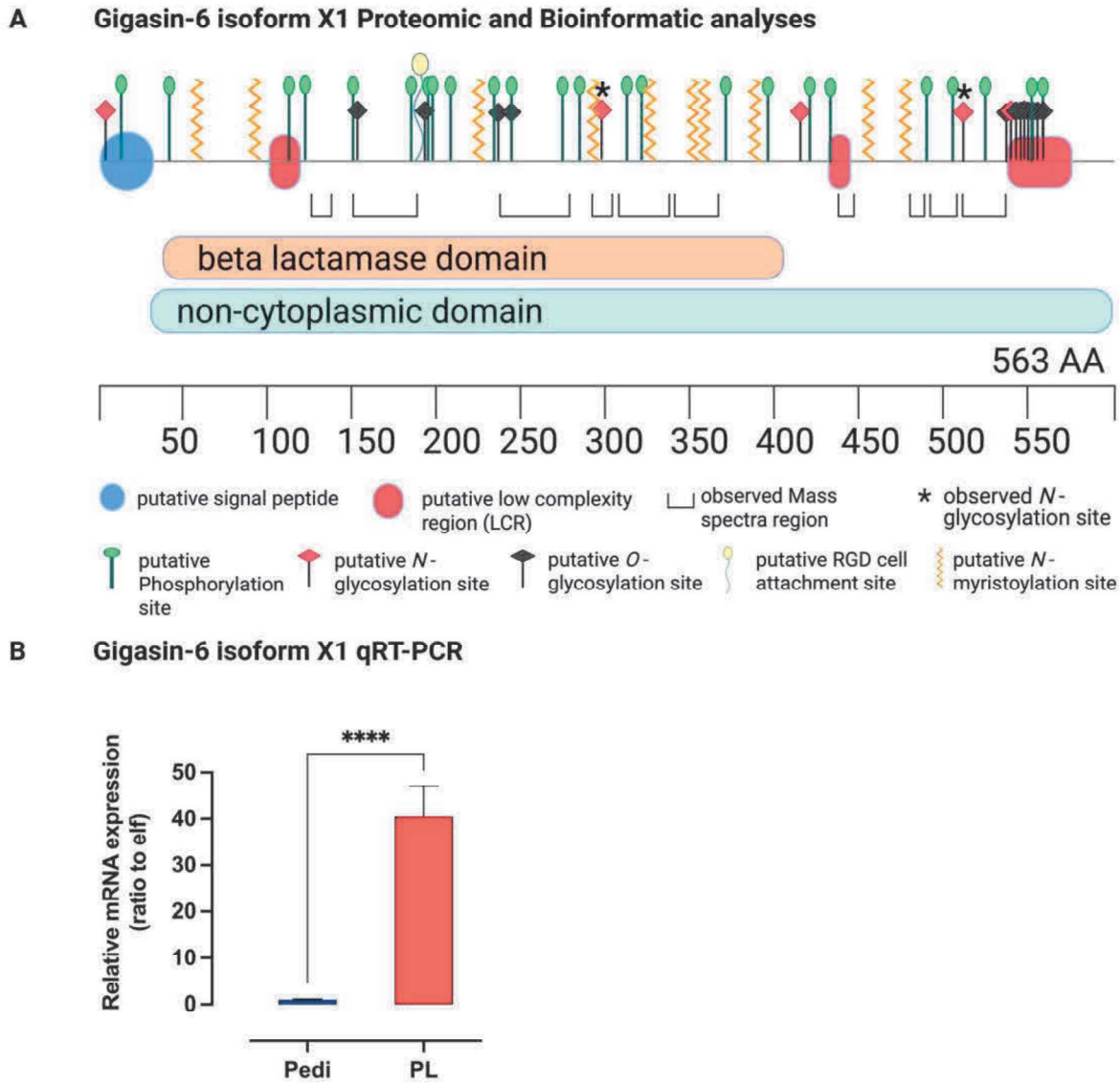

Figure 3.5. Molecular characterization of Gigasin-6 isoform X1. (**A**) Graphical representation showing its protein length, signal peptide, putative domains, motif, and posttranslational modifications identified by mass spectrometry and bioinformatic analyses. The putative phosphorylation sites (24 out of 70) in this model highlight those which may be involved in signal transduction processes. The observed region indicates the location of the mass spectra identified by mass spectrometry analysis. Figure 3.5A was created using BioRender.com (accessed on 25 July 2022). (**B**) Relative mRNA expression levels of Gigasin-6 isoform X1 gene before (Pediveliger, Pedi) and after (postlarvae, PL) settlement induction by the CgSE using quantitative real-time PCR. Data are expressed as the mean (SD) (n = 3). Values with "****" are significantly different ($p < 0.0001$). Figure 3.5B was created using GraphPad 9 software program.

To identify the actual *N*-glycosylation sites in this protein, peptide mass fingerprint patterns of the glycosylated and deglycosylated bands B5 and B7, respectively, were compared (Tables 3.1 and 3.2, Figure 3.2, Appendix text B3.2, Figure A3.7) [54]. Results revealed 4 *N*-glycosylation site-related m/z signals corresponding to two *N*-glycosylation peptide sequences (^{521}**NST**YIEAFTVDKFDDAKFER^{540} and ^{291}FMNYLLG**NGS**IPGTNDVLLAK^{309}). This

study also reports for the first time an unrecorded actual *N*-glycosylation site-related m/z 2525.21 signal from this protein. This result highlights the importance of combining biochemical, bioinformatic, and mass spectrometry techniques to explore and verify undiscovered glycan moieties that may play crucial roles in biological processes [54].

As shown in Figure 3.5 B, results of the qRT-PCR analysis showed that Gigasin-6 isoform X1 was highly expressed in the CgSE-induced postlarvae than in the pediveliger larvae ($p < 0.001$). This finding further supports the transcriptome data in a previous study where Gigasin-6 isoform X1 was differentially expressed and highly upregulated in the CgSE-induced postlarvae [47]. Also, this result is consistent with the findings of Xu and Zhang [6] where the Gigasin-6 gene was significantly upregulated in the epinephrine-induced postlarvae of *C. gigas*. It is often observed that in a large-scale culture of oyster larvae, once a batch of postlarvae begins to settle on the culture tank, other competent pediveliger larvae are also induced to settle. In this present study, the isolated Gigasin-6 isoform X1 and/or X2 found in the adult shells that could induce larval settlement was also highly expressed in the CgSE-induced postlarvae. This indicates that it may function in postlarva-larva settlement interactions such that in the presence of settled juveniles, other non-settled larvae may also be induced to settle.

Multiple sequence alignment of Gigasin-6 isoform X1 showed some amino acid residues, sequence insertions, and deletions that were unique to itself in comparison with other bivalves as shown in Figure 3.6. The identification of Gigasin-6 isoform X1 and/or X2 in the adult shells of *C. gigas* is consistent with a previous study by Mouchi et al. [78]. They have also reported this protein could interact synergistically with other acetic acid-soluble matrix proteins to regulate *in vitro* the formation of calcium carbonate through induction of polycrystalline aggregates [78].

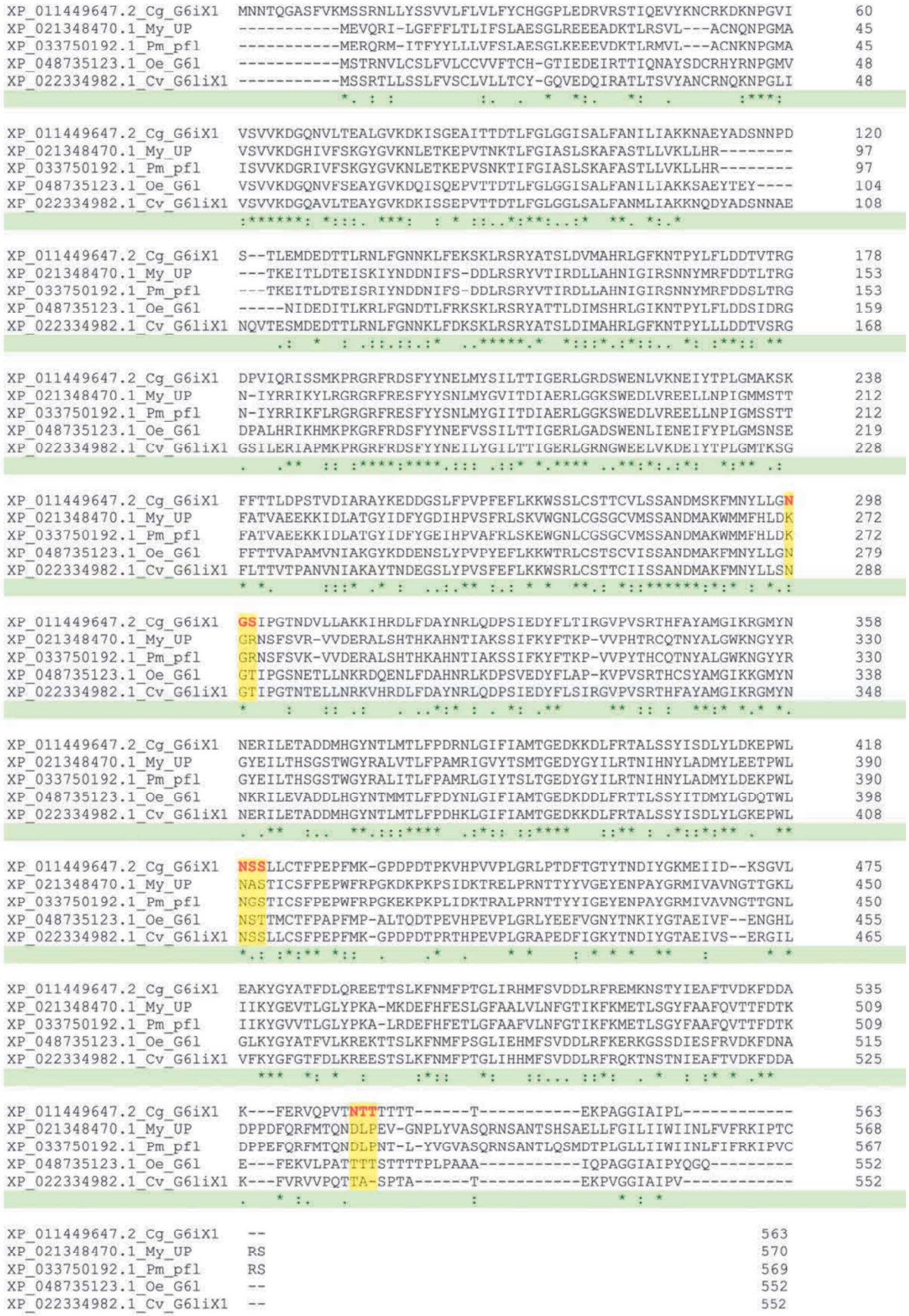

Figure 3.6. Multiple sequence alignment of Gigasin-6 isoform X1 with homologous sequences from other bivalves. Dashes denote blanks or gaps. Asterisks, colons, and dots represent

identical amino acids, conserved substitutions, and semi-conserved substitutions, respectively. Putative *N*-glycosylation sites of G6ix1 are in red font while those from other organisms are highlighted in yellow. Abbreviations: G6iX1, Gigasin-6 isoform X1 from *Crassostrea gigas*; My_UP, Uncharacterized protein LOC021348470.1 from *Mizuhopecten yessoensis*; Pm_pfl, Protein flp-like from *Pecten maximus*; Oe_G6l; Gigasin-6-like from *Ostrea edulis;* Cv_G6liX1, Gigasin-6-like isoform X1 from *Crassostrea virginica.*

III.3.3.2 Stains-all stainable acidic proteins

The close association of the stains-all stainable acidic proteins with the biomineral phase and their highly acidic nature made it difficult to confirm their identity. The in-gel and in-solution protease digests (Figures 3.1 and 3.4) of all the stains-all stainable acidic protein components did not reveal any peptide with sequences corresponding to those found in the *C. gigas* NCBI database. Asp-N digests were too fragmented that did not yield any identifiable sequences. Although mass spectrometry analysis could not successfully identify the putative settlement inducer from the stains-all stainable acidic proteins due to their large-sized, trypsin, chymotrypsin, and Glu-C digests, all bands showed similar chemical signatures, suggesting that these bands are fragments of proteins that may have similar chemical structures (data not shown). Since a stains-all stainable shell matrix protein has been identified from other species; i.e., Folian cv1 (*C. virginica*) [40], the Folian cv1 homolog of *C. gigas* was searched by BLAST and a gene encoding dentin sialophosphoprotein-like protein (CGDSP, XP_034310169.1, Appendix text C3, Figure A3.8) was identified. Prediction of intrinsically unstructured proteins and detection of low complexity regions (LCRs) were done using IUPred3 software. IUPred3 analysis of Dentin sialophosphoprotein-like protein sequence revealed that it is considered a putative intrinsically disordered protein (Appendix text C3). In my previous study, the dentin sialophosphoprotein (CGDSP) gene was found to be highly upregulated in the CgSE-induced postlarvae transcriptome [47]. Nonetheless, in this study, the ability of these isolated stains-all stainable acidic proteins to induce larval settlement in both the first and second purification approaches was demonstrated to be consistent with the report of Vasquez et al. [23] where a 55 kDa stains-all stainable acidic protein from an HCl-soluble shell extract was found to induce conspecific settlement in *C. gigas* larvae.

III.3.3.3 Surface protein P12p-like

Peptide Mass Fingerprinting of the 17 kDa band was identified as the product(s) of two isoform genes of Surface protein P12p-like (accession numbers: XP_034319257.1 and XP_034321529.1) as shown in Table 3.1. In Appendix Figure A3.9A, a sequence alignment

comparison of these two variant gene products indicates they have identical amino acid sequences except for one substitution of an amino acid residue at position 7, where a Phe is replaced by Leu, in a putative signal peptide. The protein sequence of this protein was predicted only by automated computational analysis derived from a genomic sequence in *C. gigas* [79]. Its functional role remains unexplored. This is the first report on the mass spectrometry and PTM characterization of this protein. Results in this study indicate that it is slightly glycosylated and phosphorylated. A closer look at its few putative *N*-and *O*-glycosylation and phosphorylation sites (Appendix Figure A3.9) concurs with the observed results in the SDS-PAGE analyses (Figures 3.2, 3.3). It also contains some putative *N*-myristoylation sites (Appendix text D3, Figure A3.9). To compare the *C. gigas* Surface protein P12p-like (CGS12P) with its homologs, a multiple sequence alignment was performed. Notably, Appendix Figure A3.10 shows this protein has variation and low homology with other proteins specific to *Crassostrea gigas,* but no homology was found in other organisms. Notably, this is the first report implicating the actual presence of this protein as part of the adult shell matrix in *C. gigas*.

The overall biochemical and molecular characterization of Gigasin-6 isoform X1 and/or X2, CGS12P, CGDSP, and other stains-all stainable acidic bands in this study suggest that their amino acid composition showed some amino acid residues, sequence insertions, and deletions that are unique to each of these proteins in *C. gigas* in comparison with other organisms; and that all these isolated proteins possess post-translational modifications (glycosylation and phosphorylation) that may play crucial roles in cell signaling processes and strengthen recognition processes for conspecific cue-mediated larval settlement induction.

III.4 Discussion

This study reports a novel role of several shell matrix proteins that cooperatively function as the *Crassostrea gigas* Settlement Pheromone Protein Components in adult oyster shells and is demonstrated as the biological cue responsible for gregarious settlement on conspecifics.

In this present study, a bioassay-guided protein separation approach was done to locate the key molecule from conspecific adult shells responsible for gregarious settlement in *C. gigas*. Results showed that in the first protein separation approach, a group of stains-all stainable acidic proteins was isolated. The extracted proteins were capable of inducing settlement and possess characteristics inherent to proteins with intrinsically disordered regions (IDPs) [42,51,80,81]. This result is consistent with the report of Vasquez et al. [23] where a 55 kDa stains-all stainable acidic protein from an HCl-soluble shell extract was found to induce settlement in *C. gigas* larvae. In the second protein separation approach, although Gigasin-6 isoform X1 and/or X2 which is bound to a WGA lectin resin could induce a high concentration-dependent settlement inducing activity, the unbound fraction could also elicit the same settlement-inducing pattern. This unbound fraction component contained a combination of a low amount of Gigasin-6 isoform X1 and/or X2, a moderate amount of CGS12P, and the dominant stains-all stainable acidic proteins including the 48 kDa band, putatively identified as CGDSP. One possible reason for such a settlement inducing pattern in the unbound fraction component could be that only a small amount of Gigasin-6 isoform X1 and/or X2 might be needed to contribute a settlement-inducing effect. However, the stains-all stainable acidic proteins which were shown in the first protein separation approach to induce settlement could also be contributing to the increased settlement signal in this fraction component. Moreover, whether CGS12P is a co-factor or not has to be investigated and confirmed by further study. These protein components in the unbound fraction that may not have bound to the WGA lectin indicate that it could still induce settlement by interacting with the oyster larvae via an unknown signaling pathway. Additionally, separating the dominant protein components individually showed lower settlement inducing effects. The overall results of the bioassay-guided protein separation, biochemical, and molecular analyses suggest that these dominant EDTA-soluble shell matrix proteins may be synergistically interacting together to form a macromolecular assembly capable of creating a stable and strong aggregating signal for conspecific settlement which is hereon named the *Crassostrea gigas* Settlement Pheromone Protein Components (CGSPPC). However, whether or not CGSPPC forms a `complex' or is dispersed on the shell has to be studied in the future.

One possible reason that may explain the synergistic settlement inducing effect of the CGSPPC could be the hypothesis proposed by Pancsa et al. [80]. Pancsa et al. [80] proposed that intrinsically disordered region-containing proteins do not have a function on their own, but when these proteins form dynamic and non-stoichiometric supramolecular assemblages, novel and functional properties emerge [80]. The possible presence of the putative CGDSP and other IDPs acting as scaffold proteins may provide the structural framework on which other settlement-inducing proteins could anchor via phosphorylation interactions and induced proximity [82]. IDPs which are involved in biomineralization are frequently very extended and disordered [81]. However, their disordered structure is integral to how these proteins fulfill their function [81]. IDPs are also known as assemblers which could simultaneously bind to multiple ligands as macromolecular assemblies [80,81]. The meta-stable conformation of the stains-all stainable acidic proteins in this study, as an IDP, allows them to bind to their protein partners as well as act with high specificity and relatively low affinity [81]. The capacity of the stains-all stainable acidic proteins to interact with Gigasin-6 isoform X1 and/or X2 and CGS12P may have contributed to changing or adjusting the structures and functions of all multiple partners enabling them to collectively intensify the settlement inducing signal [83]. This might be the underlying reason that explains the comparable high settlement inducing activity of the unbound fraction component despite having most of the Gigasin-6 isoform X1 and/or X2 separated into the bound fraction component. Further investigation is needed to clarify the specific binding sites of these protein components within the *C. gigas* adult shell. Notably, the stains-all stainable acidic proteins were found to be trypsin, chymotrypsin, and Glu-C resistant indicating that this set of proteins is highly modified and has no or few cleavage sites for these enzymes. In addition, Vasquez et al. (2013) have reported that the stains-all stainable acidic protein 55 kDa putative settlement cue from *C. gigas* HCl-soluble shell matrix extract was stable even at 200°C [23]. This stability of some shell matrix proteins at high temperatures has also been demonstrated in other oyster shells [84]. This proves advantageous for oysters that have to protect themselves against enzyme secretions from shell boring epibiont organisms, extreme desiccation exposure, and possible predation [42]. Furthermore, these properties of the stains-all stainable acidic proteins suggest the protective role it may confer on its interacting partners within the shell matrix. Thus, enabling this conspecific cue to elicit a strong and stable aggregating signal even under harsh environmental conditions. This is a point that may have contributed to their evolutionary success as benthic organisms.

A graphical representation of CGSPPC and its comparison with barnacle SIPC is shown in Figure 3.7. The present findings showed that several signaling molecules (CGSPPC) from *C. gigas* conspecific adult shells could induce settlement while those reported in other marine invertebrates come from a single signaling molecule such as the α-macroglobulin-like gene that encodes the Settlement Inducing Protein Complex (SIPC) in barnacles [45,46]. In Sedanza et al. [26], results indicate that induction of settlement in the oyster larvae may not only require lectin-glycan interactions [26] but also amino acids and other chemical groups [85,86] including phosphates present at the site of chemical contact between the larvae and the substrate [26]. The signaling molecules, i.e., Gigasin-6 isoform X1 and/or X2, CGDSP, and CGS12P, identified in this study showed amino acid residues, sequence insertions, and deletions that were unique in *C. gigas* in comparison with other organisms (Figure 3.6, Appendix Figures A3.8 and A3.10). This may contribute to the specificity of the conspecific cue recognition processes in *C. gigas* larval settlement. The ability of CGSPPC to induce settlement, either individually or collectively, highlights the important role of PTMs in oyster chemical signaling for conspecific recognition. The major settlement inducing component in CGSPPC which is Gigasin-6 isoform X1 and/or X2 is characterized as a phosphorylated sialoglycoprotein with varied *N*- and *O*-glycan specificities. It showed positive binding to WGA lectin which was previously reported to mediate settlement in *C. gigas* [23,24,26]. It may also function in settled juveniles as an attractant for other non-settled larvae to commence settlement. All these properties of Gigasin-6 isoform X1 and/or X2 provide support to previous findings [26] that multiple glycans may be necessary to produce high avidity binding interactions during larval settlement induction [86] and that among these glycans, GlcNAc and Sialic residues may occupy the highest ratio of these sugar moieties as evidenced by an intense band pattern of Gigasin-6 isoform X1 and/or X2 that has bound to WGA lectin resin. The differential effects of sugar types on attachment to a substrate until settlement suggests that such ratios within the settlement inducing cue could determine the maximum adhesion of a settling larva on a substrate [87]. Moreover, CGDSP together with other isolated IDPs were also demonstrated in this study as co-factors that could contribute to strengthening the inducing signal via phosphorylation crosstalk with an unidentified CGSPPC recognizing larval receptor. This suggests that the putative larval receptor may also be a phosphoprotein.

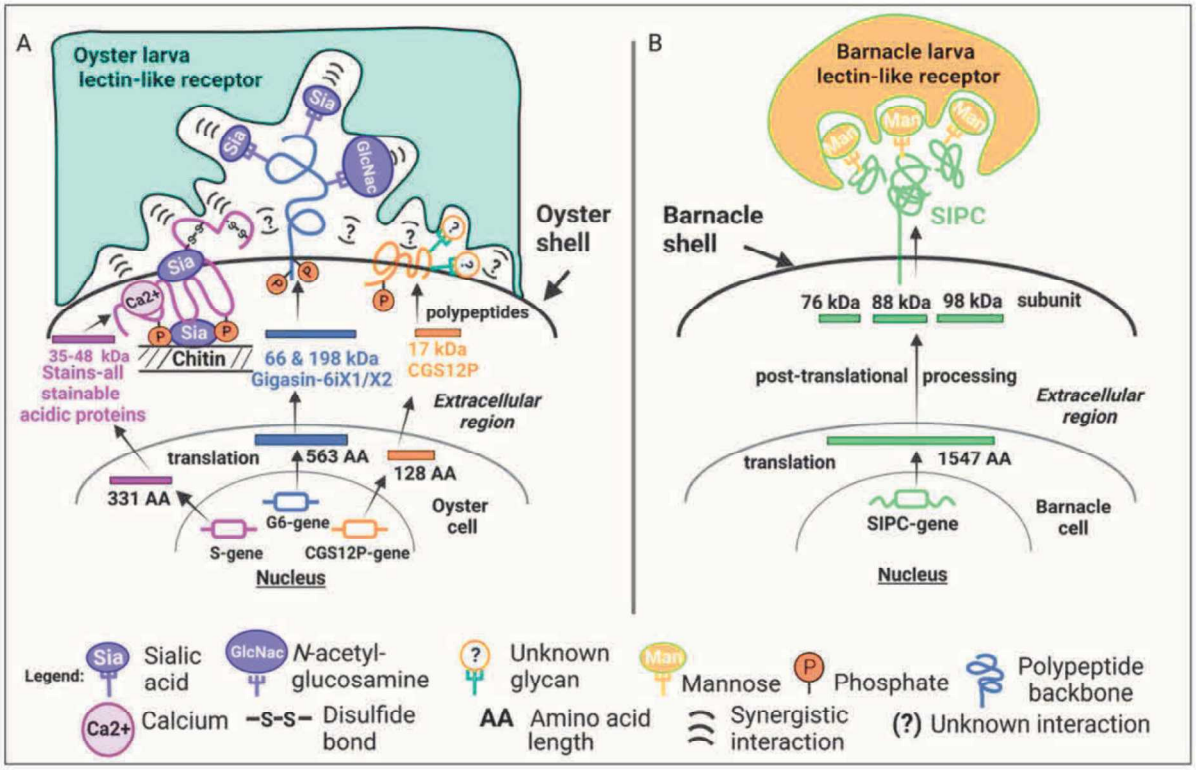

Figure 3.7. Comparison of settlement inducing cues in barnacles and oysters. (**A**) Supramolecular interaction of several unrelated genes from adult oyster shells as *Crassostrea gigas* Settlement Pheromone Protein Components (CGSPPC). (**B**) Settlement Inducing Protein Complex (SIPC) is encoded in a single gene in the barnacle, *Balanus amphitrite*. For comparison purposes, the authors illustrated the concept model in Figure 6B based on the reports by Dreanno et al. [45] and Matsumura et al. [46]. This figure was created using BioRender.com (accessed on 25 July 2022).

This present study highlights the importance played by PTMs such as phosphorylation that could contribute to the rapid cellular response of oyster larvae in finding suitable substrates similar to those reported in cypris larvae [88]. This observation on the importance of phosphorylation is consistent with the report of Thiyagarajan et al. [88,89] on barnacles. They hypothesized that protein phosphorylation is involved in regulating the cypris larval metamorphosis [89]. Exposure to a highly phosphorylated settlement cue activated phosphatase in searching cyprids which eventually led to a series of biochemical changes, attachment, and metamorphosis [88]. Furthermore, the rapid response to environmental cues and metamorphic morphogenesis of oyster larvae could be regulated by the presence of phosphorylated [88] and glycosylated [23,24,26,87] settlement cues on substrates and may explain why marine invertebrates, such as *C. gigas,* appear to require little or no *de novo* gene action during the metamorphic induction process [90]. Our previous study also supports this theory that extensive *de novo* synthesis of proteins is not necessary for the completion of metamorphosis [47] and which was found similar in other reported studies in marine

invertebrates [89,90]. Such rapid transition from pediveliger to postlarvae is to the advantage of these organisms that are particularly vulnerable to predation during this event [91]. In contrast, as shown in Figure 3.7 B, reported gregarious conspecific cues in other benthic organisms such as the barnacle SIPC presents a less complex lectin-glycan interaction during settlement induction as it is specifically modulated by LCA/Con A-binding sugars such as mannose [46,92].

These present findings show that the induction of conspecific cue mediated larval settlement in *C. gigas* requires an interplay of different amino acid groups, disulfide bonds, glycans, and phosphorylation crosstalk for recognition. On the other hand, the receptor that could recognize CGSPPC involves an unidentified WGA lectin-like larval receptor and a phosphoprotein. However, whether these sugar moieties and phosphate groups found in the components of CGSPPC could bind to the same larval receptor or not, remains to be clarified. The noncovalent bond interaction between the lectin-glycan and/or phosphate groups in both larva and surface-bound cues from substrates, i.e., shells of conspecifics, allows accumulation of transduced cues and creates a seemingly complex recognition system in *C. gigas*. This facilitates higher affinity, specificity, and complementarity between CGSPPC and its receptor to ensure greater selectivity during site selection. Moreover, our previous findings [26] and in this present study showed that in the absence of a settlement cue (seawater only), oyster larvae do not settle and undergo metamorphosis [26]. The metamorphic process that larvae need to undergo requires a lot of energy and varied morphogenetic changes. Thus, this study supports the current hypothesis that oyster larvae actively respond to a wide range of variables or cues that may provide information to secure a site appropriate for its post-settlement growth and survival [23,93]. Oysters may seem simple, but this study provides insights into the complex molecular mechanisms governing oyster substrate selection that may indicate their evolutionary success as benthic organisms. Continued study on CGSPPC and its corresponding oyster larval receptor will advance our understanding of their evolution, population dynamics, and chemical communication. Lastly, these results may find application in the development of oyster aquaculture by using this extracted compound as a surface-bound attractant to collect wild and hatchery-grown oyster larvae which could help recover declining marine species; this compound could also find application as a target of anti-fouling agents on man-made structures.

III.5 Conclusions

This study reports a novel role of several shell matrix proteins that acts as *Crassostrea gigas* Settlement Pheromone Protein Components (CGSPPC) and is demonstrated as the biological cue responsible for gregarious settlement on conspecifics. These isolated matrix proteins from the adult oyster shells include Gigasin-6 isoform X1 and/or X2, *Crassostrea gigas* Surface protein P12p-like (CGS12P), and several stains-all stainable acidic proteins, the most dominant of which was a 48 kDa band, putatively identified as *Crassostrea gigas* Dentin sialophosphoprotein (CGDSP). Each protein factor may individually induce larval settlement, but evidence also shows that they could synergistically interact with each other to form a macromolecular assemblage within the shell to create a stable and strong settlement inducing signal. Some key factors that are attributed to their ability to induce settlement, either individually or collectively, are the presence of unique amino acid groups and post-translational modifications that may play important roles in oyster chemical signaling for conspecific recognition. Gigasin-6 isoform X1 and/or X2 which showed positive binding to WGA lectin is considered the major settlement inducing cue component. Gigasin-6 isoform X1 and/or X2 plays a role in larval settlement induction on adult shells and may also function in postlarva-larva settlement interactions. CGDSP together with other stains-all stainable acidic proteins are also co-factors that could contribute to strengthening the settlement inducing signal via phosphorylation crosstalk with a CGSPPC recognizing larval receptor. CGDSP and some stains-all stainable acidic proteins isolated in this study require disulfide bonds to stabilize its settlement inducing effects and may play a role in providing the structural framework for other shell matrix proteins to anchor within the settlement protein components and provide protection to its multiple binding partners. The role of CGS12P as a co-factor in settlement induction warrants further investigation. Furthermore, conspecific cue-mediated larval settlement induction in *C. gigas* presents a complex system that requires an interplay of different amino acid groups, disulfide bonds, glycans, and phosphorylation crosstalk for conspecific recognition. On the other hand, the receptor that could recognize CGSPPC may involve a WGA lectin-like larval receptor and a phosphoprotein.

III.6 References

1. Hedgecock, D.; Gaffney, P.M.; Goulletquer, P.; Guo, X.; Reece, K.; Warr, G.W. The Case for Sequencing the Pacific Oyster Genome. *J. Shellfish Res.* 2005, *24*, 429–441, doi:10.2983/0730-8000(2005)24[429:TCFSTP]2.0.CO;2.
2. Zhang, G.; Fang, X.; Guo, X.; Li, L.; Luo, R.; Xu, F.; Yang, P.; Zhang, L.; Wang, X.; Qi, H.; et al. The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation. *Nature* 2012, *490*, 49–54, doi:10.1038/nature11413.
3. Grabowski, J.H.; Brumbaugh, R.D.; Conrad, R.F.; Keeler, A.G.; Opaluch, J.J.; Peterson, C.H.; Piehler, M.F.; Powers, S.P.; Smyth, A.R. Economic Valuation of Ecosystem Services Provided by Oyster Reefs. *Bioscience* 2012, *62*, 900–909, doi:10.1525/bio.2012.62.10.10.
4. Gentry, R.R.; Alleway, H.K.; Bishop, M.J.; Gillies, C.L.; Waters, T.; Jones, R. Exploring the Potential for Marine Aquaculture to Contribute to Ecosystem Services. *Rev. Aquac.* 2020, *12*, 499–512, doi:10.1111/raq.12328.
5. Ridlon, A.D.; Wasson, K.; Waters, T.; Adams, J.; Donatuto, J.; Fleener, G.; Froehlich, H.; Govender, R.; Kornbluth, A.; Lorda, J.; et al. Conservation Aquaculture as a Tool for Imperiled Marine Species: Evaluation of Opportunities and Risks for Olympia Oysters, *Ostrea lurida*. *PLoS One* 2021, *16*, e0252810, doi:10.1371/journal.pone.0252810.
6. Xu, F.; Zhang, G. Transcriptomic and Proteomic Dynamics during Metamorphosis of Pacific Oyster *Crassostrea gigas*. *bioRxiv* 2020, 2020.03.25.004614, doi:https://doi.org/10.1101/2020.03.25.004614.
7. Fitt, W.K.; Coon, S.L.; Walch, M.; Weiner, R.M.; Colwell, R.R.; Bonar, D.B. Settlement Behavior and Metamorphosis of Oyster Larvae (*Crassostrea gigas*) in Response to Bacterial Supernatants. *Mar. Biol.* 1990, *106*, 389–394, doi:10.1007/BF01344317.
8. Coon, S.L.; Bonar, D.B.; Weiner, R.M. Induction of Settlement and Metamorphosis of the Pacific Oyster, *Crassostrea gigas* (Thunberg), by L-DOPA and Catecholamines. *J. Exp. Mar. Bio. Ecol.* 1985, *94*, doi:10.1016/0022-0981(85)90059-0.
9. Joyce, A.; Vogeler, S. Molluscan Bivalve Settlement and Metamorphosis: Neuroendocrine Inducers and Morphogenetic Responses. *Aquaculture* 2018, *487*, 64–82, doi:10.1016/j.aquaculture.2018.01.002.
10. Yurchenko, O. V.; Skiteva, O.I.; Voronezhskaya, E.E.; Dyachuk, V.A. Nervous System Development in the Pacific Oyster, *Crassostrea gigas* (Mollusca: Bivalvia). *Front. Zool.* 2018, *15*, 10, doi:10.1186/s12983-018-0259-8.

11. Marie, B.; Zanella-Cléon, I.; Guichard, N.; Becchi, M.; Marin, F. Novel Proteins from the Calcifying Shell Matrix of the Pacific Oyster *Crassostrea gigas*. *Mar. Biotechnol.* 2011, *13*, 1159–1168, doi:10.1007/s10126-011-9379-2.
12. Wang, X.; Song, X.; WANG, T.; Zhu, Q.; Miao, G.; Chen, Y.; Fang, X.; Que, H.; Li, L.; Zhang, G. Evolution and Functional Analysis of the Pif97 Gene of the Pacific Oyster *Crassostrea gigas*. *Curr. Zool.* 2013, *59*, 109–115, doi:10.1093/czoolo/59.1.109.
13. De Wit, P.; Durland, E.; Ventura, A.; Langdon, C.J. Gene Expression Correlated with Delay in Shell Formation in Larval Pacific Oysters (*Crassostrea gigas*) Exposed to Experimental Ocean Acidification Provides Insights into Shell Formation Mechanisms. *BMC Genomics* 2018, *19*, 160, doi:10.1186/s12864-018-4519-y.
14. Beck, M.W.; Brumbaugh, R.D.; Airoldi, L.; Carranza, A.; Coen, L.D.; Crawford, C.; Defeo, O.; Edgar, G.J.; Hancock, B.; Kay, M.C.; et al. Oyster Reefs at Risk and Recommendations for Conservation, Restoration, and Management. *Bioscience* 2011, *61*, 107–116, doi:10.1525/bio.2011.61.2.5.
15. Hadfield, M.; Paul, V. Natural Chemical Cues for Settlement and Metamorphosis of Marine-Invertebrate Larvae. In *Marine Chemical Ecology*; McClintok, J.B., Baker, B.J., Eds.; CRC Press: Boca Raton: FL, USA, 2001; pp. 431–461 ISBN 9781420036602.
16. Prendergast, G.S. Settlement and Behaviour of Marine Fouling Organisms. In *Biofouling*; Wiley-Blackwell: Oxford, UK, 2010; pp. 30–59 ISBN 9781444315462.
17. Hadfield, M.G. Biofilms and Marine Invertebrate Larvae: What Bacteria Produce That Larvae Use to Choose Settlement Sites. *Ann. Rev. Mar. Sci.* 2011, *3*, 453–470, doi:10.1146/annurev-marine-120709-142753.
18. Highsmith, R.C. Induced Settlement and Metamorphosis of Sand Dollar (*Dendraster ecentricus*) Larvae in Predator-Free Sites: Adult Sand Dollar Beds. *Ecology* 1982, *63*, 329–337, doi:10.2307/1938950.
19. McQuaid, C.D.; Miller, K. Larval Supply and Dispersal. In *Biofouling*; Wiley-Blackwell: Oxford, UK, 2010; pp. 16–29 ISBN 9781444315462.
20. Thorson, G. Reproductive and Larval Ecology of Marine Bottom Invertebrates. *Biol. Rev.* 1950, *25*, 1–45, doi:10.1111/j.1469-185X.1950.tb00585.x.
21. Hay, M.E. Marine Chemical Ecology: Chemical Signals and Cues Structure Marine Populations, Communities, and Ecosystems. *Ann. Rev. Mar. Sci.* 2009, *1*, 193–212, doi:10.1146/annurev.marine.010908.163708.
22. Bayne, B.L. The Gregarious Behaviour of the Larvae of *Ostrea edulis* L. at Settlement. *J. Mar. Biol. Assoc. United Kingdom* 1969, *49*, 327–356,

doi:10.1017/S0025315400035943.

23. Vasquez, H.E.; Hashimoto, K.; Yoshida, A.; Hara, K.; Imai, C.C.; Kitamura, H.; Satuito, C.G. A Glycoprotein in Shells of Conspecifics Induces Larval Settlement of the Pacific Oyster *Crassostrea gigas*. *PLoS One* 2013, *8*, e82358, doi:10.1371/journal.pone.0082358.
24. Vasquez, H.E.; Hashimoto, K.; Kitamura, H.; Satuito, C.G. Wheat Germ Agglutinin-Binding Glycoprotein Extract from Shells of Conspecifics Induces Settlement of Larvae of the Pacific Oyster *Crassostrea gigas* (Thunberg). *J. Shellfish Res.* 2014, *33*, 415–423, doi:10.2983/035.033.0210.
25. Diederich, S. Differential Recruitment of Introduced Pacific Oysters and Native Mussels at the North Sea Coast: Coexistence Possible? *J. Sea Res.* 2005, *53*, 269–281, doi:10.1016/j.seares.2005.01.002.
26. Sedanza, M.G.; Kim, H.-J.; Seposo, X.; Yoshida, A.; Yamaguchi, K.; Satuito, C.G. Regulatory Role of Sugars on the Settlement Inducing Activity of a Conspecific Cue in Pacific Oyster *Crassostrea gigas*. *Int. J. Mol. Sci.* 2021, *22*, 3273, doi:10.3390/ijms22063273.
27. Crisp, D.J. Chemical Factors Inducing Settlement in *Crassostrea virginica* (Gmelin). *J. Anim. Ecol.* 1967, *36*, 329–335, doi:10.2307/2916.
28. Poirier, L.A.; Clements, J.C.; Davidson, J.D.P.; Miron, G.; Davidson, J.; Comeau, L.A. Sink before You Settle: Settlement Behaviour of Eastern Oyster (*Crassostrea virginica)* Larvae on Artificial Spat Collectors and Natural Substrate. *Aquac. Reports* 2019, *13*, doi:10.1016/j.aqrep.2019.100181.
29. Tamburri, M.N.; Luckenbach, M.W.; Breitburg, D.L.; Bonniwell, S.M. Settlement of *Crassostrea ariakensis* Larvae: Effects of Substrate, Biofilms, Sediment and Adult Chemical Cues. *J. Shellfish Res.* 2008, *27*, 601–608, doi:10.2983/0730-8000(2008)27[601:SOCALE]2.0.CO;2.
30. Hirata, Y.; Tamura, Y.; Nagasawa, K. Influence of Presoaking Conditions of Spat Collectors in Seawater Containing Adult Pacific Oysters (*Crassostrea gigas*) on Their Larval Settlement. *Nippon Suisan Gakkaishi (In Japanese)* 2008, *74*, 1017–1023, doi:10.2331/suisan.74.1017.
31. Hirata, Y. Advanced Techniques for Artificial Setting of Oyster *Crassostrea gigas*. *Nippon Suisan Gakkaishi (In Japanese)* 2005, *71*, 725–728, doi:10.2331/suisan.71.725.
32. Hirata, Y. Effective Pretreatments of Spat-Collectors for Induced Settlement of the Pacific Oyster, *Crassostrea gigas*. *Nippon Suisan Gakkaishi (In Japanese)* 1998, *64*,

610–617, doi:10.2331/suisan.64.610.

33. Zimmer-Faust, R.K.; Tamburri, M.N. Chemical Identity and Ecological Implications of a Waterborne, Larval Settlement Cue. *Limnol. Oceanogr.* 1994, *39*, 1075–1087, doi:10.4319/lo.1994.39.5.1075.

34. Turner, E.J.; Zimmer-Faust, R.K.; Palmer, M.A.; Luckenbach, M.; Pentchef, N.D. Settlement of Oyster (*Crassostrea virginica*) Larvae: Effects of Water Flow and a Water-soluble Chemical Cue. *Limnol. Oceanogr.* 1994, *39*, 1579–1593, doi:10.4319/lo.1994.39.7.1579.

35. Carroll, J.M.; Riddle, K.; Woods, K.E.; Finelli, C.M. Recruitment of the Eastern Oyster, *Crassostrea virginica*, in Response to Settlement Cues and Predation in North Carolina. *J. Exp. Mar. Bio. Ecol.* 2015, *463*, 1–7, doi:10.1016/j.jembe.2014.10.024.

36. Pascual, M.S.; Zampatti, E.A. Evidence of a Chemically Mediated Adult-Larval Interaction Triggering Settlement in *Ostrea puelchana*: Applications in Hatchery Production. *Aquaculture* 1995, *133*, 33–44, doi:10.1016/0044-8486(95)00005-M.

37. Veitch, F.P.; Hidu, H. Gregarious Setting in the American Oyster *Crassostrea virginica* Gmelin: I. Properties of a Partially Purified "Setting Factor." *Chesap. Sci.* 1971, *12*, 173, doi:10.2307/1350777.

38. Coon, S.L.; Walch, M.; Fitt, W.K.; Weiner, R.M.; Bonar, D.B. Ammonia Induces Settlement Behavior in Oyster Larvae. *Biol. Bull.* 1990, *179*, 297–303, doi:10.2307/1542321.

39. Myers, J.M.; Veis, A.; Sabsay, B.; Wheeler, A.P. A Method for Enhancing the Sensitivity and Stability of Stains-All for Phosphoproteins Separated in Sodium Dodecyl Sulfate-Polyacrylamide Gels. *Anal. Biochem.* 1996, *240*, 300–302, doi:10.1006/abio.1996.0361.

40. Johnstone, M.B.; Wheeler, A.P.; Falwell, E.P.; Staton, M.E.; Saski, C.A.; Mount, A.S. Folian-Cv1 Is a Member of a Highly Acidic Phosphoprotein Class Derived From the Foliated Layer of the Eastern Oyster (*Crassostrea virginica*) Shell and Identified in Hemocytes and Mantle. *Front. Mar. Sci.* 2019, *6*, 1–21, doi:10.3389/fmars.2019.00366.

41. Marin, F.; Luquet, G. Molluscan Shell Proteins. *Comptes Rendus Palevol* 2004, *3*, 469–492, doi:10.1016/j.crpv.2004.07.009.

42. Marin, F.; Le Roy, N.; Marie, B. The Formation and Mineralization of Mollusk Shell. *Front. Biosci. - Sch.* 2012, *4 S*, 1099–1125, doi:10.2741/s321.

43. Gotliv, B.-A.; Addadi, L.; Weiner, S. Mollusk Shell Acidic Proteins: In Search of Individual Functions. *Chembiochem* 2003, *4*, 522–529, doi:10.1002/cbic.200200548.

44. Rivera-Pérez, C.; Hernández-Saavedra, N.Y. Review: Post-Translational Modifications

of Marine Shell Matrix Proteins. *Comp. Biochem. Physiol. B. Biochem. Mol. Biol.* 2021, *256*, 110641, doi:10.1016/j.cbpb.2021.110641.

45. Dreanno, C.; Matsumura, K.; Dohmae, N.; Takio, K.; Hirota, H.; Kirby, R.R.; Clare, A.S. An α 2 -Macroglobulin-like Protein Is the Cue to Gregarious Settlement of the Barnacle *Balanus amphitrite*. *Proc. Natl. Acad. Sci.* 2006, *103*, 14396–14401, doi:10.1073/pnas.0602763103.
46. Matsumura, K.; Nagano, M.; Fusetani, N. Purification of a Larval Settlement-Inducing Protein Complex (SIPC) of the Barnacle, *Balanus amphitrite*. *J. Exp. Zool.* 1998, *281*, doi:10.1002/(SICI)1097-010X(19980501)281:1<12::AID-JEZ3>3.0.CO;2-F.
47. Sedanza, M.G.; Alshaweesh, J.; Gao, Y.; Yoshida, A.; Kim, H.; Yamaguchi, K.; Satuito, C.G. Transcriptome Dynamics of an Oyster Larval Response to a Conspecific Cue-Mediated Settlement Induction in the Pacific Oyster *Crassostrea gigas*. *Diversity* 2022, *14*, 559, doi:10.3390/d14070559.
48. Laemmli, U, K. Cleavage of Structural Proteins during the Assembly of the Head of Bacteriophage T4. *Nature* 1970, *227*, 680–685.
49. Yang, Z.; Hancock, W.S. Approach to the Comprehensive Analysis of Glycoproteins Isolated from Human Serum Using a Multi-Lectin Affinity Column. *J. Chromatogr. A* 2004, *1053*, 79–88, doi:10.1016/j.chroma.2004.08.150.
50. Campbell, K.P.; MacLennan, D.H.; Jorgensen, A.O. Staining of the Ca2+-Binding Proteins, Calsequestrin, Calmodulin, Troponin C, and S-100, with the Cationic Carbocyanine Dye “Stains-All”. *J. Biol. Chem.* 1983, *258*, 11267–11273, doi:10.1016/S0021-9258(17)44413-9.
51. Marin, F. Mollusc Shellomes: Past, Present and Future. *J. Struct. Biol.* 2020, *212*, 107583, doi:10.1016/j.jsb.2020.107583.
52. Unno, H.; Nakamura, A.; Mori, S.; Goda, S.; Yamaguchi, K.; Hiemori, K.; Tateno, H.; Hatakeyama, T. Identification, Characterization, and X-Ray Crystallographic Analysis of a Novel Type of Lectin AJLec from the Sea Anemone *Anthopleura japonica*. *Sci. Rep.* 2018, *8*, doi:10.1038/s41598-018-29498-0.
53. Yamaguchi, K. Preparation and Proteomic Analysis of Chloroplast Ribosomes. *Methods Mol. Biol.* 2011, *775*, doi:10.1007/978-1-61779-237-3_13.
54. León, I.R.; da Costa Neves-Ferreira, A.G.; da Rocha, S.L.G.; de Oliveira Trugilho, M.R.; Perales, J.; Valente, R.H. Using Mass Spectrometry to Explore the Neglected Glycan Moieties of the Antiophidic Proteins DM43 and DM64. *Proteomics* 2012, *12*, doi:10.1002/pmic.201200062.

55. Mann, K.; Edsinger-Gonzales, E.; Mann, M. In-Depth Proteomic Analysis of a Mollusc Shell: Acid-Soluble and Acid-Insoluble Matrix of the Limpet *Lottia gigantea*. *Proteome Sci.* 2012, *10*, 28, doi:10.1186/1477-5956-10-28.

56. Teufel, F.; Almagro Armenteros, J.J.; Johansen, A.R.; Gíslason, M.H.; Pihl, S.I.; Tsirigos, K.D.; Winther, O.; Brunak, S.; von Heijne, G.; Nielsen, H. SignalP 6.0 Predicts All Five Types of Signal Peptides Using Protein Language Models. *Nat. Biotechnol.* 2022, doi:10.1038/s41587-021-01156-3.

57. Blum, M.; Chang, H.Y.; Chuguransky, S.; Grego, T.; Kandasaamy, S.; Mitchell, A.; Nuka, G.; Paysan-Lafosse, T.; Qureshi, M.; Raj, S.; et al. The InterPro Protein Families and Domains Database: 20 Years On. *Nucleic Acids Res.* 2021, *49*, doi:10.1093/nar/gkaa977.

58. de Castro, E.; Sigrist, C.J.A.; Gattiker, A.; Bulliard, V.; Langendijk-Genevaux, P.S.; Gasteiger, E.; Bairoch, A.; Hulo, N. ScanProsite: Detection of PROSITE Signature Matches and ProRule-Associated Functional and Structural Residues in Proteins. *Nucleic Acids Res.* 2006, *34*, doi:10.1093/nar/gkl124.

59. Schultz, J.; Milpetz, F.; Bork, P.; Ponting, C.P. SMART, a Simple Modular Architecture Research Tool: Identification of Signaling Domains. *Proc. Natl. Acad. Sci. U. S. A.* 1998, *95*, doi:10.1073/pnas.95.11.5857.

60. Letunic, I.; Khedkar, S.; Bork, P. SMART: Recent Updates, New Developments and Status in 2020. *Nucleic Acids Res.* 2021, *49*, doi:10.1093/nar/gkaa937.

61. Krogh, A.; Larsson, B.; Von Heijne, G.; Sonnhammer, E.L.L. Predicting Transmembrane Protein Topology with a Hidden Markov Model: Application to Complete Genomes. *J. Mol. Biol.* 2001, *305*, doi:10.1006/jmbi.2000.4315.

62. Blom, N.; Gammeltoft, S.; Brunak, S. Sequence and Structure-Based Prediction of Eukaryotic Protein Phosphorylation Sites. *J. Mol. Biol.* 1999, *294*, doi:10.1006/jmbi.1999.3310.

63. Gupta, R.; Brunak, S. Prediction of Glycosylation across the Human Proteome and the Correlation to Protein Function. *Pac. Symp. Biocomput.* 2002, doi:10.1142/9789812799623_0029.

64. Steentoft, C.; Vakhrushev, S.Y.; Joshi, H.J.; Kong, Y.; Vester-Christensen, M.B.; Schjoldager, K.T.B.G.; Lavrsen, K.; Dabelsteen, S.; Pedersen, N.B.; Marcos-Silva, L.; et al. Precision Mapping of the Human O-GalNAc Glycoproteome through SimpleCell Technology. *EMBO J.* 2013, *32*, doi:10.1038/emboj.2013.79.

65. Erdos, G.; Pajkos, M.; Dosztányi, Z. IUPred3: Prediction of Protein Disorder Enhanced

with Unambiguous Experimental Annotation and Visualization of Evolutionary Conservation. *Nucleic Acids Res.* 2021, *49*, doi:10.1093/nar/gkab408.

66. Madeira, F.; Pearce, M.; Tivey, A.R.N.; Basutkar, P.; Lee, J.; Edbali, O.; Madhusoodanan, N.; Kolesnikov, A.; Lopez, R. Search and Sequence Analysis Tools Services from EMBL-EBI in 2022. *Nucleic Acids Res.* 2022, doi:10.1093/nar/gkac240.
67. Gao, Y.L.; Yoshida, A.; Liu, J.Y.; Shimizu, T.; Shirota, K.; Shiina, Y.; Osatomi, K. Molecular Cloning and Expression Dynamics of UNC-45B upon Heat Shock in the Muscle of Yellowtail. *Aquaculture* 2021, *541*, doi:10.1016/j.aquaculture.2021.736827.
68. Huan, P.; Wang, H.; Liu, B. A Label-Free Proteomic Analysis on Competent Larvae and Juveniles of the Pacific Oyster (*Crassostrea gigas*). *PLoS One* 2015, *10*, e0135008, doi:10.1371/journal.pone.0135008.
69. R Core Team R: A Language and Environment for Statistical Computing. *R Found. Stat. Comput. Vienna, Austria. ISBN 3-900051-07-0, Available online http//www.R-project.org/.* 2016, 2013.
70. Cabra, V.; Arreguin, R.; Vazquez-Duhalt, R.; Farres, A. Effect of Temperature and PH on the Secondary Structure and Processes of Oligomerization of 19 KDa Alpha-Zein. *Biochim. Biophys. Acta - Proteins Proteomics* 2006, *1764*, 1110–1118, doi:10.1016/j.bbapap.2006.04.002.
71. Tatsuno, R.; Yamaguchi, K.; Takatani, T.; Arakawa, O. RT-PCR- and MALDI-TOF Mass Spectrometry-Based Identification and Discrimination of Isoforms Homologous to Pufferfish Saxitoxin- and Tetrodotoxin-Binding Protein in the Plasma of Non-Toxic Cultured Pufferfish (*Takifugu rubripes*). *Biosci. Biotechnol. Biochem.* 2013, *77*, 208–212, doi:10.1271/bbb.120701.
72. Choi, Y.H.; Yamaguchi, K.; Oda, T.; Nam, T.J. Chemical and Mass Spectrometry Characterization of the Red Alga *Pyropia yezoensis* Chemoprotective Protein (PYP): Protective Activity of the N-Terminal Fragment of PYP1 against Acetaminophen-Induced Cell Death in Chang Liver Cells. *Int. J. Mol. Med.* 2015, *35*, 271–276, doi:10.3892/ijmm.2014.1992.
73. Magnelli, P.E.; Bielik, A.M.; Guthrie, E.P. Identification and Characterization of Protein Glycosylation Using Specific Endo- and Exoglycosidases. *J. Vis. Exp.* 2011, e3749, doi:10.3791/3749.
74. Dwek, R.A.; Edge, C.J.; Harvey, D.J.; Wormald, M.R.; Parekh, R.B. Analysis of Glycoprotein-Associated Oligosaccharides. *Annu. Rev. Biochem.* 1993, *62*, 65–100, doi:10.1146/annurev.bi.62.070193.000433.

75. Marie, B.; Luquet, G.; Pais De Barros, J.-P.; Guichard, N.; Morel, S.; Alcaraz, G.; Bollache, L.; Marin, F. The Shell Matrix of the Freshwater Mussel *Unio pictorum* (Paleoheterodonta, Unionoida). *FEBS J.* 2007, *274*, 2933–2945, doi:10.1111/j.1742-4658.2007.05825.x.

76. Sellinger, O.Z.; Wolfson, M.F. Carboxylmethylation Affects the Proteolysis of Myelin Basic Protein by *Staphylococcus aureus* V8 Proteinase. *Biochim. Biophys. Acta - Protein Struct. Mol. Enzymol.* 1991, *1080*, 110–118, doi:10.1016/0167-4838(91)90136-N.

77. Miyashita, T.; Takagi, R.; Okushima, M.; Nakano, S.; Miyamoto, H.; Nishikawa, E.; Matsushiro, A. Complementary DNA Cloning and Characterization of Pearlin, a New Class of Matrix Protein in the Nacreous Layer of Oyster Pearls. *Mar. Biotechnol.* 2000, *2*, 409–418, doi:10.1007/s101260000013.

78. Mouchi, V.; Lartaud, F.; Guichard, N.; Immel, F.; de Rafélis, M.; Broussard, C.; Crowley, Q.G.; Marin, F. Chalky versus Foliated: A Discriminant Immunogold Labelling of Shell Microstructures in the Edible Oyster *Crassostrea gigas*. *Mar. Biol.* 2016, *163*, 256, doi:10.1007/s00227-016-3040-6.

79. Schmitt, P.; Gueguen, Y.; Desmarais, E.; Bachère, E.; de Lorgeril, J. Molecular Diversity of Antimicrobial Effectors in the Oyster *Crassostrea gigas*. *BMC Evol. Biol.* 2010, *10*, 23, doi:10.1186/1471-2148-10-23.

80. Pancsa, R.; Schad, E.; Tantos, A.; Tompa, P. Emergent Functions of Proteins in Non-Stoichiometric Supramolecular Assemblies. *Biochim. Biophys. acta. Proteins proteomics* 2019, *1867*, 970–979, doi:10.1016/j.bbapap.2019.02.007.

81. Wojtas, M.; Dobryszycki, P.; Oyhar, A. Intrinsically Disordered Proteins in Biomineralization. In *Advanced Topics in Biomineralization*; InTech, 2012; pp. 1–32.

82. Alberts, B.; Johnson, A.; Lewis, J.; Morgan, D.; Raff, M.; Roberts, K.; Walter, P. *Molecular Biology of the Cell*; Wilson, J., Hunt, T., Eds.; 6th ed.; W.W. Norton & Company, 2017; ISBN 9781315735368.

83. Tompa, P. The Interplay between Structure and Function in Intrinsically Unstructured Proteins. *FEBS Lett.* 2005, *579*, 3346–3354, doi:10.1016/j.febslet.2005.03.072.

84. Totten, D.K.; Davidson, F.D.; Wyckoff, R.W. Amino-Acid Composition of Heated Oyster Shells. *Proc. Natl. Acad. Sci. U. S. A.* 1972, *69*, 784–785, doi:10.1073/pnas.69.4.784.

85. Tommasone, S.; Allabush, F.; Tagger, Y.K.; Norman, J.; Köpf, M.; Tucker, J.H.R.; Mendes, P.M. The Challenges of Glycan Recognition with Natural and Artificial Receptors. *Chem. Soc. Rev.* 2019, *48*, 5488–5505, doi:10.1039/C8CS00768C.

86. Taylor, M.E.; Drickamer, K.; Schnaar, R.L.; Etzler, M.E.; Varki, A. Discovery and Classification of Glycan-Binding Proteins. In *Essentials of Glycobiology [Internet]*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2017 ISBN 9780879697709.

87. Neal, A.L.; Yule, A.B. The Effects of Dissolved Sugars Upon the Temporary Adhesion of Barnacle Cyprids. *J. Mar. Biol. Assoc. United Kingdom* 1996, *76*, 649–655, doi:10.1017/S0025315400031350.

88. Thiyagarajan, V. A Review on the Role of Chemical Cues in Habitat Selection by Barnacles: New Insights from Larval Proteomics. *J. Exp. Mar. Bio. Ecol.* 2010, *392*, 22–36, doi:10.1016/j.jembe.2010.04.030.

89. Thiyagarajan, V.; Wong, T.; Qian, P.-Y. 2D Gel-Based Proteome and Phosphoproteome Analysis During Larval Metamorphosis in Two Major Marine Biofouling Invertebrates. *J. Proteome Res.* 2009, *8*, 2708–2719, doi:10.1021/pr800976u.

90. Hadfield, M.G.; Carpizo-Ituarte, E.J.; Carmen, K. Del; Nedved, B.T. Metamorphic Competence, a Major Adaptive Convergence in Marine Invertebrate Larvae. *Am. Zool.* 2001, *41*, 1123–1131, doi:10.1093/icb/41.5.1123.

91. Calado, R.; Leal, M.C. Trophic Ecology of Benthic Marine Invertebrates with Bi-Phasic Life Cycles: What Are We Still Missing? *Adv. Mar. Biol.* 2015, *71*, 1–70, doi:10.1016/bs.amb.2015.07.001.

92. Matsumura, K.; Mori, S.; Nagano, M.; Fusetani, N. Lentil Lectin Inhibits Adult Extract-Induced Settlement of the Barnacle, *Balanus amphitrite*. *J. Exp. Zool.* 1998, *280*, 213–219, doi:10.1002/(SICI)1097-010X(19980215)280:3<213::AID-JEZ2>3.0.CO;2-Q.

93. Steinberg, P.D.; De Nys, R.; Kjelleberg, S. Chemical Cues for Surface Colonization. *J. Chem. Ecol.* 2002, *28*, 1935–1951, doi:https://doi.org/10.1023/A:1020789625989.

Chapter IV: Transcriptome dynamics of an oyster larval response to conspecific cue-mediated settlement induction in Pacific oyster *Crassostrea gigas*

IV.1 Introduction

The Pacific oyster *Crassostrea gigas* is a benthic bivalve mollusk with a biphasic life cycle. It is one of the most well-studied organisms due to its biology and aquaculture importance [1,2]. The biphasic life cycle of *C. gigas* is composed of a planktonic larval stage for dispersion and a sessile adult stage, with metamorphosis marking the transition between the two life stages in a relatively short time (generally less than 48 h) [3–5]. As oyster larvae reach the pediveliger stage, they begin a series of swimming motions by sinking to the bottom and extending their foot in search of a suitable surface to settle on [6,7]. Competent pediveliger larvae manifest such behaviors. When environmental cues meet the requirements for a suitable settlement surface, larvae undergo irreversible metamorphic changes that include loss of the velum, foot, and eyespot, development of gills, and the production of adult shells [8]. In addition, the metamorphic transition often involves morphological, physiological, structural, and functional changes that are governed by a gene regulatory network [9].

A variety of chemical cues from the natural environment have been implicated to play a role in the induction of larval settlement in many marine invertebrate species. These cues may come from microbial or bacterial biofilms [10–16], macroalgal hosts [17,18], and conspecifics [19–24]. Of the chemical cues reported from conspecifics, surface-bound cues from various oyster shells have been reported to induce larval settlement in *C. gigas* [25–27], *C. virginica* [20,28], and *C. ariakensis* [29]. These studies show that chemical cues from conspecifics may play a role in the gregarious behavior exhibited by oyster larvae upon settlement [20,25–29]. However, the exact nature of these cues is not yet fully understood and characterized. Understanding the effects of these cues on the larvae can help clarify the larval settlement mechanism of oysters, and these compounds may find application in the development of oyster aquaculture.

Several studies have also explored possible mechanisms underlying larval settlement in oysters, which include the involvement of hormones, neurotransmitters, genes, and signaling pathways [4,8,11,30–36]. The first coherent theory for settlement and metamorphosis was made by Bonar et al. [4], in which two independent nervous-system-related pathways were necessary for oyster species: first, the dopaminergic pathway, which governs settlement behavior and attachment to surfaces. This often involves larval exposure to L-DOPA,

epinephrine, norepinephrine, and dopamine compounds. It is then followed by a second subsequent step, the adrenergic pathway, which regulates metamorphosis. However, they have also reported that exposure to norepinephrine or epinephrine led the oyster larvae to metamorphose directly without attaching to any surface, thereby bypassing the dopaminergic pathway. This theory was later supported by various pharmacological studies [8,11,37,38]. However, there is a lack of molecular evidence to connect these findings to actual pathways.

Recent advances in sequencing technology have paved the way for us to further explore a wider and deeper view of changes in gene expression patterns and signal pathways involved in the various larval development stages [32,39,40]. Transcriptome analysis is one of the useful technologies that may benefit us in gaining a comprehensive view of animal development and growth [30,32,41–43]. There are a growing number of molecular studies implicating the involvement of nuclear receptors and the ecdysone signaling pathway in oyster settlement [30,33,36,44]. Using the available genome and transcriptome databases, Vogeler et al. [8] have shown that the ecdysone receptor (EcR), as well as other EcR-cascade-related transcription factors, are present in *C. gigas*. They have also demonstrated that *C. gigas* EcR is highly expressed in the eyed pediveliger just before metamorphosis, suggesting its potential role in the induction of metamorphosis in bivalves. Despite this notable advance in our understanding of the gene regulatory network involved in oyster larval settlement, most of these studies use pharmacological inducing compounds, such as epinephrine, as reported in *C. gigas* [33] and *C. angulata* [31], as well as biofilm inducing cues for *C. gigas* [5]. Moreover, oysters are observed to attach and grow on top of each other's shells and build oyster reef-like structures in the natural environment. Most oyster farms still use scallops or oyster shells as settlement substrates for oyster culture. These shells are often put in cages for raft or long-line culture. Mechanisms that facilitate gregarious larval settlement under this condition are still fragmentary, especially regarding their molecular underpinnings. Previous studies by Vasquez et al. [25,26] and Sedanza et al. [27] have shown that a protein extract from the shells of conspecifics could induce larval attachment and its subsequent metamorphosis in *C. gigas* competent pediveligers. Currently, we are continuing to identify the exact nature of this conspecific cue from shells of *C. gigas* [27]. To date, there has been no available report yet that explores the effect, at a molecular level, of the exposure of competent oyster larvae to a natural settlement-inducing cue from conspecifics, such as those protein extracts found in shells, that may play a role in mediating settlement. Hence, our aim in this study is to fill in some gaps in our knowledge on the underlying specific cellular and physiological processes involved in the development of marine invertebrates, specifically on *C. gigas*, with the aid of the 'omics'

approach, such as transcriptomics. In the present study, we applied DNBseq-G400RS high-throughput sequencing technology to evaluate the effect of settlement induction by a conspecific cue on *C. gigas* larvae, before and after its recognition, and compare the transcriptomes of the pediveliger and conspecific cue-induced postlarvae. We also carried out gene function annotation, Gene Ontology enrichment, and real-time PCR verification analyses. Moreover, we identified key genes and molecular events involved in this process and proposed the possible involvement of a neuroendocrine-biomineralization crosstalk in *C. gigas* larval settlement.

IV.2 Materials and Methods

IV.2.1 Larval Culture and Shell Extract Preparation

Artificial fertilization and larval culture were conducted as previously described by Sedanza et al. [27], where two male and five female adult oysters were used as parents to create a high genetic diversity. Stripped gametes from a combination of several parents can avoid genetic diversity being too low [5]. In brief, gonads were stripped from all the adult oysters to collect gametes [45]. Eggs and sperm were separately suspended in 2 L glass beakers containing GF/C (Whatman glass fiber filter; pore size: 1.2 mm) filtered seawater (FSW) adjusted to 27 °C. Eggs were washed several times with FSW through repeated decantation and were then fertilized with a small volume of sperm suspension. Thirty minutes after artificial fertilization, fertilized eggs were collected in a 20 μm net, washed four to five times with FSW, and re-suspended in 2 L glass beakers containing FSW. Fertilized eggs were kept at 27 ± 1 °C in an incubator for 24 h. Then, the swimming straight-hinged larvae were collected in a 40 μm mesh, gently washed with FSW, and stocked in four 30 L tanks filled with 10 L FSW, at a density of 5 larvae mL^{-1}. The *C. gigas* larvae were cultured in these tanks under dark conditions with constant aeration. The larvae were fed with *Chaetoceros calcitrans* (10,000–50,000 cells/mL/day) from day 1 to day 5, fed a combination of *C. calcitrans* (25,000 cells/mL/day) and *Chaetoceros gracilis* (25,000 cells/mL/day) from day 6 to day 10, and were then fed *C. gracilis* (50,000 cells/mL/day) from day 11 onward during the culture period. Water quality parameters during the larval culture period were salinity (30–32 psu), pH (8.02–8.08), and temperature (25–27 °C). The seawater was renewed daily prior to feeding. Small larvae and empty shells that passed through 100 and 200 μm mesh screens in the 1st week and 2nd weeks of culture, respectively, were removed so as to obtain individuals that were of uniform size until they reached a competent stage. Competent larvae reached the eyed pediveliger stage on the 16^{th} to 20^{th} day of culture, when they reached 270 μm [46,47], which was consistent with the growth in hatcheries in Japan. Competent larvae were used for settlement experiments 24 days after the start of culture when they reached an average shell height of 322 μm [27]. Competence is defined as the ability to metamorphose following exposure of larvae to a demonstrated natural inducer [48]. All the larvae were actively swimming prior to conducting the settlement assay.

The preparation of the *C. gigas* crude shell EDTA extract (CgSE) was carried out following the method outlined by Sedanza et al. [27]. In brief, clean and dried *C. gigas* shells were crushed with a hammer until shell chips of about 0.5 to 1.0 mm in size were collected.

These shell chips (150 g) were decalcified with 1 L of 0.8 M ethylenediaminetetraacetic acid (EDTA, pH 8.0) for 60 h at 4 °C with continuous agitation. The supernatant was collected by centrifugation at 10,000× *g* for 60 min at 4 °C. Then, it was subsequently filtered and dialyzed against distilled water at 4 °C for 3 days. CgSE was lyophilized and the resultant powder was dissolved in as small a volume as possible of distilled water and stored at -20 °C until further use.

IV.2.2. Sample Collection and Research Design

The research design workflow is shown in Figure 4.1. This study was framed to evaluate the effect on larval settlement induction when a natural settlement-inducing cue is present, such as those from the shells of conspecifics (CgSE). Under laboratory conditions, my previous study showed that, in the absence of this conspecific cue, no larvae were induced to settle when immersed in 0.22 μm filtered seawater (FSW) [27]. Whereas when it is present, more than 80% of the larvae were induced to attach to the substrate and metamorphose after 24 h incubation [27]. Thus, after 24 h, cement substances could be observed surrounding one side of the oyster shell attached to the substrate and new shell growth in the postlarvae [27]. Results in this previous study were consistent with those reported by Coon et al. [48]. They have determined the temporal relationship between the onset of behavioral and morphogenetic competence in *C. gigas* using a neuroactive compound and have demonstrated that, in the absence of chemical stimulation, cultured oyster larvae were able to delay metamorphosis and maintain competence for at least 30 days [48]. In addition, no metamorphosed *C. gigas* larvae were recorded in the absence of chemical inducers in the study of Beiras and Widdows [37]. Hence, in this present study, the transcriptome profiles were compared between the larval batches (a) before exposure to the presence of the conspecific cue, which means they were not exposed to a chemical stimulant and, as such, hereon designated as the control and as pediveliger (Pedi), and (b) settled individuals after exposure to the conspecific cue, CgSE, hereon called postlarvae (PL). Therefore, it was necessary to have an equal start by using larval batches of the same age. Hence, the Pedi sample group was collected for RNA-seq analysis on the 24th DOC, while the PL sample group was exposed to CgSE at 24 DOC and collected after 24 h. In this study, it was not possible to carry out sampling at subsequent time points, which would have been an interesting approach, as it was hard to collect transitioning samples in real time because oyster settlement processes are quick and usually nonsynchronous. Hence, only PL samples after 24 h exposure were collected.

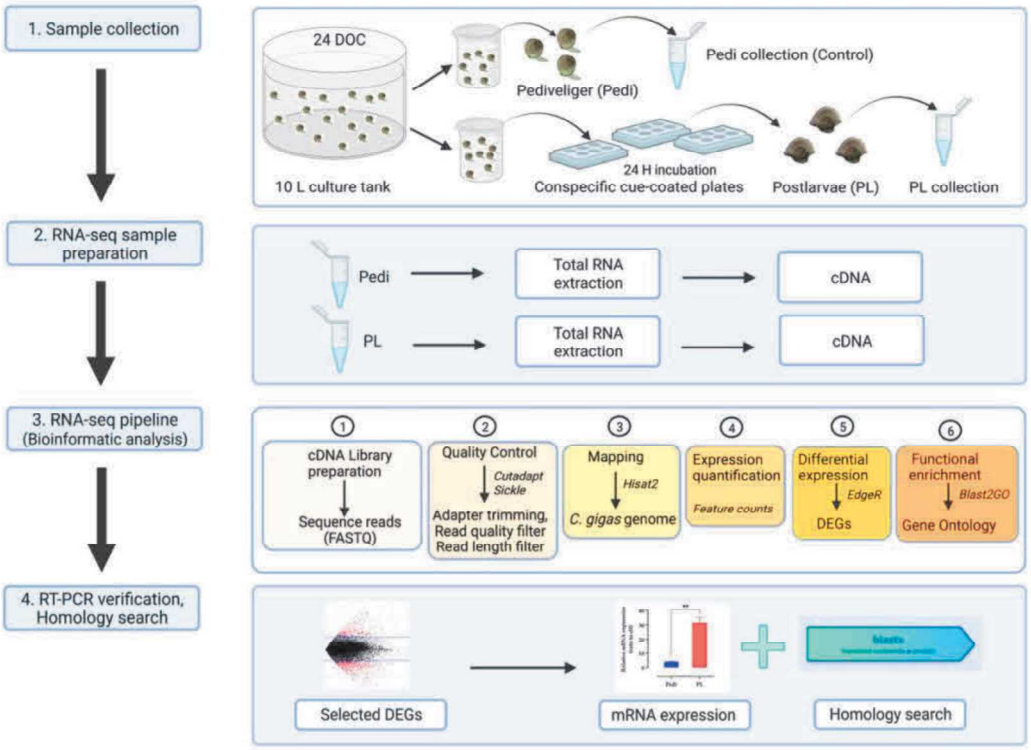

Figure 4.1. Research design workflow. This figure was created using BioRender.com.

The method design was as follows: On the day of the assay, all larvae were not fed. These larvae were at 24 days post-fertilization (or 24 days of culture, DOC) when approximately 70–80% have developed an eyespot and started to crawl. To ensure that all the larvae were at a competent stage, larvae from multiple 10-L FSW-filled tanks were sieved over a 300 µm mesh-covered tip nestled on a glass bowl filled with FSW. They were carefully washed in Millipore FSW (0.22 µm) and transferred to several 3-L beaker containers newly filled with FSW. After filtration, more than 95% of the larvae were verified under the microscope to have reached the pediveliger stage. The larval samples were then divided into two batches. Beakers containing the control batch of competent pediveliger larvae used for RNA-seq analysis were siphoned over a 300 µm mesh-covered tip nestled on a glass bowl filled with FSW, carefully washed with distilled water, and collected as a pooled Pedi sample. Approximately 100,000 individuals of pooled pediveliger larvae (183 mg wet weight) in a microcentrifuge tube were treated immediately with RNAlater Stabilization Solution (Invitrogen, Vilnius, Lithuania) for 24 h at 4 °C following the manufacturer's protocol. After which, RNAlater was removed and the pooled Pedi batch was stored at -80 °C prior to the subsequent experiments. For the PL batch, the method previously described by Sedanza et al. [27] was adopted for the induction of oyster larvae with CgSE. Briefly, 6-well plates were pre-coated with 50 µg CgSE and subsequently

filled with 10 mL of filtered seawater per well. Twenty larvae were released in each well and induced to settle on the same day as the collection of the Pedi larvae, i.e., at 24 DOC. In this study, we define settlement as the sequential transition of a competent pediveliger larva to a postlarva, which includes the attachment to a substrate and then eventual metamorphosis.

Postlarvae were confirmed under the microscope as individuals that secreted cement substances or those with postlarval shell growth. We observed about 80% settlement success after induction. The remaining larvae that were not induced to settle were found actively swimming after 24 h. No mortality was observed. All postlarvae that metamorphosed after 24 h were carefully removed from the 6-well plate bases. After removal of the postlarvae from each well, they were pipetted out to a 300 μm mesh-covered tip nestled on a glass bowl filled with FSW, washed gently several times with distilled water to remove any traces of the protein extract, and were transferred to a clean microcentrifuge tube. After the removal of the excess liquid on the tube, the postlarvae were immediately treated with RNAlater. This whole process was repeated several times until all the samples were collected. After 24 h, a similar procedure to that of the Pedi batch was carried out, wherein all the postlarvae were pooled and RNAlater was removed. Approximately 3000 individuals of postlarvae, equivalent to about 100 mg of pooled sample, were stored at -80 °C until further use.

IV.2.3. RNA Extraction, cDNA Library Preparation, Sequencing, and Bioinformatic Analysis

Total RNA extraction was performed by treating one hundred milligrams of each pediveliger (Pedi) and postlarvae (PL) samples with ISOGEN II (Nippongene, Tokyo, Japan) following the manufacturer's instructions. A 100 mg sample of each pooled Pedi and PL batches contained approximately 50,000 and 3000 individuals, respectively. Total RNA concentration was measured using a Quantus Fluorometer and QuantiFluor RNA system (Promega, Madison, WI, USA). The quality was then confirmed using the 5200 Fragment Analyzer System and the Agilent HS RNA kit (Agilent Technologies, Santa Clara, CA, USA). The cDNA library was prepared using the MGIEasy RNA Directional Library Prep Set (MGI Tech Co., Ltd., Shenzhen, China) following the instruction manual. The Qubit 3.0 Fluorometer and dsDNA HS Assay kit (Thermo Scientific, Waltham, MA, USA) were used to measure the concentration of the prepared library solution. The Fragment Analyzer and dsDNA Reagent 915 kit (Advanced Analytical Technologies, Ames, IA, USA) were used to check the quality of the prepared libraries. The DNA-depleted and purified RNA was used to construct the single-stranded circular DNA library using the MGIEasy Circulation kit (MGI Tech Co., Ltd., Shenzhen, China) to generate the DNA nanoball (DNB)-based libraries. DNB libraries were

sequenced and analyzed using the DNBSEQ-G400RS High-throughput sequencing kit and platform (MGI Tech Co., Ltd., Shenzhen, China), with a set 2 × 100 bp condition.

Adapter sequences were removed using cutadapt (version 1.9.1) (Martin, M. [49], Dortmund, Germany); then, sickle (ver 1.33) [50], was used to remove bases with a quality score of less than 20 and paired reads with less than 40 bases. Cleaned and high-quality sequencing reads were mapped to the *C. gigas* genome (https://www.ncbi.nlm.nih.gov/assembly/GCF_902806645.1/, accessed on 25 February 2022) with Hisat2 (version 2.2.1, Pertea et al. [51], Maryland, USA). Featurecounts (version 2.0.0, Liao et al. [52], Melbourne, Australia) was used to count the reads that mapped to the gene regions of the reference sequence. EdgeR [53] was used to perform differential analysis of sequence read count data using the R software [54]. EdgeR uses the quantile-adjusted conditional maximum likelihood (qCML) method for experiments with a single factor [53]; qCML is the most reliable in terms of bias on a wide range of conditions and, specifically, performs best in the situation of many small samples with a common dispersion, the model which is applicable to the Next-Gen sequencing data [53]. Transcripts with more than 2 M count-per-million (CPM) were considered expressed and were used for further analysis. A trimmed mean of M-values (TMM) was used to determine the scale factors in normalizing the library sizes and produce an effective library size that minimizes the log-fold changes between the samples. To test for differentially expressed genes (DEGs), a dispersion value for the biological coefficient of variation (BCV or square-root dispersion) was set at 0.25 [53]. Likewise, the threshold values of the false discovery rate (FDR) < 0.01, the absolute value of the log-fold of change |log2 Ratio| ≥ 2 for DEGs analysis, and |log2 Ratio| > 1.5 for Gene Ontology enrichment analysis were set as conditions to determine the upregulated and downregulated transcripts. These sets of selected transcripts from each batch, i.e., Pedi and PL, were then subjected to Gene Ontology (GO) enrichment analysis using Blast2GO [55]. Fischer's exact test was used to evaluate the significance of enrichment levels of DEGs under each GO term. Multiple testing p-values were adjusted based on the Benjamini and Hochberg method (adjusted $p < 0.01$) [56]. In this study, DEGs were identified at the transcript level [57].

The top thirty transcripts, and those that were significantly enriched after GO analysis and were annotated as uncharacterized were further subjected to a homology search and used as a query to search the non-redundant protein databases available at the National Center for Biotechnology Information (https://blast.ncbi.nlm.nih.gov/Blast.cgi/, accessed on 17 April 2022) using the BlastX algorithm with an E-value ≤ 10^{-3} [58].

IV.2.4. Quantitative Real-Time Polymerase Chain Reaction (qRT-PCR)

The total RNA used in this experiment was reverse-transcribed following the method of Gao et al. [59]. In brief, all RNA samples were measured for quantity and quality by a NanoDrop ND-1000 spectrometer (Thermo Fisher Scientific, Wilmington, DE, USA). The cDNA was synthesized by treating the total RNA sample (500 ng) with oligo-dT primer CDS-BR (5-GGCCACGCGTCGACTAGTAC(T)16–3′), random primer (5-NNNNNNNNN-3′), and M-MLV reverse transcriptase (Promega, Madison, WI, USA). Six candidate transcripts were chosen for real-time PCR verification to evaluate the transcriptome sequence integrity and expression trends indicated by RNA-seq. Selected differentially expressed transcripts were those that showed a log-fold of change (logFC) ≥ 2 and FDR < 0.01 and those that were significantly associated with GO-enriched categories. We designed gene-specific primers for real-time PCR by Primer-BLAST (https://www.ncbi.nlm.nih.gov/tools/primer-blast/, accessed on 17 March 2022) with default settings, except for PCR product size (100–150 bp). The list of the genes and their respective primers is found in Table 4.1.

Table 4.1. Primers of selected genes used for qPCR verification.

Accession No.	Annotation	Primer sequence (5' – 3')	Product size	Reference
LOC105332577	Chit1-SP	AGGAAATGGGAACAGAAGTTAG	139	-
	Chit1-AP	AGGACGGTTTGTCAATGGGAGG		
LOC105327998	Calm A-SP	GAGTATTCCGTGTTTCCAATG	150	-
	Calm A-AP	TCTGTTAGCTGTTCTGCCATGT		
LOC105335791	A-A4-SP	ATTATACAGGTCTTGGGAAAC	155	-
	A-A4-AP	GCAGGGGGTCTCATACAA		
LOC105331104	CA1-XII-SP	GACGAAATACGGGCACTCAAAG	145	-
	CA1-XII-AP	TTGTCTCCAATGTGACTC		
LOC105327922	VWC2-SP	TTGCTGTCCAGTGTGTCCTAAC	114	-
	VWC2-AP	CCGAAATGTGAACAATGAC		
LOC105335286	Wif-1-SP	CGAATGAACGACCACTGCAACAAGTC	133	Huan et al. [5]
	Wif-1-AP	TGTCTGTCTGTCCAGGCTGTAGGC		
LOC105338957	Elf-SP	CGAGAAGGAAGCTGCTGAGATGG	208	Huan et al. [5]
	Elf-AP	ACAGTCAGCCTGTGAAGTTCCTGTA		

Abbreviations: Chit1—putative chitinase 1; Calm A—calmodulin A; A-A4—Annexin A4, CA1-XII—Collagen alpha 1 (XII) chain; VWC2L—von Willebrand factor C domain-containing protein 2-like; Wif-1—Wnt inhibitory factor 1/EGF-like domain-containing protein 2 (primers used were as reported by Huan et al. [5]); *elf*—elongation factor 1 (as reference gene, primers used were as reported by Huan et al. [5]).

The qRT-PCR assay was conducted using the Power SYBR Green PCR Master Mix and a QuantStudio 3 Real-Time PCR System (Applied Biosystems, Thermo Fisher Scientific, Rockford, IL, USA) following the manufacturer's instructions. Thermocycling was performed as follows: enzyme activation for 10 min at 95 °C, denature step for 15 s at 95 °C, and annealing and extension for 1 min at 60 °C for 40 cycles. All data were normalized with the *C. gigas elf*

gene as an internal reference. The mRNA expression levels were quantified by the standard curve method, and the relative mRNA expression levels were calculated by the ratio to the *elf* reference gene. All data were presented as the mean ± standard deviation (SD). This experiment was conducted in triplicate. The comparison of the relative mRNA expression levels between pediveliger and postlarvae samples of each DEG was analyzed using a Student's t-test and was determined using the GraphPad 9 software program (GraphPad Prism, Software Inc., San Diego, CA, USA). The figure generated in the qRT-PCR assay was also performed using the same software program.

IV.3 Results

IV.3.1 Transcriptome Sequencing and Mapping

In this study, the cDNA pool of *C. gigas* pediveliger (Pedi) and postlarvae (PL) samples were subjected to RNA-seq analysis and were compared to obtain a comprehensive view of the transcriptome profile upon larval settlement induction. The statistical information of the sequencing and mapping results is given in Table 4.2. The average number of pair reads from each library was 22,987,478 and the average number of high-quality pair reads was 22,495,268. The average total mapping rate of the libraries at Q20 and Q30 values were 97.1% and 89.4%, respectively. There was also a high percentage of the cleaned and filtered number of reads in both samples that were successfully mapped to the *C. gigas* reference genome (Table 4.2). The results showed that the RNA-seq data could be used for further analysis.

Table 4.2. Statistics summary of the Pacific oyster, *Crassostrea gigas*, transcriptome.

Sample	Sequencing Results			Quality Filter Results	Summary of Mapping Results		
	No. of Pair Reads	>=Q20 (%)	>=Q30 (%)	High-Quality Pair Reads	No. of Reads Not Mapped to the Reference Array	No. of Reads Mapped to the Reference Array (One Site)	No. of Reads Mapped to the Reference Array (Multiple Sites)
Pediveliger (Pedi)	23,141,856	97.0	89.0	22,628,121	8,184,886 (36.17%)	13,473,074 (59.54%)	970,161 (4.29%)
Post larvae (PL)	22,833,100	97.2	89.7	22,362,414	8,126,324 (36.34%)	13,440,437 (60.10%)	795,653 (3.56%)
Average	22,987,478	97.1	89.4	22,495,268	8,155,605 (36.25%)	13,456,756 (59.82%)	882,907 (3.93%)

Different normalization models were also used to find the effective library size in analyzing the differentially expressed transcripts in the transcriptome data. The supplemental Figure S1 file shows in the histogram profile that the average log CPM of each transcript was four CPM. The trimmed mean of M-values (TMM) in this figure shows the best normalization model fit for the downstream analysis of differentially expressed transcripts in this experiment. The raw reads of the RNA-seq data are openly available in the public domain NCBI Sequence Read Archive, BioProject ID: PRJNA841258 (https://www.ncbi.nlm.nih.gov/sra/PRJNA841258, accessed 22 May 2022).

IV.3.2 Identification and Analysis of the Differentially Expressed Genes (DEGs) in the Pediveliger (Pedi) and Postlarvae (PL)

Post quality control analysis yielded 39,108 transcripts. The transcripts with differential expression were identified in the Pedi and PL batches by setting a threshold filtering value of $|\log2\ Ratio| \geq 2$ and Benjamini–Hochberg false discovery rate correction (FDR) < 0.01 after TMM normalization. Transcriptome results showed a total of 2383 transcripts were differentially expressed between pre-settlement pediveliger larvae (Pedi) and postlarvae (PL). There were 740 upregulated and 1643 downregulated transcripts after settlement, while the rest of the 36,725 transcripts were found common in Pedi and PL batches. The full list of transcripts tested for differential expression is attached in Appendix Table A4.1, which is expected to provide support for related research. The MA plot and heat map profiles of the transcripts with differential expression are found in Figure 4.2. A partial list showing the top thirty transcripts with differential expression in the Pedi and PL batches is shown in Table 4.3. A blast homology search was performed on some of the uncharacterized genes in this list and its full annotation is provided in Appendix Table A4.2.

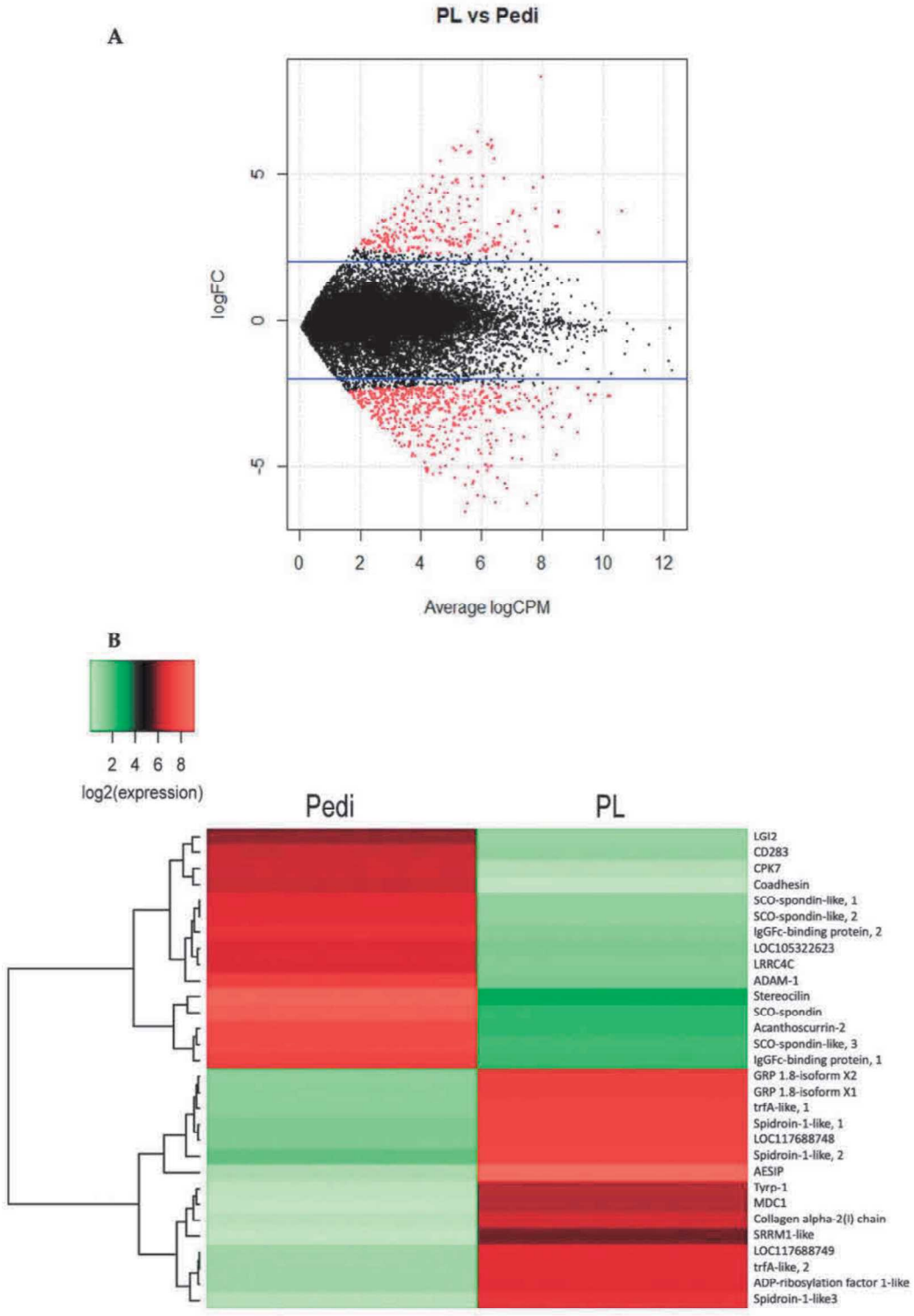

Figure 4.2. MA plot of transcripts with differential expression in *Crassostrea gigas* after settlement induction by a conspecific cue (**A**) and heat map expression pattern of the top thirty transcripts with differential expression (**B**). (**A**) Blue lines indicate transcripts that were filtered with a threshold value set at FDR < 0.01. The red dots indicate those transcripts that were considered as differentially expressed and met the two filtration threshold criteria: log-fold of change (logFC) ≥ 2 and FDR < 0.01. Red dots shown with a positive and negative logFC denote transcripts that were upregulated and downregulated, respectively, after settlement induction. Abbreviation: Average log CPM = log count-per-million; (**B**) The list of the top thirty transcripts with differential expressions were those that showed the highest logFC and lowest

FDR values. The color intensity from green to red indicates low to highest logFC of expression in Pedi and PL batches. Transcripts with similar expression levels were clustered as shown in the tree on the left of Figure 4.2. The unabbreviated names of the transcripts in this heat map are provided in Table 4.3.

Table 4.3. Top thirty transcripts with differential expression between pre-settlement pediveliger larvae (Pedi) and postlarvae (PL) in Pacific oyster, *C. gigas*.

Accession No.[a]	Annotation	Reference species	Evalue[b]	logFC[c]	FDR[d]
XP_011437564.2	Abscisic acid and environmental stress-inducible protein (AESIP)	*Crassostrea gigas*	0	9	1.05×10^{-26}
XP_034329440.1	Coadhesin	*Crassostrea gigas*	0	–6.56	3.83×10^{-18}
XP_034330544.1	SCO-spondin	*Crassostrea gigas*	0	–6.29	3.83×10^{-18}
XP_045201952.1	General transcriptional corepressor trfA-like, 1 (trfA-like, 1)	*Mercenaria mercenaria*	3×10^{-29}	6.44	3.83×10^{-18}
XP_034329445.1	A disintegrin and metalloproteinase with thrombospondin motifs adt-1 (ADAM-1)	*Crassostrea gigas*	0	–6.23	1.07×10^{-17}
XP_011427257.2	Calcium-dependent protein kinase 7 (CPK7)	*Crassostrea gigas*	2×10^{-37}	–6.28	1.44×10^{-17}
XP_034323067.1	Spidroin-1-like, 1	*Crassostrea gigas*	2×10^{-08}	6.12	2.88×10^{-17}
OWF40357.1	Stereocilin	*Mizuhopecten yessoensis*	0	–6.02	2.90×10^{-17}
XP_011422608.2	IgGFc-binding protein	*Crassostrea gigas*	6×10^{-135}	–6.05	4.55×10^{-17}
XP_022339848.1	SCO-spondin-like, 1	*Crassostrea virginica*	0	–5.99	7.63×10^{-17}
XP_022339848.1	SCO-spondin-like, 2	*Crassostrea virginica*	0	–5.99	7.63×10^{-17}
XP_045201952.1	General transcriptional corepressor trfA-like, 2 (trfA-like, 2)	*Mercenaria mercenaria*	9×10^{-31}	5.96	7.63×10^{-17}
XP_034323070.1	Spidroin-1-like, 2	*Crassostrea gigas*	7×10^{-13}	5.97	7.63×10^{-17}
XP_034322956.1	uncharacterized LOC117688748	*Crassostrea gigas*	0	5.95	7.71×10^{-17}
XP_034322955.1	Uncharacterized LOC117688749	*Crassostrea gigas*	0	5.85	1.81×10^{-16}
XP_022339848.1	SCO-spondin-like, 3	*Crassostrea virginica*	0	–5.79	2.11×10^{-16}
XP_034313686.1	ADP-ribosylation factor 1-like	*Crassostrea gigas*	3×10^{-136}	5.93	2.28×10^{-16}
XP_011426386.2	Mediator of DNA damage checkpoint protein 1 (MDC1)	*Crassostrea gigas*	0	5.87	5.08×10^{-16}
XP_034323479.1	Glycine-rich cell wall structural protein 1.8 transcript isoform X2 (GRP 1.8- isoform X2)	*Crassostrea gigas*	0	5.74	7.33×10^{-16}
XP_011450501.3	Glycine-rich cell wall structural protein 1.8 transcript isoform X1 (GRP 1.8-isoform X1)	*Crassostrea gigas*	0	5.74	7.33×10^{-16}
APC92582.1	Tyrp-1	*Hyriopsis cumingii*	2×10^{-157}	5.74	7.66×10^{-16}
XP_011436430.2	Collagen alpha-2(I) chain	*Crassostrea gigas*	6×10^{-37}	5.70	1.03×10^{-15}
XP_011422608.2	IgGFc-binding protein	*Crassostrea gigas*	0	–5.58	1.40×10^{-15}
XP_011452538.2	Leucine-rich repeat-containing protein 4C (LRRC4C)	*Crassostrea gigas*	0	–5.63	2.00×10^{-15}

XP_011419747.2	Uncharacterized LOC105322623	*Crassostrea gigas*	0	–5.56	2.75×10^{-15}
XP_034323071.1	Spidroin-1-like, 3	*Crassostrea gigas*	3×10^{-08}	5.51	2.83×10^{-15}
CAG2230452.1	CD283	*Mytilus edulis*	3×10^{-59}	–5.49	5.61×10^{-15}
XP_011415271.2	Acanthoscurrin-2	*Crassostrea gigas*	0	–5.40	7.66×10^{-15}
XP_011452536.2	Leucine-rich repeat LGI family member 2 (LGI2)	*Crassostrea gigas*	0	–5.38	2.24×10^{-14}
XP_021357253.1	Serine/arginine repetitive matrix protein 1-like (SRRM1-like)	*Mizuhopecten yessoensis*	9×10^{-62}	5.44	2.88×10^{-14}

a: protein name according to NCBI database; b: Blastx Expect value; c: log fold change of transcripts; d: false discovery rate, FDR

IV.3.2.1 Pre-Settlement Pediveliger Larvae

Some transcripts found in the top thirty list were upregulated in the pediveliger larvae but were found downregulated after settlement code for a variety of functions: immunity and stress response (ADAM-1, CD283, Acanthoscurrin-2, 2 IgGFc-binding genes, LRRC4C); nervous system remodeling (1 SCO-spondin, 3 SCO-spondin-like genes); structural protein/cell adhesion (Coadhesin, Stereocilin); signal regulation (CPK7, LGI2), and an uncharacterized LOC105322623 gene with unknown function. Another interesting set of differentially expressed transcripts that were downregulated after settlement, which were not within the top thirty list, code for the regulation of ciliary motility (dynein regulatory complex subunit 5 (LOC105344378)), ion transport (transient receptor potential cation channel subfamily M member-like 2 (LOC105348393)), as well as cell signaling related receptors: octopamine receptor b-2R (LC105343310), leucine-rich repeat-containing G-protein coupled receptor 4 (LOC105325936), nuclear receptor subfamily 2 stage E member 1 (LOC105336095), neuronal acetylcholine receptors (LOC105331790, LOC105117683432, LOC105335119), FMRFamide receptor (LOC105331946), and ecdysone receptor (LOC105320474).

IV.3.2.2 Postlarvae

Among those in the top thirty list, some transcripts which were upregulated after settlement were found to hold a variety of functional roles: protein synthesis and degradation (2 trfA-like transcription factor genes), structural proteins (3 Spidroin-1-like and 2 GRP1.8 variant genes), immune and stress response (AESIP), ion transporter (ADP-ribosylation factor 1-like), oxidoreductase activity (Tyrp-1), cytoskeleton/cell adhesion (Collagen alpha-2(I) chain), and two uncharacterized LOC117688748 and LOC117688749 transcripts with unknown function. Notably, other upregulated transcripts after settlement that were not shown in this list include transcription factors (ATP-dependent RNA helicase *elf* 4A (LOC105324495), ETS-related transcription factor Elf3-like (LOC105325143), activating

transcription factor 3 (LOC105337015), cyclic AMP-dependent transcription factor ATF-3 (LOC105337016)), an enzyme for hormone synthesis and breakdown (cytochrome P450 (LOC105325143)), and structural proteins (highly expressed transcripts: gigasin-6 isoforms X1 and X2 (LOC105343850), dentin sialophosphoprotein-like (LOC105323395), and clumping factor A (LOC105346538)). Interestingly, gigasin-6 isoforms X1 and X2, as well as dentin sialophosphoprotein-like, were upregulated in the conspecific cue-induced PL batch. Gigasin-6 and its isoforms have been reported as a component in the adult shell matrix in *C. gigas* [60,61], while the BLAST sequence homology search on dentin sialophosphoprotein-like showed close homology to folian-cv1 in *C. virginica* [62]. Others were related to signal regulation processes such as calmodulin A (LOC105327998), Wnt inhibitory factor 1/EGF-like domain-containing protein 2 (Wif-1) (LOC105335286), tyrosine-protein kinase receptor Tie (LOC105338560), beta-2 adrenergic receptor (LOC105336903), nuclear receptors (LOC105338842, LOC105345246), and FMFRamide receptor (LOC105343238). The differential expression of beta-2 adrenergic receptor in the conspecific cue-induced PL batch is of particular interest. This is in contrast with the findings of Coon et al. [63] wherein the endogenous alpha-1 adrenergic receptor was implicated in *C. gigas* larval settlement induction processes through pharmacological evidence. In addition, as observed by Wang et al. [64], the larval metamorphosis in *Meretrix meretrix* competent clam was successfully induced or inhibited by a beta-adrenoreceptor agonist and antagonist, respectively. The possible role of beta-2 adrenergic receptors in *C. gigas* larval settlement warrants further investigation. It is also noteworthy that in this study, some differentially expressed transcripts that manifested some form of isoform switching [57] were identified, wherein the FMFRamide receptor (LOC105331946) was downregulated, while the FMFRamide receptor (LOC105343238) was upregulated after settlement. An isoform of Calmodulin A (LOC105328000) was downregulated, but the precursor form of Calmodulin A (LOC105327998) was upregulated after settlement (RNA-seq) and was supported by qRT-PCR. This may also imply that some genes play unique functions during oyster larval settlement, similar to those reported in other invertebrates [65,66].

IV.3.3 Gene Ontology (GO) Enrichment Analysis

To understand how the presence of a conspecific cue, CgSE, impacts larval settlement responses and to interpret the differential expression (DE) results in terms of higher-order biological processes or molecular pathways, we counted the number of DE transcripts that were annotated with each possible GO term [67]. GO terms that occurred frequently in the list of DE

genes are said to be over-represented or enriched [67]. The filtering threshold values in this analysis were set at FDR < 0.01 and |log2 Ratio| > 1.5. Table 4.4 presents the results of the GO enrichment analysis while Appendix Table A4.3 displays the full annotation on the sequence homology search for the uncharacterized transcripts associated with each of the GO terms.

Table 4.4. Gene Ontology analysis of transcripts with differential expression in Pedi and PL batches before and after conspecific cue settlement induction. Underlined gene transcripts are those found in two GO categories.

Description	Chitin-binding	Calcium ion binding	Extracellular region
GO ID	GO:0008061	GO:0005509	GO:0005576
FDR value	8.13×10^{-05}	0.0024	5.67×10^{-06}
Ontology	Molecular function (MF)	Molecular function (MF)	Cellular component (CC)
Count	9	17	17
Gene	cgPif97 (↑), Peritrophin-44 (↓), Lactahedrin (↓), Collagen alpha 1 (XII) chain (↓), Peritrophin-1-like,1 (↓), Peritrophin-like,2 (↓), Cleavage and polyadenylation factor 1 subunit 1 (Clip1) (↑), Putative chitinase 1 precursor (↑), Putative chitinase (↑)	DNA ligase 1 (↓), EF-hand domain-containing protein 1 (↓), Calmodulin-4 (↓), Calcium-binding protein E-63-1-like (↓), Calmodulin-A isoforms (↓), Leucine-rich repeat-containing protein 74B (↓), Sarcoplasmic calcium-binding protein (↑), Myosin (↑), EGF-like repeat and discoidin-I domain-containing protein 3 (↓), Mammalian ependymin-related protein 1 (↓), EF-hand domain-containing family member B (↓), Sarcoplasmic calcium-binding protein (↑), Neurocalcin homolog,1,(↓), Neurocalcin homolog,2 (↑), Regucalcin,1 (↓), Regucalcin,2 (↓), Annexin A4 (↓)	Poly(3-hydroxyalkanoate) depolymerase C (↓), cgPif97 (↑), Trithorax group protein osa-like (↑), NPC intracellular cholesterol transporter 2 (↓), Mammalian ependymin-related protein 1 (↓), Peritrophin-44 (↓), Lactadherin (↓), Golgi-associated plant pathogenesis-related protein 1 (↓), Collagen alpha 1 (XII) chain (↓), Dermatopontin (↓), Metalloproteinase inhibitor 3 (↓), Peritrophin-1-like, 1 (↓), Peritrophin-1-like, 2 (↓), inactive pancreatic lipase-related protein 1 (↓), pancreatic lipase-related protein 2 (↓), adenosine deaminase AGSA (↓), von Willebrand factor C domain-containing protein 2-like (VWC2L) (↓)

A total of 43 transcripts had a GO functional annotation associated with them. They were assigned into two major functional categories: molecular functions and cellular components. Overall, 26 transcripts were classified into molecular function processes, that is, chitin-binding (9) and calcium ion binding (17), while 17 transcripts were assigned to the extracellular region,

under the cellular component processes. Seven transcripts were also found in two categories in both molecular function and cellular components. These results reflect the difficulty in associating a possible function and an appropriate match [2] with existing GO term annotations. One possible explanation for why it was difficult to parse the differentially expressed transcripts into discrete GO terms could be explained by the findings of Zhang et al. [2], wherein they found that the *C. gigas* genome exhibited a high degree of polymorphism and an abundance of repetitive sequences, of which over 62% could not be assigned to known categories. This reflects not only the presence of numerous isoforms of many genes but also the paucity of genomic information from this species [2]. Our present findings are also in agreement with a previous study by Huan et al. [5], where they examined the proteomic profiles between the pediveliger larvae and biofilm-induced postlarvae in *C. gigas*. They found no significant result in the GO analysis that could associate the DE proteins with particular biological processes owing to their functional diversity and the relatively small number of proteins [5]. They resorted to manually dividing the proteins into different groups according to their potential functions based on sequence homology [5], and which we also applied in this present study for those uncharacterized transcripts so as to know their possible function.

Under the chitin-binding pathway (GO:0008061; Q = 8.13 x 10^{-05}), we found four significantly enriched transcripts were upregulated while five transcripts were downregulated after settlement. The potential functions of these upregulated transcripts were as structural proteins (cgPif97) and enzymes (Clip1, 2 Putative chitinases), while those of the downregulated transcripts indicate roles in the cytoskeleton and cell adhesion (peritrophin-44, 2 lactadherin, collagen alpha 1 (XII) chain, peritrophin-1-like).

Three upregulated and fourteen downregulated transcripts were significantly enriched in the calcium ion binding pathway (GO:0005509; Q = 0.0024) after settlement. The associated transcripts under this category highlight upregulated transcripts with functional roles such as those related to the cytoskeleton and cell adhesion (sarcoplasmic calcium-binding protein), ion transport (neurocalcin homolog), and tissue development (myosin), while the rest of the downregulated transcripts may have potential functions involved in protein synthesis and degradation (DNA ligase 1), signal regulation (EF-hand domain-containing protein 1, calmodulin-4, calcium-binding protein E63-1-like, calmodulin-A, EGF-like repeat and discoidin I-like domain-containing protein 3, EF-hand domain-containing family member B, Annexin A4), immunity and stress response (LRR74), tissue development (mammalian ependymin-related protein), and ion transport (neurocalcin homolog, regucalcin).

Lastly, two upregulated transcripts and fifteen downregulated transcripts after settlement were significantly enriched under the cellular component processes associated with the extracellular region (G0:0005576; Q = 5.67 x 10^{-06}). The upregulated transcripts associated with this category include shell matrix protein (cgPif) and another transcript related to protein synthesis (Trithorax stage protein osa-like). However, the fifteen downregulated transcripts code for a variety of functions such as enzymes (poly(3-hydroxyalkanoate) depolymerase C, inactive pancreatic lipase-related protein 1, pancreatic lipase-related protein 2, adenosine deaminase AGSA), ion transport (NPC intracellular cholesterol transporter 2), particular tissue development (mammalian ependymin-related protein 1), cytoskeleton and cell adhesion (peritrophin-44, lactadherin, collagen alpha 1 (XII) chain, dermatopontin, 2 peritrophin-1-like genes), a protease inhibitor (metalloproteinase inhibitor 3), and signal regulation (Golgi-associated plant pathogenesis-related protein 1, von Willebrand factor C domain-containing protein 2-like).

IV.3.4 Validation of Transcriptomic Data via qRT-PCR

The qRT-PCR validation experiments were performed on six selected transcripts that were differentially expressed before and after settlement, as well as those found in the GO-enriched categories in the transcriptome dataset. Figure 4.3 illustrates their relative mRNA expression levels. Three upregulated transcripts before settlement but downregulated after settlement such as VWC2L (GO-extracellular region), collagen alpha 1 (XII) (GO-extracellular region), and annexin A4 (GO-calcium ion binding) were consistent in the qRT-PCR results, showing higher relative mRNA expression levels before settlement but lower expression levels after settlement. Likewise, three downregulated transcripts before settlement but upregulated after settlement such as putative chitinase 1 (GO-Chitin binding), calmodulin A, and Wnt inhibitory factor 1 were supported by qRT-PCR, showing lower relative mRNA expression levels before settlement but with higher expression after settlement.

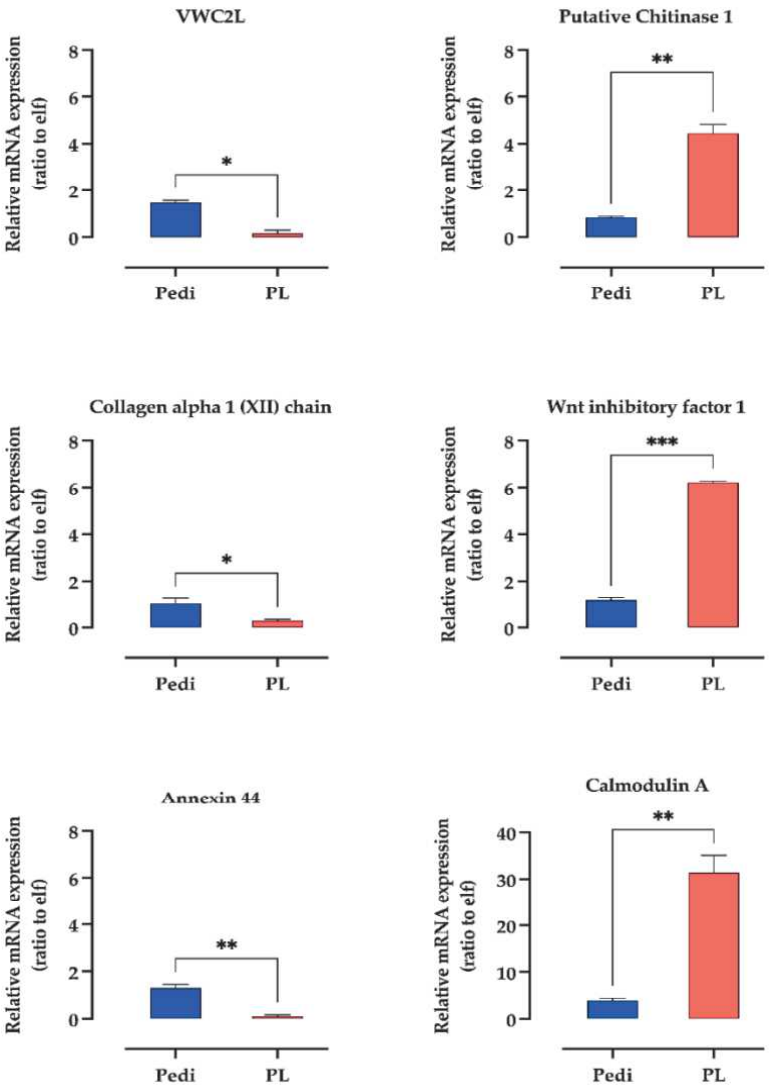

Figure 4.3. Relative expression levels of selected genes before (pediveliger, Pedi) and after (postlarvae, PL) settlement induction by a conspecific cue from shells in *C. gigas* using quantitative real-time PCR. Data are expressed as the mean + SD (n = 3). Values with "*, **, ***" are significantly different ($p < 0.05$, <0.01, <0.001, respectively). Abbreviation: VWC2L = von Willebrand factor C domain-containing protein 2-like.

IV.4 Discussion

Oysters have a complex life cycle in which they exhibit an active, pelagic phase and an attached, benthic phase. The transition from one phase to the next is one of the most life-changing events common to many mollusks, such as oysters. Physiological mechanisms associated with the induction of molluscan larval settlement, including oysters, are hypothesized to involve the nervous system as a result of stimulation of different neural circuits by external cues [48,68–73]. Previous studies have shown that irreversible larval attachment to a substrate and metamorphosis, which comprise the entire settlement process in oyster larvae, can be induced and experimentally separated using pharmacological agents [48]. Coon et al. [3] demonstrated that L-3,4-dihydroxyphenylalanine (L-DOPA) induced settlement behavior that led to the cementation of the *C. gigas* larvae to a substrate followed by metamorphosis, while the use of epinephrine led to the induction of metamorphosis bypassing the attachment of the larvae on a substrate [74] by interacting with endogenous α1-adrenergic receptors [63]. In this present study, we provide new insights into the possible involvement of a hormonal signaling system associated with the onset of metamorphic processes upon settlement induction by a natural settlement inducer from shells of conspecifics. Our understanding of the neuroendocrine-biomineralization gene activities that accompany oyster larval settlement could provide insights into pathways and mechanisms and help to understand how the larvae respond to external cues to initiate these processes [8,75]. To the best of my knowledge, this is the first time that omics technologies as a tool, such as transcriptomics, have been applied to understand the response of oyster larvae to a natural settlement-inducing cue from the shells of conspecifics. I hope that this data provides a new direction to studies aimed at understanding the gregarious phenomenon among oysters.

This present study aims to provide transcriptomics insights on differential expression patterns that happen within an oyster larva as it encounters the presence of a conspecific cue from shell protein extracts and is induced to settle. Hence, transcriptome profile comparisons were analyzed between the larval batches before and after conspecific cue exposure which led to their transition from being a pediveliger larva into a postlarva. The discussion here will elaborate inferences on a snapshot of one group of pediveliger larvae that settled first. The biological relevance of this fraction of the total population is worth considering and introducing as future research topics.

In this study, a total of 2,383 transcripts were found to be differentially expressed between the pre-settlement pediveliger larvae (Pedi) and postlarvae (PL). There were 740

upregulated and 1,643 downregulated transcripts after settlement. These identified transcripts may play important roles during the conspecific cue-mediated induction of *C. gigas* larval settlement. It would also be essential to elucidate their individual functions for future studies.

A graphical representation of the key molecular events in conspecific cue-mediated *C. gigas* larval settlement induction is shown in Figure 4.4. Comparison of the transcriptomes before and after exposure of the competent larvae to a conspecific cue represents a wide range of functions and indicates the complex regulatory mechanisms during the *C. gigas* larval settlement processes. Before the onset of metamorphosis, several transcripts related to the cytoskeleton, cell adhesion, and ECM were differentially expressed and may implicate muscular tissue development and structural reorganization of body patterns [76]. Furthermore, transcripts associated with the cytoskeleton may facilitate cell and tissue assembly and the movement and disassembly of muscle filaments during the metamorphosis [76].

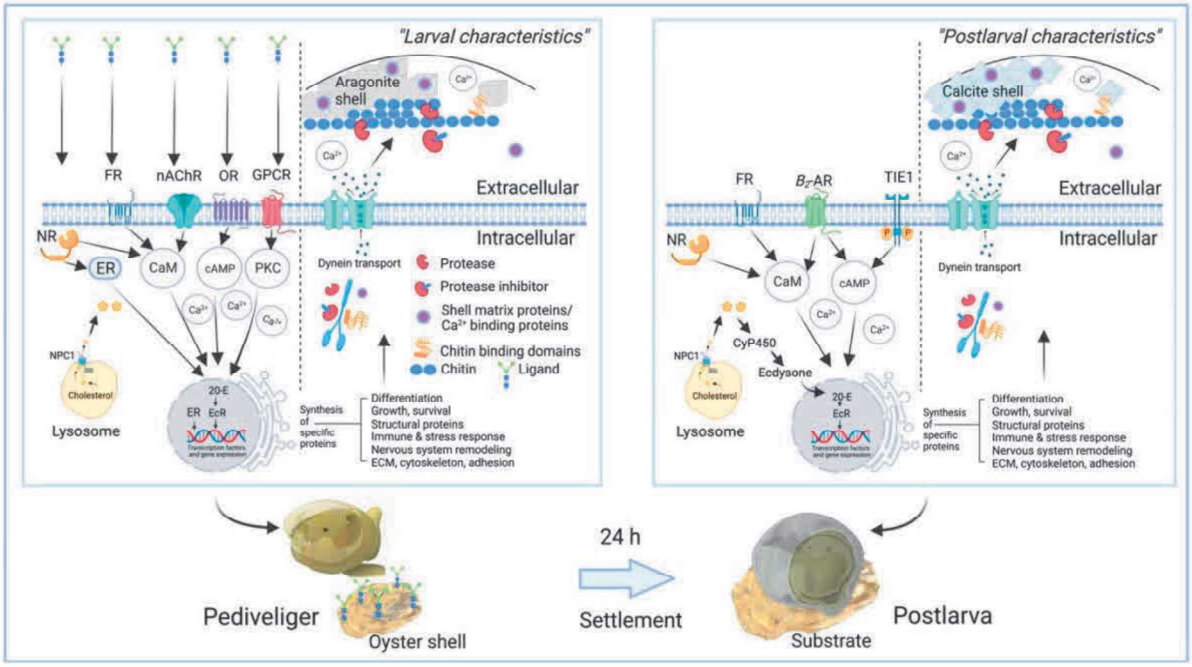

Figure 4.4. Schematic of the neuroendocrine-biomineralization crosstalk in *C. gigas* conspecific cue-induced larval settlement. This indicates the biological and molecular processes identified through the differential transcript expression in the competent pediveliger as it transitioned into postlarvae by conspecific cue-mediated settlement induction. As the pediveliger commences its searching behavior, it makes use of its external chemoreceptors to detect signals for a suitable location for settlement. Once it detects a ligand and other signaling cues coming from shells of conspecifics, various intracellular signaling receptors act in a concerted way and a molecular switch is turned on for the initiation of downstream signal cascades into the nucleus with the possible involvement of the ecdysone signaling pathway. Then, various transcription factors and gene expressions for the synthesis of specific proteins are activated. Subsequently, rapid morphological, physiological, structural, and functional changes occur as the competent pediveliger undergoes its metamorphosis into a postlarva within 24 h. Abbreviations: NR, nuclear receptor; FR, FRMF (Phe-Met-Arg-Phe) amide receptor; nAChR, neuronal acetylcholine receptor; OR, octopamine receptor; GPCR, leucine-

repeat-containing GPCR 4; ER, estrogen receptor; CaM, calmodulin groups; cAMP, cAMP-regulated D2 protein; PKC, protein kinase C; Ca^{2+}, calcium ion; NPC1, NPC cholesterol transporter 1; 20-E, hydroxyecdysone; EcR, ecdysone receptor; B_2AR, β_2 adrenoreceptor; TIE 1, tyrosine-protein kinase receptor TIE; CyP450, cytochrome P450; and P, phosphate. Some concept models were adapted and modified to illustrate processes of shell formation [77], chitin-mediated biomineralization [78], and dynein transport [79] that reflect the transcriptome larval settlement dynamics in our study. This figure was created using BioRender.com.

On the other hand, immune- and stress-related genes may play a role in the regulation of apoptosis, protein degradation, cell proliferation, and cell processes [76,80,81], and may be required to counter the damage caused by oxidative stress [76]. Genes related to nervous system remodeling may allow the oyster larvae greater sensitivity to external cues as it reaches a metamorphic competent stage. Thus, taking all these data together implies that the competent pediveliger larvae become equipped to transduce external cues from the environment into internal processes with neural and/or hormonal elements as it tries to find a suitable substrate for settlement [8,11,75,82]. Once the larvae have accumulated a significant amount of cue from a suitable settlement site, the onset of settlement is triggered. In *C. gigas*, just like other marine invertebrates, this occurs in a rapid process. In our case, within 24 h, the induced larvae acquired postlarval characteristics such as permanent attachment to a substrate and new shell growth, to name a few. The fast transition from pediveliger to postlarvae is necessary as these organisms are particularly vulnerable to predation during this event [83]. Hadfield et al. [84] proposed that marine invertebrates appear to require little or no *de novo* gene action during the metamorphic induction process which makes them capable of responding rapidly to environmental cues to settlement and undergoing very rapid metamorphic morphogenesis [84]. They then put in contrast, the slow, hormonally regulated metamorphic transitions of vertebrates and insects. Our present data show, however, that several elements related to various hormonal signaling pathways were also differentially expressed in the pediveliger larvae. This may indicate their involvement in oyster metamorphosis. Recent molecular studies have shown that Lophotrochozoans, such as *C. gigas*, may contain integrated hormone systems [33,85] such as ecdysone [8,33,85] and thyroid hormone [86] (for the hormonal system); octopamine [34] and epinephrine [63] (for the nervous system) were reported in the development of *C. gigas*. We then hypothesized that the differential expression of some elements related to various hormonal signaling pathways in *C. gigas* may be necessary for preparation for the onset of metamorphosis. This variety of hormones may act in a concerted manner when activated by their respective receptors. They can then act as transcription factors that unleash cascades of *de novo* transcription and synthesis of new proteins, which may be similar to the mode of action reported in insects and amphibians [84,87]. Simultaneous to the

ongoing loss of larva-specific structures at the onset of metamorphosis, postlarval structures are also being formed [84,87]. This present data also shows that several differential transcript expressions were found related to transcription factors, protein synthesis, structural proteins including shell matrix proteins often found in adult shells, biomineralization-related enzymes, ion transport, chitin, and calcium ion binding pathways in the postlarval stage. After 24 h of exposure of the oyster larvae to the settlement-inducing conspecific cue, shell matrix proteins such as cgPif97 indicate the start of the restructuring of the shell composition from aragonite (dominant in the pediveliger stage) to calcite (dominant in postlarvae to adults) [2]. As the postlarvae develop to the adult stage, this calcified shell becomes thicker, and sturdy enough to protect it from harsh and dynamically changing environments, predation, and desiccation [2]. Moreover, this transcriptomics analysis also shows several similar genes to those reported in epinephrine-induced *C. gigas* larvae, especially on the possible involvement of ecdysteroid signaling pathways [33]. Furthermore, in this present study, several elements involved in the ecdysone signal pathway, including the ecdysone receptor, NPC cholesterol transporter, cytochrome P450, tyrosine-protein kinase receptor Tie, tyrosine (Tryp-1), and other nuclear receptors, were differentially expressed. Likewise, receptors of neuronal acetylcholine, octopamine b-2R like, estrogen, FMRFamide, and GPCRs were found to be differentially expressed. A similar observation was made by Zheng et al. [30], having compared *P. fucata* and *C. gigas* proteomes. Broadly, these results provide support to the hypothesis proposed by Zheng et al. [30], that a gene regulatory network may have closely linked dynamics with shell formation and may involve the interplay of various hormone receptors, neurotransmitters, and neuropeptide receptors working together in a concerted way during larval settlement-induction processes. Hence, conspecific cue-mediated larval settlement induction may involve a neuroendocrine-biomineralization crosstalk during the settlement processes in *C. gigas*. In comparison to these present findings, the presence of elements related to the ecdysone signal pathway in *C. gigas*, was further evidenced by the study of Xu and Zhang [33] on the transcriptomic and proteomic dynamics profile of larvae exposed to epinephrine stimulation. Likewise, a recent study by Xiong et al. [44], demonstrated that the ecdysone signal pathway participates in the shell formation of the Pearl oyster *P. fucata martensii*. They have reported that ecdysone could activate the ecdysone receptor complex to regulate the expression of biomineralization-related genes. Hence, on the basis of the current findings and previous reports, the possibility of a neuroendocrine-biomineralization crosstalk in *C. gigas* conspecific cue-induced larval settlement was proposed.

These present findings demonstrate the mode of action of a naturally occurring inducer

such as the chemical cue from shells of conspecifics, which complements and show some common differential transcript expression that was reported in neuroactive compound-induced and biofilm-induced *C. gigas* larvae and in other marine invertebrate larvae, which may have potential roles in larval settlement. However, these results are based on the comparison of two pooled populations in two RNA-seq samples representing Pedi and PL batches. This finding may warrant further study using multiple samples. Nonetheless, these present findings showed that expression trends of selected representatives from the most significantly expressed genes were consistent between the real time-PCR and transcriptome data. Hence, the lack of replicates did not affect the results and the model of edgeR could compensate for it. On the other hand, conditions for RNA-seq larval sample preparation may also impact gene expression trends, and this warrants future investigation [79].

This data can be a useful baseline in further identifying the natural settlement-inducing compound from shells of conspecifics and elucidating their genetic mechanisms. In addition, a more in-depth examination of the various pathways is essential by cloning their essential components in such a way that the receptors may be utilized as targets to define pathways. Knowledge of receptor characterization and the downstream signaling cascade for each pathway might lead to significant advances in our knowledge of larval settlement as well as the neuroendocrine functions in oysters. Further identification and characterization of this settlement-inducing cue from shells in conspecifics can enable us to utilize its properties as a possible attractant for wild and hatchery-grown oyster larvae to attach to artificial substrates for aquaculture. Development in our knowledge and techniques in this field may be useful in increasing hatchery production of potential mollusk species, developing new antifouling compounds capable of deliberately inhibiting bivalve attachment, and helping other related industries.

IV.5 Conclusions

A transcriptome profile of an oyster larval response to a conspecific cue-mediated settlement induction in *Crassostrea gigas* is reported for the first time. Transcriptomics analysis has identified key transcripts that were differentially expressed in the pediveliger and conspecific cue-induced postlarvae in *C. gigas*. Quantitative analysis revealed 2,383 candidate transcripts that may play important roles during larval settlement processes. Several transcripts with differential expression were significantly enriched in GO analysis associated with chitin-binding, calcium-ion binding, and extracellular region-related categories. Differential expressions of six candidate transcripts representing Pedi and PL batches were validated by qRT-PCR, which showed relative expression trends were consistent between qRT-PCR and transcriptome data. By comparing the expression patterns of the Pedi and conspecific cue-induced PL batches, we hypothesized that shell formation shows closely linked dynamics with a gene regulatory network in *C. gigas* that may involve the interplay of various hormone receptors, neurotransmitters, and neuropeptide receptors working together in a concerted way in the induction of conspecific cue-mediated larval settlement. These results highlight the transcriptome dynamics underlying the gregarious settlement of oysters on conspecific adult shells and show the potential use of this settlement-inducing cue as an attractant for wild and hatchery-grown oyster larvae to attach to artificial substrates for aquaculture. Furthermore, this study hopes to provide hints and support for future studies related to the involvement of the ecdysone signal pathway and the linkages within the neuroendocrine-biomineralization crosstalk in *C. gigas* settlement.

IV.6 References

1. Hedgecock, D.; Gaffney, P.M.; Goulletquer, P.; Guo, X.; Reece, K.; Warr, G.W. The Case for Sequencing the Pacific Oyster Genome. *J. Shellfish Res.* 2005, *24*, 429–441. https://doi.org/10.2983/0730-8000(2005)24[429:TCFSTP]2.0.CO;2.
2. Wang, J.; Zhang, G.; Fang, X.; Guo, X.; Li, L.; Luo, R.; Xu, F.; Yang, P.; Zhang, L.; Wang, X.; et al. The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation. *Nature* 2012, *490*, 49–54. https://doi.org/10.1038/nature11413.
3. Coon, S.L.; Bonar, D.B.; Weiner, R.M. Induction of Settlement and Metamorphosis of the Pacific Oyster, *Crassostrea gigas* (Thunberg), by L-DOPA and Catecholamines. *J. Exp. Mar. Bio. Ecol.* 1985, *94*, 211–221. https://doi.org/10.1016/0022-0981(85)90059-0.
4. Bonar, D.B.; Coon, S.L.; Walch, M.; Weiner, R.M.; Fitt, W. Control of Oyster Settlement and Metamorphosis by Endogenous and Exogenous Chemical Cues. *Bull. Mar. Sci.* 1990, *46*, 484–498.
5. Huan, P.; Wang, H.; Liu, B. A Label-Free Proteomic Analysis on Competent Larvae and Juveniles of the Pacific Oyster (*Crassostrea gigas*). *PLoS ONE* 2015, *10*, e0135008. https://doi.org/10.1371/journal.pone.0135008.
6. Laing, I. Bivalve Molluscs: Biology, Ecology and Culture. *Aquaculture* 2004, *229*, 507–508. https://doi.org/10.1016/s0044-8486(03)00412-5.
7. Gosling, E. Reproduction, Settlement and Recruitment. In *Marine Bivalve Molluscs*; John Wiley & Sons, Ltd.: Chichester, UK, 2015; pp. 157–202.
8. Joyce, A.; Vogeler, S. Molluscan Bivalve Settlement and Metamorphosis: Neuroendocrine Inducers and Morphogenetic Responses. *Aquaculture* 2018, *487*, 64–82. https://doi.org/10.1016/j.aquaculture.2018.01.002.
9. Heyland, A.; Moroz, L.L. Signaling Mechanisms Underlying Metamorphic Transitions in Animals. *Integr. Comp. Biol.* 2006, *46*, 743–759. https://doi.org/10.1093/icb/icl023.
10. Fitt, W.K.; Labare, M.P.; Fuqua, W.C.; Walch, M.; Coon, S.L.; Bonar, D.B.; Colwell, R.R.; Weiner, R.M. Factors Influencing Bacterial Production of Inducers of Settlement Behavior of Larvae of the Oyster *Crassostrea gigas*. *Microb. Ecol.* 1989, *17*, 287–298. https://doi.org/10.1007/BF02012841.
11. Fitt, W.K.; Coon, S.L.; Walch, M.; Weiner, R.M.; Colwell, R.R.; Bonar, D.B. Settlement Behavior and Metamorphosis of Oyster Larvae (*Crassostrea gigas*) in Response to Bacterial Supernatants. *Mar. Biol.* 1990, *106*, 389–394. https://doi.org/10.1007/BF01344317.

12. Tamburri, M.N.; Zimmer-Faust, R.K.; Tamplin, M.L. Natural Sources and Properties of Chemical Inducers Mediating Settlement of Oyster Larvae: A Re-Examination. *Biol. Bull.* 1992, *183*, 327–338. https://doi.org/10.2307/1542218.
13. Hashimoto, K.; Vasquez, H.E.; Kitamura, H.; Satuito, C.G. Larval Settlement of the Pacific Oyster *Crassostrea gigas* in Response to Marine Microbial Films and Monospecies Bacterial Films. *Sess. Org.* 2017, *34*, 21–27. https://doi.org/10.4282/sosj.34.21.
14. Swanson, R.L.; De Nys, R.; Huggett, M.J.; Green, J.K.; Steinberg, P.D. In Situ Quantification of a Natural Settlement Cue and Recruitment of the Australian Sea Urchin *Holopneustes purpurascens*. *Mar. Ecol. Prog. Ser.* 2006, *314*, 1–14. https://doi.org/10.3354/meps314001.
15. Huggett, M.J.; Williamson, J.E.; De Nys, R.; Kjelleberg, S.; Steinberg, P.D. Larval Settlement of the Common Australian Sea Urchin *Heliocidaris erythrogramma* in Response to Bacteria from the Surface of Coralline Algae. *Oecologia* 2006, *149*, 604–619. https://doi.org/10.1007/s00442-006-0470-8.
16. Huggett, M.J.; Crocetti, G.R.; Kjelleberg, S.; Steinberg, P.D. Recruitment of the Sea Urchin *Heliocidaris erythrogramma* and the Distribution and Abundance of Inducing Bacteria in the Field. *Aquat. Microb. Ecol.* 2008, *53*, 161–171. https://doi.org/10.3354/ame01239.
17. Hadfield, M.; Paul, V. Natural Chemical Cues for Settlement and Metamorphosis of Marine-Invertebrate Larvae. In *Marine Chemical Ecology*; McClintock, J.B., Baker, B.J., Eds., CRC Press: Boca Raton, FL, USA, 2001; pp. 431–461; ISBN 9781420036602.
18. Hay, M.E. Marine Chemical Ecology: Chemical Signals and Cues Structure Marine Populations, Communities, and Ecosystems. *Ann. Rev. Mar. Sci.* 2009, *1*, 193–212. https://doi.org/10.1146/annurev.marine.010908.163708.
19. Bayne, B.L. The Gregarious Behaviour of the Larvae of *Ostrea edulis* L. at Settlement. *J. Mar. Biol. Assoc. U. K.* 1969, *49*, 327–356. https://doi.org/10.1017/S0025315400035943.
20. Crisp, D.J. Chemical Factors Inducing Settlement in *Crassostrea virginica* (Gmelin). *J. Anim. Ecol.* 1967, *36*, 329–335. https://doi.org/10.2307/2916.
21. Hidu, H. Gregarious Setting in the American Oyster *Crassostrea virginica* Gmelin. *Chesap. Sci.* 1969, *10*, 85–92. https://doi.org/10.2307/1350836.
22. Hirata, Y.; Tamura, Y.; Nagasawa, K. Influence of Presoaking Conditions of Spat Collectors in Seawater Containing Adult Pacific Oysters (*Crassostrea gigas*) on Their Larval Settlement. *Nippon Suisan Gakkaishi (Jpn.)* 2008, *74*, 1017–1023. https://doi.org/10.2331/suisan.74.1017.

23. Pascual, M.S.; Zampatti, E.A. Evidence of a Chemically Mediated Adult-Larval Interaction Triggering Settlement in *Ostrea puelchana*: Applications in Hatchery Production. *Aquaculture* 1995, *133*, 33–44. https://doi.org/10.1016/0044-8486(95)00005-M.
24. Dworjanyn, S.A.; Pirozzi, I. Induction of Settlement in the Sea Urchin *Tripneustes gratilla* by Macroalgae, Biofilms and Conspecifics: A Role for Bacteria? *Aquaculture* 2008, *274*, 268–274. https://doi.org/10.1016/j.aquaculture.2007.11.030.
25. Vasquez, H.E.; Hashimoto, K.; Yoshida, A.; Hara, K.; Imai, C.C.; Kitamura, H.; Satuito, C.G. A Glycoprotein in Shells of Conspecifics Induces Larval Settlement of the Pacific Oyster *Crassostrea gigas*. *PLoS ONE* 2013, *8*, e82358. https://doi.org/10.1371/journal.pone.0082358.
26. Vasquez, H.E.; Hashimoto, K.; Kitamura, H.; Satuito, C.G. Wheat Germ Agglutinin-Binding Glycoprotein Extract from Shells of Conspecifics Induces Settlement of Larvae of the Pacific Oyster *Crassostrea gigas* (Thunberg). *J. Shellfish Res.* 2014, *33*, 415–423. https://doi.org/10.2983/035.033.0210.
27. Sedanza, M.G.; Kim, H.-J.; Seposo, X.; Yoshida, A.; Yamaguchi, K.; Satuito, C.G. Regulatory Role of Sugars on the Settlement Inducing Activity of a Conspecific Cue in Pacific Oyster *Crassostrea gigas*. *Int. J. Mol. Sci.* 2021, *22*, 3273. https://doi.org/10.3390/ijms22063273.
28. Poirier, L.A.; Clements, J.C.; Davidson, J.D.P.; Miron, G.; Davidson, J.; Comeau, L.A. Sink before You Settle: Settlement Behaviour of Eastern Oyster (*Crassostrea virginica*) Larvae on Artificial Spat Collectors and Natural Substrate. *Aquac. Rep.* 2019, *13*, 100181. https://doi.org/10.1016/j.aqrep.2019.100181.
29. Tamburri, M.N.; Luckenbach, M.W.; Breitburg, D.L.; Bonniwell, S.M. Settlement of *Crassostrea ariakensis* Larvae: Effects of Substrate, Biofilms, Sediment and Adult Chemical Cues. *J. Shellfish Res.* 2008, *27*, 601–608. https://doi.org/10.2983/0730-8000(2008)27[601:SOCALE]2.0.CO;2.
30. Zheng, Z.; Hao, R.; Xiong, X.; Jiao, Y.; Deng, Y.; Du, X. Developmental Characteristics of Pearl Oyster *Pinctada fucata martensii*: Insight into Key Molecular Events Related to Shell Formation, Settlement and Metamorphosis. *BMC Genom.* 2019, *20*, 122. https://doi.org/10.1186/s12864-019-5505-8.
31. Di, G.; Xiao, X.; Tong, M.H.; Chen, X.; Li, L.; Huang, M.; Zhou, L.; Ke, C. Proteome of Larval Metamorphosis Induced by Epinephrine in the Fujian Oyster *Crassostrea angulata*. *BMC Genom.* 2020, *21*, 675. https://doi.org/10.1186/s12864-020-07066-z.

32. Li, H.; Zhang, B.; Huang, G.; Liu, B.; Fan, S.; Zhang, D.; Yu, D. Differential Gene Expression during Larval Metamorphic Development in the Pearl Oyster, *Pinctada fucata*, Based on Transcriptome Analysis. *Int. J. Genom.* 2016, *2016*, 1–15. https://doi.org/10.1155/2016/2895303.
33. Xu, F.; Zhang, G. Transcriptomic and Proteomic Dynamics during Metamorphosis of Pacific Oyster *Crassostrea gigas*. *bioRxiv* 2020, bioRxiv:2020.03.25.004614. https://doi.org/10.1101/2020.03.25.004614.
34. Ji, P.; Xu, F.; Huang, B.; Li, Y.; Li, L.; Zhang, G. Molecular Characterization and Functional Analysis of a Putative Octopamine/Tyramine Receptor during the Developmental Stages of the Pacific Oyster, *Crassostrea gigas*. *PLoS ONE* 2016, *11*, e0168574. https://doi.org/10.1371/journal.pone.0168574.
35. Foulon, V.; Boudry, P.; Artigaud, S.; Guérard, F.; Hellio, C. In Silico Analysis of Pacific Oyster (*Crassostrea gigas*) Transcriptome over Developmental Stages Reveals Candidate Genes for Larval Settlement. *Int. J. Mol. Sci.* 2019, *20*, 197. https://doi.org/10.3390/ijms20010197.
36. Vogeler, S.; Bean, T.P.; Lyons, B.P.; Galloway, T.S. Dynamics of Nuclear Receptor Gene Expression during Pacific Oyster Development. *BMC Dev. Biol.* 2016, *16*, 33. https://doi.org/10.1186/s12861-016-0129-6.
37. Beiras, R.; Widdows, J. Induction of Metamorphosis in Larvae of the Oyster *Crassostrea gigas* Using Neuroactive Compounds. *Mar. Biol.* 1995, *123*, 327–334. https://doi.org/10.1007/BF00353624.
38. Murthy, P.S.; Venugopalan, V.P.; Nair, K.V.K.; Subramoniam, T. Larval Settlement and Surfaces: Implications in Development of Antifouling Strategies. In *Marine and Industrial Biofouling*; Springer: Berlin/Heidelberg, Germany, 2008; pp. 233–263. https://doi.org/10.1007/978-3-540-69796-1_13.
39. Fiedler, T.J.; Hudder, A.; McKay, S.J.; Shivkumar, S.; Capo, T.R.; Schmale, M.C.; Walsh, P.J. The Transcriptome of the Early Life History Stages of the California Sea Hare *Aplysia californica*. *Comp. Biochem. Physiol. Part D Genom. Proteom.* 2010, *5*, 165–170. https://doi.org/10.1016/j.cbd.2010.03.003.
40. Huan, P.; Wang, H.; Liu, B. Transcriptomic Analysis of the Clam *Meretrix meretrix* on Different Larval Stages. *Mar. Biotechnol.* 2012, *14*, 69–78. https://doi.org/10.1007/s10126-011-9389-0.

41. Qin, J.; Huang, Z.; Chen, J.; Zou, Q.; You, W.; Ke, C. Sequencing and *de Novo* Analysis of *Crassostrea angulata* (Fujian Oyster) from 8 Different Developing Phases Using 454 GSFlx. *PLoS ONE* 2012, *7*, e43653. https://doi.org/10.1371/journal.pone.0043653.
42. Bassim, S.; Tanguy, A.; Genard, B.; Moraga, D.; Tremblay, R. Identification of *Mytilus edulis* Genetic Regulators during Early Development. *Gene* 2014, *551*, 65–78. https://doi.org/10.1016/j.gene.2014.08.042.
43. Zhang, J.; Xiong, X.; Deng, Y.; Zheng, Z.; Yang, C.; Du, X. Integrated Application of Transcriptomics and Metabolomics Provides Insights into the Larval Metamorphosis of Pearl Oyster (*Pinctada fucata martensii*). *Aquaculture* 2021, *532*, 736067. https://doi.org/10.1016/j.aquaculture.2020.736067.
44. Xiong, X.; Cao, Y.; Li, Z.; Huang, R.; Du, X.; Zheng, Z. Ecdysone Signal Pathway Participates in Shell Formation in Pearl Oysters *Pinctada fucata martensii*. *J. Steroid Biochem. Mol. Biol.* 2022, *217*, 106045. https://doi.org/10.1016/j.jsbmb.2021.106045.
45. Allen, S.K.; Bushek, D. Large-Scale Production of Triploid Oysters, *Crassostrea virginica* (Gmelin), Using "Stripped" Gametes. *Aquaculture* 1992, *103*, 241–251. https://doi.org/10.1016/0044-8486(92)90170-P.
46. Sugawara, Y.; Koganezawa, A. Oysters. In *Oyster, Scallop and Abalone—Advances in Culture of Molluscs and Related Fields of Research (Japanese)*; Nomura, T., Sugawara, Y., Mori, K., M., Takeuchi, M., Nomachi, K., Matsutani T., Eds.; Kouseisha Kouseikaku: Tokyo, Japan, 1991; pp. 1–17.
47. Sugawara, Y. Oysters. In *Settlement Mechanism of Marine Organisms (Japanese)*; Kajiwara, T., Research Institute of Marine Invertebrates, Eds.; Kouseisha Kouseikaku: Tokyo, Japan, 1991; pp. 62–75; ISBN 978-4-7699-1205-7w.
48. Coon, S.L.; Fitt, W.K.; Bonar, D.B. Competence and Delay of Metamorphosis in the Pacific Oyster *Crassostrea gigas*. *Mar. Biol.* 1990, *106*, 379–387. https://doi.org/10.1007/BF01344316.
49. Martin, M. Cutadapt Removes Adapter Sequences from High-Throughput Sequencing Reads. *EMBnet. J.* 2011, *17*, 10–12. https://doi.org/10.14806/ej.17.1.200, accessed on 25 February 2022.
50. Joshi, N.; Fass, J. Sickle: A Sliding-Window, Adaptive, Quality-Based Trimming Tool for FastQ Files (Version 1.33) [Software], 2011. Available online: https://github.com/najoshi/sickle (accessed on 25 February 2022).

51. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-Level Expression Analysis of RNA-Seq Experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 2016, *11*, 1650–1667. https://doi.org/10.1038/nprot.2016.095, accessed on 25 February 2022.
52. Liao, Y.; Smyth, G.K.; Shi, W. FeatureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* 2014, *30*, 923–930. https://doi.org/10.1093/bioinformatics/btt656, accessed on 25 February 2022.
53. Chen, Y.; McCarthy, D.; Ritchie, M.; Robinson, M.; Smyth, G. EdgeR: Differential Analysis of Sequence Read Count Data. *R Package* 2008, 1–119. https://www.bioconductor.org/packages/devel/bioc/vignettes/edgeR/inst/doc/edgeRUsersGuide.pdf.
54. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2013; ISBN 3-900051-07-0. Available online: http//www.R-project.org, (accessed on 25 February 2022).
55. Conesa, A.; Götz, S.; García-Gómez, J.M.; Terol, J.; Talón, M.; Robles, M. Blast2GO: A Universal Tool for Annotation, Visualization and Analysis in Functional Genomics Research. *Bioinformatics* 2005, *21*, 3674–3676. https://doi.org/10.1093/bioinformatics/bti610.
56. Benjamini, Y.; Hochberg, Y. Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. *J. R. Stat. Soc. Ser. B* 1995, *57*, 289–300. https://doi.org/10.1111/j.2517-6161.1995.tb02031.x.
57. Trapnell, C.; Hendrickson, D.G.; Sauvageau, M.; Goff, L.; Rinn, J.L.; Pachter, L. Differential Analysis of Gene Regulation at Transcript Resolution with RNA-Seq. *Nat. Biotechnol.* 2013, *31*, 46–53. https://doi.org/10.1038/nbt.2450.
58. Madden, T. The BLAST Sequence Analysis Tool. In *The NCBI Handbook [Internet]*, 2nd ed.; National Center for Biotechnology Information (US): Bethesda, MD, USA, 2013; pp. 1–17.
59. Gao, Y.-L.; Yoshida, A.; Liu, J.-Y.; Shimizu, T.; Shirota, K.; Shiina, Y.; Osatomi, K. Molecular Cloning and Expression Dynamics of UNC-45B upon Heat Shock in the Muscle of Yellowtail. *Aquaculture* 2021, *541*, 736827. https://doi.org/10.1016/j.aquaculture.2021.736827.
60. Marie, B.; Zanella-Cléon, I.; Guichard, N.; Becchi, M.; Marin, F. Novel Proteins from the Calcifying Shell Matrix of the Pacific Oyster *Crassostrea gigas*. *Mar. Biotechnol.* 2011, *13*, 1159–1168. https://doi.org/10.1007/s10126-011-9379-2.

61. Mouchi, V.; Lartaud, F.; Guichard, N.; Immel, F.; de Rafélis, M.; Broussard, C.; Crowley, Q.G.; Marin, F. Chalky versus Foliated: A Discriminant Immunogold Labelling of Shell Microstructures in the Edible Oyster *Crassostrea gigas*. *Mar. Biol.* 2016, *163*, 256. https://doi.org/10.1007/s00227-016-3040-6.
62. Johnstone, M.B.; Wheeler, A.P.; Falwell, E.P.; Staton, M.E.; Saski, C.A.; Mount, A.S. Folian-Cv1 Is a Member of a Highly Acidic Phosphoprotein Class Derived From the Foliated Layer of the Eastern Oyster (*Crassostrea virginica*) Shell and Identified in Hemocytes and Mantle. *Front. Mar. Sci.* 2019, *6*, 1–21. https://doi.org/10.3389/fmars.2019.00366.
63. Coon, S.L.; Bonar, D.B. Pharmacological Evidence That Alpha1.-Adrenoceptors Mediate Metamorphosis of the Pacific Oyster, *Crassostrea gigas*. *Neuroscience* 1987, *23*, 1169–1174. https://doi.org/10.1016/0306-4522(87)90190-4.
64. Wang, G.; Liu, B.; Tang, B.; Zhang, T.; Xiang, J. Pharmacological and Immunocytochemical Investigation of the Role of Catecholamines on Larval Metamorphosis by β-Adrenergic-like Receptor in the Bivalve *Meretrix meretrix*. *Aquaculture* 2006, *258*, 611–618. https://doi.org/10.1016/j.aquaculture.2006.04.031.
65. Chen, Z.F.; Matsumura, K.; Wang, H.; Arellano, S.M.; Yan, X.; Alam, I.; Archer, J.A.C.; Bajic, V.B.; Qian, P.Y. Toward an Understanding of the Molecular Mechanisms of Barnacle Larval Settlement: A Comparative Transcriptomic Approach. *PLoS ONE* 2011, *6*, e22913. https://doi.org/10.1371/journal.pone.0022913.
66. Huang, L.X.; Gong, Y.J.; Gu, J.; Zeng, B.J.; Huang, L.H.; Feng, Q.L. Expression, Subcellular Localization and Protein-Protein Interaction of Four Isoforms of EcR/USP in the Common Cutworm. *Insect Sci.* 2015, *22*, 95–105. https://doi.org/10.1111/1744-7917.12101.
67. Chen, Y.; Lun, A.T.L.; Smyth, G.K. From Reads to Genes to Pathways: Differential Expression Analysis of RNA-Seq Experiments Using Rsubread and the EdgeR Quasi-Likelihood Pipeline. *F1000Research* 2016, *5*, 1438. https://doi.org/10.12688/F1000RESEARCH.8987.2.
68. Burke, R.D. The Induction of Metamorphosis of Marine Invertebrate Larvae: Stimulus and Response. *Can. J. Zool.* 1983, *61*, 1701–1719. https://doi.org/10.1139/z83-221.
69. Morse, D.E. Neurotransmitter-Mimetic Inducers of Larval Settlement and Metamorphosis. *Bull. Mar. Sci.* 1985, *37*, 697–706.

70. Trapido-Rosenthal, H.; Morse, D. Regulation of Receptor-Mediated Settlement and Metamorphosis in Larvae of a Gastropod Mollusc (*Haliotis rufescens*). *Bull. Mar. Sci.* 1986, *39*, 383–392.
71. Hirata, K.Y.; Hadfield, M.G. The Role of Choline in Metamorphic Induction of *Phestilla* (Gastropoda, Nudibranchia). *Comp. Biochem. Physiol. Part C Comp. Pharmacol.* 1986, *84*, 15–21. https://doi.org/10.1016/0742-8413(86)90158-1.
72. Satuito, C.G.; Natoyama, K.; Yamazaki, M.; Shimizu, K.; Fusetani, N. Induction of Metamorphosis in the Pediveliger Larvae of the Mussel *Mytilus galloprovincialis* by Neuroactive Compounds. *Fish. Sci.* 1999, *65*, 384–389. https://doi.org/10.2331/fishsci.65.384.
73. Yang, J.L.; Li, Y.F.; Bao, W.Y.; Satuito, C.G.; Kitamura, H. Larval Metamorphosis of the Mussel *Mytilus galloprovincialis* Lamarck, 1819 in Response to Neurotransmitter Blockers and Tetraethylammonium. *Biofouling* 2011, *27*, 193–199. https://doi.org/10.1080/08927014.2011.553717.
74. Coon, S.L.; Bonar, D.B.; Weiner, R.M. Chemical Production of Cultchless Oyster Spat Using Epinephrine and Norepinephrine. *Aquaculture* 1986, *58*, 255–262. https://doi.org/10.1016/0044-8486(86)90090-6.
75. Croll, R.P. Developing Nervous Systems in Molluscs: Navigating the Twists and Turns of a Complex Life Cycle. *Brain. Behav. Evol.* 2009, *74*, 164–176. https://doi.org/10.1159/000258664.
76. Chandramouli, K.H.; Sun, J.; Mok, F.S.; Liu, L.; Qiu, J.W.; Ravasi, T.; Qian, P.Y. Transcriptome and Quantitative Proteome Analysis Reveals Molecular Processes Associated with Larval Metamorphosis in the Polychaete *Pseudopolydora vexillosa*. *J. Proteome Res.* 2013, *12*, 1344–1358. https://doi.org/10.1021/pr3010088.
77. Song, X.; Liu, Z.; Wang, L.; Song, L. Recent Advances of Shell Matrix Proteins and Cellular Orchestration in Marine Molluscan Shell Biomineralization. *Front. Mar. Sci.* 2019, *6*, 16. https://doi.org/10.3389/fmars.2019.00041.
78. Chan, V.B.S.; Johnstone, M.B.; Wheeler, A.P.; Mount, A.S. Chitin Facilitated Mineralization in the Eastern Oyster. *Front. Mar. Sci.* 2018, *5*, 12. https://doi.org/10.3389/fmars.2018.00347.
79. De Wit, P.; Durland, E.; Ventura, A.; Langdon, C.J. Gene Expression Correlated with Delay in Shell Formation in Larval Pacific Oysters (*Crassostrea gigas*) Exposed to Experimental Ocean Acidification Provides Insights into Shell Formation Mechanisms. *BMC Genom.* 2018, *19*, 160. https://doi.org/10.1186/s12864-018-4519-y.

80. Chandramouli, K.H.; Al-Aqeel, S.; Ryu, T.; Zhang, H.; Seridi, L.; Ghosheh, Y.; Qian, P.Y.; Ravasi, T. Transcriptome and Proteome Dynamics in Larvae of the Barnacle *Balanus amphitrite* from the Red Sea. *BMC Genom.* 2015, *16*, 1–17. https://doi.org/10.1186/s12864-015-2262-1.
81. Chandramouli, K.H.; Mok, F.S.; Wang, H.; Qian, P.Y. Phosphoproteome Analysis during Larval Development and Metamorphosis in the Spionid Polychaete *Pseudopolydora vexillosa*. *BMC Dev. Biol.* 2011, *11*, 31. https://doi.org/10.1186/1471-213X-11-31.
82. Cranfield, H.J. Observations on the Function of the Glands of the Foot of the Pediveliger of *Ostrea edulis* during Settlement. *Mar. Biol.* 1973, *22*, 211–223. https://doi.org/10.1007/BF00389175.
83. Calado, R.; Leal, M.C. Trophic Ecology of Benthic Marine Invertebrates with Bi-Phasic Life Cycles: What Are We Still Missing? *Adv. Mar. Biol.* 2015, *71*, 1–70. https://doi.org/10.1016/bs.amb.2015.07.001.
84. Hadfield, M.G.; Carpizo-Ituarte, E.J.; Del Carmen, K.; Nedved, B.T. Metamorphic Competence, a Major Adaptive Convergence in Marine Invertebrate Larvae. *Am. Zool.* 2001, *41*, 1123–1131. https://doi.org/10.1093/icb/41.5.1123.
85. Lafont, R.; Koolman, J. Diversity of Ecdysteroids in Animal Species. In *Ecdysone: Structures and Functions*; Smagghe, G., Eds.; Springer: Dordrecht, The Netherlands, 2009; pp. 47–71. https://doi.org/10.1007/978-1-4020-9112-4_2.
86. Huang, W.; Xu, F.; Qu, T.; Zhang, R.; Li, L.; Que, H.; Zhang, G. Identification of Thyroid Hormones and Functional Characterization of Thyroid Hormone Receptor in the Pacific Oyster *Crassostrea gigas* Provide Insight into Evolution of the Thyroid Hormone System. *PLoS ONE* 2015, *10*, e0144991. https://doi.org/10.1371/journal.pone.0144991.
87. Gilbert, L.I., Tata, J.R., Atkinson, B.G., (Eds.) *Metamorphosis: Postembryonic Reprogramming of Gene Expression in Amphibian and Insect Cells*; Academic Press: Cambridge, MA, USA; 1996; Volume 103; p. 231. https://doi.org/10.1006/gcen.1996.0114. *Gen. Comp. Endocrinol.*

Chapter V: General Discussion and Conclusion

The role of chemical cues in habitat selection is crucial for marine invertebrates. Habitat selection is one of the key factors that determine the spatial-temporal distribution, especially for marine benthic invertebrates that must attach to their chosen settlement site for the rest of their lives [46]. Larval settlement induction is influenced by a variety of endogenous and exogenous factors [10,46]. In this present study, the role of chemical cues from shells of conspecifics in *Crassostrea gigas* larval settlement induction was investigated. Understanding the molecular mechanisms governing oyster substrate selection may provide new insights into the molecular evolution of regulatory signal molecules, chemical sensing mechanisms, and their roles in determining the distribution and abundance of benthic organisms [46]. A discussion is made on the key findings of the three phases of research conducted on the role of sugars in lectin-glycan interaction effects on conspecific cue-mediated settlement induction and its localization on oyster larval tissues (Chapter II); the identification and characterization of the biological cue responsible for gregarious settlement on conspecifics in *C. gigas* (Chapter III); and lastly, the transcriptome dynamics of the oyster larval response to the conspecific cue-mediated settlement induction in *C. gigas* (Chapter IV).

In Chapter II, the involvement of lectin-glycan interaction is demonstrated and the important role played by sugar compounds in settlement site selection by *C. gigas* larvae. Sugar compounds were found to influence the ability of the larvae to identify and discriminate suitable substrata in response to chemical cues in either inhibiting or enhancing lectin-mediated processes involved in biorecognition. Sugars in the absence of a conspecific cue did not promote settlement, whereas, in the presence of the cue showed varied effects, most of which were found inhibitory at different concentrations (Figure 2.2). This indicates that induction of settlement in the larvae may not only require the presence of sugars alone; but also amino acids and other chemical groups, present at the site of chemical contact between the larvae and the substrate [47,48], as observed in the high settlement response on the settlement inducing conspecific cue. This may also imply the need for multivalent interactions between multiple carbohydrate recognizing domains (CRDs) from lectins and multiple glycans that are often required to produce high-avidity binding interactions [48] for settlement induction to take effect. Hence, the high percentage of settlement on the conspecific cue could be in part, due to an increased avidity effect [48]. In addition, significant settlement inhibiting sugars indicate that a lectin-like receptor on the larvae seems to bind with sugars from different specificities: hexose sugar (αMDM), pentose sugar (D-arabinose), ketose sugar (D-

fructose), amino sugar (GlcNAc), sialic acid (Neu5Ac), reducing sugars (lactose, maltose) and non-reducing sugar (sucrose). The differential effects of sugar types on attachment to a substrate until settlement and metamorphosis suggest that the ratios of these sugars within a settlement inducing chemical cue could determine the maximum adhesion of settling larvae on a substrata [49]. The binding affinity of selected inhibitory sugars such as GlcNAc, Neu5Ac, and lactose is influenced not only by its concentration but also by exposure time, with 2 h duration eliciting the maximum inhibitive effect on oyster larval settlement in the presence of a conspecific cue (Figure 2.3). This evidence points to the possible involvement not only of GlcNAc sugars but also that of Neu5Ac residues, present in the presumed larval settlement inducer from a conspecific cue in shells of *C. gigas*. Results in this study further suggest that lactose may play a role as well in the settlement induction of *C. gigas*.

Blocking of WGA-binding sites on the conspecific cue from the shell extract significantly reduced its settlement induction efficiency and the presence of GlcNAc sugar binding to WGA lectin-like receptors distributed on different *C. gigas* larval tissues-tagged by WGA-FITC conjugate was confirmed (Figure 2.4). New findings in this study suggest that a WGA lectin-like receptor and its endogenous ligand are both found in the larval chemoreceptors (foot, mantle, and velum). At the same time, the conspecific cue from the shell CgSE has also been hypothesized to contain both a settlement-inducing endogenous ligand, as well as an endogenous lectin-like receptor that may complementarily work together to allow the larvae greater selectivity during site selection. It also provides new insights into how carbohydrate-lectin interaction could play an important role in the natural environment when larvae encounter different glycoproteins adsorbed to surfaces and are also exposed to dissolved cues. The effect of sugars on settlement is significant in the study of chemical communication and could have substantial implications for oyster ecology.

In chapter III, a novel role of several shell matrix proteins that acts as Settlement Pheromone Protein Components (CGSPPC) in adult shells is demonstrated as the biological cue responsible for gregarious settlement on *Crassostrea gigas* conspecifics. A bioassay-guided protein separation approach aided by biochemical and molecular analyses reveals that the isolated matrix proteins from the EDTA-soluble extract include Gigasin-6 isoform X1 and/or X2, *Crassostrea gigas* Surface protein P12p-like (CGS12P), and several stains-all stainable acidic proteins, the most dominant of which was a 48 kDa band, putatively identified as *Crassostrea gigas* Dentin sialophosphoprotein-like (CGDSP) (Figure 3.1, Table 3.1). Each protein factor may individually induce larval settlement, but evidence also shows that two or more gene products could cooperatively function together within the shell to create a stable

and strong settlement inducing signal. Some key factors that are attributed to their ability to induce settlement, either individually or collectively, are the presence of post-translational modifications (Figures 3.2-3.4) and unique sequences of amino acid groups (Figure 3.6, Appendix Figures A3.8 and A3.10) that may play important roles in oyster chemical signaling for conspecific recognition. Gigasin-6 isoform X1 and/or X2 which showed positive binding to WGA lectin is considered the major settlement inducing cue component. It plays a role in larval settlement induction on adult shells and may also function in postlarva-larva settlement interactions (Figure 3.5). CGDSP together with other stains-all stainable acidic proteins are also co-factors that could contribute to strengthening the settlement inducing signal via phosphorylation crosstalk with a CGSPPC recognizing larval receptor. CGDSP and some isolated stains-all stainable acidic proteins in this study require disulfide bonds to stabilize its settlement inducing effects and may play a role in providing the structural framework for other shell matrix proteins to anchor within the settlement protein components and provide protection to its multiple binding partners. The role of CGS12P as a co-factor in settlement induction warrants further investigation. Furthermore, conspecific cue-mediated larval settlement induction in *C. gigas* presents a complex system that requires an interplay of different amino acid groups, disulfide bonds, glycans, and phosphorylation crosstalk for conspecific recognition (Figure 3.7). On the other hand, the receptor that could recognize CGSPPC may involve a WGA lectin-like larval receptor and a phosphoprotein.

These present findings introduce another novel concept in chemical ecology demonstrating several protein factors/signaling molecules (CGSPPC), from *C. gigas* conspecific adult shells, that could induce larval settlement (Figure 3.7). Whereas, those reported in other marine invertebrates come from a single signaling molecule such as the α-macroglobulin-like gene that encodes the Settlement Inducing Protein Complex (SIPC) in the barnacles [24,50]. Notably, conspecific cue mediated larval settlement induction in *C. gigas* presents a complex system that requires an interplay of different glycans, disulfide bonds, amino acid groups, and phosphorylation crosstalk for recognition. On the other hand, the receptor that could recognize CGSPPC may involve a WGA lectin-like larval receptor and a phosphoprotein. In contrast, as shown in Figure 3.7 B, reported gregarious conspecific cues in other benthic organisms such as the barnacle SIPC presents a less complex lectin-glycan interaction during settlement induction as it is specifically modulated by LCA/Con A-binding sugars such as mannose [50,51].

In Chapter IV, the mode of action of a naturally occurring inducer such as the chemical cue from shells of conspecifics (CGSPPC) is demonstrated. A transcriptome profile of an oyster larval response to CGSPPC, a conspecific cue-mediated settlement induction in *C. gigas* is reported for the first time. Transcriptomics analysis has identified key transcripts that were differentially expressed in the pediveliger and conspecific cue-induced postlarvae in *C. gigas*. Quantitative analysis revealed 2,383 candidate transcripts, with 740 upregulated and 1,643 downregulated transcripts after settlement, that may play important roles during larval settlement processes (Figure 4.2, Table 4.3). Several transcripts with differential expression were significantly enriched in GO analysis associated with chitin-binding, calcium-ion binding, and extracellular region-related categories (Table 4.4). Differential expressions of six candidate transcripts representing Pedi and PL batches were validated by quantitative Real Time-PCR, which showed relative expression trends were consistent between qRT-PCR and transcriptome data (Figure 4.3). By comparing the expression patterns of the Pedi and conspecific cue-induced PL batches, shell formation was hypothesized to show closely linked dynamics with a gene regulatory network in *C. gigas* that may involve the interplay of various hormone receptors, neurotransmitters, and neuropeptide receptors working together in a concerted way in the induction of conspecific cue-mediated larval settlement.

Based on these findings, the molecular mechanism of the transcriptome dynamics of an oyster larval response to a conspecific cue-mediated settlement induction in *C. gigas* is proposed (Figure 4.4): As the pediveliger commences its searching behavior, it makes use of its external chemoreceptors to detect signals for a suitable location for settlement. Once it detects a ligand and other signaling cues coming from shells of conspecifics, various intracellular signaling receptors act in a concerted way and a molecular switch is turned on for the initiation of downstream signal cascades into the nucleus with the possible involvement of the ecdysone signaling pathway. Then, various transcription factors and gene expressions for the synthesis of specific proteins are activated. Subsequently, rapid morphological, physiological, structural, and functional changes occur as the competent pediveliger undergoes its metamorphosis into a postlarva within 24 h. These results highlight the transcriptome dynamics underlying the gregarious settlement of oysters on conspecific adult shells. Furthermore, this study hopes to provide hints and support for future studies related to the involvement of the ecdysone signal pathway and the linkages within the neuroendocrine-biomineralization crosstalk in *C. gigas* settlement.

Having discussed the individual and synergistic roles of the protein components within CGSPPC, the regulatory role of sugars and CGSPPC, and the transcriptome dynamics in oyster

larval responses to conspecific cue mediated larval settlement induction, the molecular basis of the conspecific cue-mediated larval settlement mechanism is proposed herein as shown in Figure 5.1:

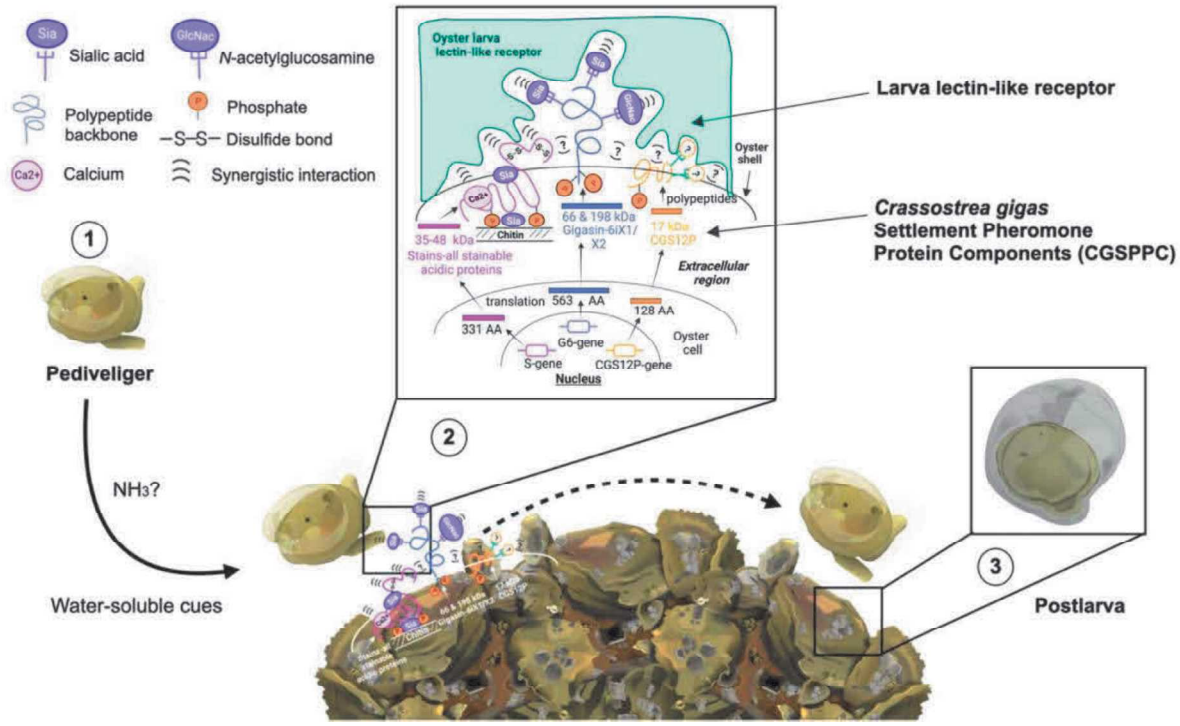

Figure 5.1. Schematic of the *Crassostrea gigas* Settlement Pheromone Protein Components (CGSPPC)-mediated larval settlement induction mechanism. (**1**) As the oyster larva becomes a competent pediveliger, with a fully developed eye and foot, it begins to exhibit searching behaviors. This searching behavior is influenced by the detection of water-soluble cues from metabolites, including glycans, NH_3, and other amino acids, from bacteria or their conspecifics. (**2**) By locating the source of the inducing cue, the larva recognizes the amino acid groups, glycan moieties, and phosphate groups on the surface-bound cue, CGSPPC, that are found on shells of conspecific postlarvae and adult oysters. The noncovalent bond interaction between the lectin-glycan and/or phosphate groups in both the larva and substrate cue,i.e., shells of conspecifics, allows the accumulation of transduced cues and creates a seemingly complex recognition system in *C. gigas*. This allows higher affinity, specificity, and complementarity between CGSPPC and its receptor to ensure greater selectivity during site selection. (**3**) Having accumulated enough signals from a suitable settlement site, a cascade of biochemical changes is triggered that leads to the attachment and its eventual metamorphosis into a postlarva. The postlarva develops newly formed gills and a new layer of an adult shell while its eye, foot, and velum disappear. This figure was partly created using BioRender.com (accessed on 25 July 2022).

(1) When the oyster larva reaches a competent stage wherein the eye and the foot fully develop. The larva begins to exhibit searching behaviors such as swimming motions with a combination of foot extensions and localized crawling movements until it finds a suitable

substratum, i.e., shells of conspecifics [39]. The presence of dissolved water-soluble cues from the environment such as dissolved sugars [52,53], NH_3 from metabolites, including glycans and amino acids from bacteria [41,54] or those from conspecifics [41] induces larval settlement behavior. It has been reported by Bonar et al. (1990) [23] that elevated levels of NH_3 produced by a dense assemblage of oysters such as those of an oyster bed trigger the oyster larva to swim in search of the source of the corresponding cue. (2) As the larva locates the source of the inducing cue, it uses its external chemical receptors, such as those found in the velum and foot, to transduce the surface-bound cues it encounters on the substrata [25]. It then recognizes the amino acid groups, sugar moieties, and phosphate groups attached to the protein components in CGSPPC both from the newly attached juveniles and adult oysters. In this study, the settlement inducing activity of CGSPPC was concentration-dependent implying that as the larva encounters more of its conspecifics, an increased settlement rate was observed. Furthermore, the interaction between the CGSPPC recognizing larval receptor and this conspecific cue is reversible since their interaction is characterized by the formation of noncovalent bonds [28,52]. This allows the oyster larva to discriminate cues and gather confirmatory signals from the putative settlement substratum. (3) As the larva continues to exhibit a searching behavior along this area, a point is reached that enough levels of accumulated signals (in the form of amino acid/sugar/phosphate groups) from the settlement inducing cue, CGSPPC, trigger a cascade of biochemical changes that leads to the attachment and its eventual metamorphosis into a post larva. In this rapid metamorphic change, the larva experiences the irreversible physiological process that includes a change from a swimming larva to a sessile juvenile, including loss of the velum, eyespot, and foot; development of gills, and production of adult shell [10].

In conclusion, oysters may seem simple, but this study provides insights into the complex chemical signaling processes and molecular mechanisms governing conspecific cue-mediated larval settlement induction that may indicate their evolutionary success as benthic organisms. Continued study on CGSPPC and its corresponding oyster larval receptor will advance our understanding of their evolution, population dynamics, and chemical communication. These results may find application in the development of oyster aquaculture by using this extracted compound as a surface-bound attractant to collect wild and hatchery-grown oyster larvae which could help recover declining marine species; this compound could also find application as a target of anti-fouling agents on man-made structures.

Chapter VI. General References

1. Guo, X.; Li, Q.; Wang, Q.Z.; Kong, L.F. Genetic Mapping and QTL Analysis of Growth-Related Traits in the Pacific Oyster. *Mar. Biotechnol.* 2012, *14*, doi:10.1007/s10126-011-9405-4.
2. Zhang, G.; Fang, X.; Guo, X.; Li, L.; Luo, R.; Xu, F.; Yang, P.; Zhang, L.; Wang, X.; Qi, H.; et al. The Oyster Genome Reveals Stress Adaptation and Complexity of Shell Formation. *Nature* 2012, *490*, 49–54, doi:10.1038/nature11413.
3. Hedgecock, D.; Gaffney, P.M.; Goulletquer, P.; Guo, X.; Reece, K.; Warr, G.W. The Case for Sequencing the Pacific Oyster Genome. *J. Shellfish Res.* 2005, *24*, 429–441, doi:10.2983/0730-8000(2005)24[429:TCFSTP]2.0.CO;2.
4. Huan, P.; Wang, H.; Liu, B. A Label-Free Proteomic Analysis on Competent Larvae and Juveniles of the Pacific Oyster (*Crassostrea gigas*). *PLoS One* 2015, *10*, e0135008, doi:10.1371/journal.pone.0135008.
5. Xu, F.; Zhang, G. Transcriptomic and Proteomic Dynamics during Metamorphosis of Pacific Oyster *Crassostrea gigas*). *bioRxiv* 2020, 2020.03.25.004614, doi:https://doi.org/10.1101/2020.03.25.004614.
6. Fitt, W.K.; Coon, S.L.; Walch, M.; Weiner, R.M.; Colwell, R.R.; Bonar, D.B. Settlement Behavior and Metamorphosis of Oyster Larvae (*Crassostrea gigas*) in Response to Bacterial Supernatants. *Mar. Biol.* 1990, *106*, 389–394, doi:10.1007/BF01344317.
7. Coon, S.L.; Bonar, D.B.; Weiner, R.M. Induction of Settlement and Metamorphosis of the Pacific Oyster, *Crassostrea gigas* (Thunberg), by L-DOPA and Catecholamines. *J. Exp. Mar. Bio. Ecol.* 1985, *94*, doi:10.1016/0022-0981(85)90059-0.
8. Gosling, E. Reproduction, Settlement and Recruitment. In *Marine Bivalve Molluscs*; John Wiley & Sons, Ltd: Chichester, UK, 2015; pp. 157–202.
9. Beiras, R.; Widdows, J. Induction of Metamorphosis in Larvae of the Oyster *Crassostrea gigas* Using Neuroactive Compounds. *Mar. Biol.* 1995, *123*, 327–334, doi:10.1007/BF00353624.
10. Joyce, A.; Vogeler, S. Molluscan Bivalve Settlement and Metamorphosis: Neuroendocrine Inducers and Morphogenetic Responses. *Aquaculture* 2018, *487*, 64–82, doi:10.1016/j.aquaculture.2018.01.002.
11. Yurchenko, O. V.; Skiteva, O.I.; Voronezhskaya, E.E.; Dyachuk, V.A. Nervous System Development in the Pacific Oyster, *Crassostrea gigas* (Mollusca: Bivalvia). *Front. Zool.* 2018, *15*, 10, doi:10.1186/s12983-018-0259-8.

12. Marie, B.; Zanella-Cléon, I.; Guichard, N.; Becchi, M.; Marin, F. Novel Proteins from the Calcifying Shell Matrix of the Pacific Oyster *Crassostrea gigas*. *Mar. Biotechnol.* 2011, *13*, 1159–1168, doi:10.1007/s10126-011-9379-2.
13. Wang, X.; Song, X.; WANG, T.; Zhu, Q.; Miao, G.; Chen, Y.; Fang, X.; Que, H.; Li, L.; Zhang, G. Evolution and Functional Analysis of the Pif97 Gene of the Pacific Oyster *Crassostrea gigas*. *Curr. Zool.* 2013, *59*, 109–115, doi:10.1093/czoolo/59.1.109.
14. Song, X.; Wang, X.; Li, L.; Zhang, G. Identification Two Novel Nacrein-like Proteins Involved in the Shell Formation of the Pacific Oyster *Crassostrea gigas*. *Mol. Biol. Rep.* 2014, *41*, doi:10.1007/s11033-014-3298-z.
15. De Wit, P.; Durland, E.; Ventura, A.; Langdon, C.J. Gene Expression Correlated with Delay in Shell Formation in Larval Pacific Oysters (*Crassostrea gigas*) Exposed to Experimental Ocean Acidification Provides Insights into Shell Formation Mechanisms. *BMC Genomics* 2018, *19*, 160, doi:10.1186/s12864-018-4519-y.
16. Grabowski, J.H.; Brumbaugh, R.D.; Conrad, R.F.; Keeler, A.G.; Opaluch, J.J.; Peterson, C.H.; Piehler, M.F.; Powers, S.P.; Smyth, A.R. Economic Valuation of Ecosystem Services Provided by Oyster Reefs. *Bioscience* 2012, *62*, 900–909, doi:10.1525/bio.2012.62.10.10.
17. Gentry, R.R.; Alleway, H.K.; Bishop, M.J.; Gillies, C.L.; Waters, T.; Jones, R. Exploring the Potential for Marine Aquaculture to Contribute to Ecosystem Services. *Rev. Aquac.* 2020, *12*, 499–512, doi:10.1111/raq.12328.
18. Ridlon, A.D.; Wasson, K.; Waters, T.; Adams, J.; Donatuto, J.; Fleener, G.; Froehlich, H.; Govender, R.; Kornbluth, A.; Lorda, J.; et al. Conservation Aquaculture as a Tool for Imperiled Marine Species: Evaluation of Opportunities and Risks for Olympia Oysters, *Ostrea lurida*. *PLoS One* 2021, *16*, e0252810, doi:10.1371/journal.pone.0252810.
19. FAO. *The State of World Fisheries and Aquaculture 2020. Sustainability in Action. Rome. Https://Doi.Org/10.4060/Ca9229en*; 2020;
20. Robert, R.; Gerard, A. Bivalve Hatchery Technology: The Current Situation for the Pacific Oyster and the Scallop in France. *Aquat. Living Resour.* 1999, *12*, 121–130, doi:10.1016/S0990-7440(99)80021-7.
21. Helm, M.M.; Bourne, N. *Hatchery Culture of Bivalves. A Practical Manual*; Lovatelli, A., Ed.; FAO: Rome, Italy, 2004; Vol. 471; ISBN 92-5-105224-7.
22. Beck, M.W.; Brumbaugh, R.D.; Airoldi, L.; Carranza, A.; Coen, L.D.; Crawford, C.; Defeo, O.; Edgar, G.J.; Hancock, B.; Kay, M.C.; et al. Oyster Reefs at Risk and Recommendations for Conservation, Restoration, and Management. *Bioscience* 2011,

61, 107–116, doi:10.1525/bio.2011.61.2.5.

23. Bonar, D.B.; Coon, S.L.; Walch, M.; Weiner, R.M.; Fitt, W. Control of Oyster Settlement and Metamorphosis by Endogenous and Exogenous Chemical Cues. *Bull. Mar. Sci.* 1990, *46*, 484–498.
24. Dreanno, C.; Matsumura, K.; Dohmae, N.; Takio, K.; Hirota, H.; Kirby, R.R.; Clare, A.S. An α 2 -Macroglobulin-like Protein Is the Cue to Gregarious Settlement of the Barnacle *Balanus amphitrite*. *Proc. Natl. Acad. Sci.* 2006, *103*, 14396–14401, doi:10.1073/pnas.0602763103.
25. Hadfield, M.G. Biofilms and Marine Invertebrate Larvae: What Bacteria Produce That Larvae Use to Choose Settlement Sites. *Ann. Rev. Mar. Sci.* 2011, *3*, 453–470, doi:10.1146/annurev-marine-120709-142753.
26. Heyland, A.; Moroz, L.L. Signaling Mechanisms Underlying Metamorphic Transitions in Animals. *Integr. Comp. Biol.* 2006, *46*, 743–759, doi:10.1093/icb/icl023.
27. Vasquez, H.E.; Hashimoto, K.; Yoshida, A.; Hara, K.; Imai, C.C.; Kitamura, H.; Satuito, C.G. A Glycoprotein in Shells of Conspecifics Induces Larval Settlement of the Pacific Oyster *Crassostrea gigas*. *PLoS One* 2013, *8*, e82358, doi:10.1371/journal.pone.0082358.
28. Vasquez, H.E.; Hashimoto, K.; Kitamura, H.; Satuito, C.G. Wheat Germ Agglutinin-Binding Glycoprotein Extract from Shells of Conspecifics Induces Settlement of Larvae of the Pacific Oyster *Crassostrea gigas* (Thunberg). *J. Shellfish Res.* 2014, *33*, 415–423, doi:10.2983/035.033.0210.
29. Crisp, D.J. Chemical Factors Inducing Settlement in *Crassostrea virginica* (Gmelin). *J. Anim. Ecol.* 1967, *36*, 329–335, doi:10.2307/2916.
30. Poirier, L.A.; Clements, J.C.; Davidson, J.D.P.; Miron, G.; Davidson, J.; Comeau, L.A. Sink before You Settle: Settlement Behaviour of Eastern Oyster (*Crassostrea virginica*) Larvae on Artificial Spat Collectors and Natural Substrate. *Aquac. Reports* 2019, *13*, doi:10.1016/j.aqrep.2019.100181.
31. Tamburri, M.N.; Luckenbach, M.W.; Breitburg, D.L.; Bonniwell, S.M. Settlement of Crassostrea Ariakensis Larvae: Effects of Substrate, Biofilms, Sediment and Adult Chemical Cues. *J. Shellfish Res.* 2008, *27*, 601–608, doi:10.2983/0730-8000(2008)27[601:SOCALE]2.0.CO;2.
32. Bayne, B.L. The Gregarious Behaviour of the Larvae of *Ostrea edulis* L. at Settlement. *J. Mar. Biol. Assoc. United Kingdom* 1969, *49*, 327–356, doi:10.1017/S0025315400035943.

33. Hirata, Y.; Tamura, Y.; Nagasawa, K. Influence of Presoaking Conditions of Spat Collectors in Seawater Containing Adult Pacific Oysters (*Crassostrea gigas*) on Their Larval Settlement. *Nippon Suisan Gakkaishi (In Japanese)* 2008, *74*, 1017–1023, doi:10.2331/suisan.74.1017.
34. Zimmer-Faust, R.K.; Tamburri, M.N. Chemical Identity and Ecological Implications of a Waterborne, Larval Settlement Cue. *Limnol. Oceanogr.* 1994, *39*, 1075–1087, doi:10.4319/lo.1994.39.5.1075.
35. Turner, E.J.; Zimmer-Faust, R.K.; Palmer, M.A.; Luckenbach, M.; Pentchef, N.D. Settlement of Oyster (*Crassostrea virginica*) Larvae: Effects of Water Flow and a Water-soluble Chemical Cue. *Limnol. Oceanogr.* 1994, *39*, 1579–1593, doi:10.4319/lo.1994.39.7.1579.
36. Carroll, J.M.; Riddle, K.; Woods, K.E.; Finelli, C.M. Recruitment of the Eastern Oyster, *Crassostrea virginica*, in Response to Settlement Cues and Predation in North Carolina. *J. Exp. Mar. Bio. Ecol.* 2015, *463*, 1–7, doi:10.1016/j.jembe.2014.10.024.
37. Pascual, M.S.; Zampatti, E.A. Evidence of a Chemically Mediated Adult-Larval Interaction Triggering Settlement in *Ostrea puelchana*: Applications in Hatchery Production. *Aquaculture* 1995, *133*, 33–44, doi:10.1016/0044-8486(95)00005-M.
38. Veitch, F.P.; Hidu, H. Gregarious Setting in the American Oyster *Crassostrea virginica* Gmelin: I. Properties of a Partially Purified "Setting Factor." *Chesap. Sci.* 1971, *12*, 173, doi:10.2307/1350777.
39. Coon, S.L.; Walch, M.; Fitt, W.K.; Weiner, R.M.; Bonar, D.B. Ammonia Induces Settlement Behavior in Oyster Larvae. *Biol. Bull.* 1990, *179*, 297–303, doi:10.2307/1542321.
40. Fitt, W.K.; Labare, M.P.; Fuqua, W.C.; Walch, M.; Coon, S.L.; Bonar, D.B.; Colwell, R.R.; Weiner, R.M. Factors Influencing Bacterial Production of Inducers of Settlement Behavior of Larvae of the Oyster *Crassostrea gigas*. *Microb. Ecol.* 1989, *17*, 287–298, doi:10.1007/BF02012841.
41. Tamburri, M.N.; Zimmer-Faust, R.K.; Tamplin, M.L. Natural Sources and Properties of Chemical Inducers Mediating Settlement of Oyster Larvae: A Re-Examination. *Biol. Bull.* 1992, *183*, 327–338, doi:10.2307/1542218.
42. Hashimoto, K.; Vasquez, H.E.; Kitamura, H.; Satuito, C.G. Larval Settlement of the Pacific Oyster *Crassostrea gigas* in Response to Marine Microbial Films and Monospecies Bacterial Films. *Sess. Org.* 2017, *34*, doi:10.4282/sosj.34.21.
43. Christensen, B.E. The Role of Extracellular Polysaccharides in Biofilms. *J. Biotechnol.*

1989, *10*, 181–202, doi:10.1016/0168-1656(89)90064-3.

44. Neu, T.R.; Marshall, K.C. Bacterial Polymers: Physicochemical Aspects of Their Interactions at Interfaces. *J. Biomater. Appl.* 1990, *5*, 107–133, doi:10.1177/088532829000500203.
45. Weiner, R.; Langille, S.; Quintero, E. Structure, Function and Immunochemistry of Bacterial Exopolysaccharides. *J. Ind. Microbiol.* 1995, *15*, 339–346, doi:10.1007/BF01569989.
46. Thiyagarajan, V. A Review on the Role of Chemical Cues in Habitat Selection by Barnacles: New Insights from Larval Proteomics. *J. Exp. Mar. Bio. Ecol.* 2010, *392*, 22–36, doi:10.1016/j.jembe.2010.04.030.
47. Tommasone, S.; Allabush, F.; Tagger, Y.K.; Norman, J.; Köpf, M.; Tucker, J.H.R.; Mendes, P.M. The Challenges of Glycan Recognition with Natural and Artificial Receptors. *Chem. Soc. Rev.* 2019, *48*, 5488–5505, doi:10.1039/C8CS00768C.
48. Taylor, M.E.; Drickamer, K.; Schnaar, R.L.; Etzler, M.E.; Varki, A. Discovery and Classification of Glycan-Binding Proteins. In *Essentials of Glycobiology [Internet]*; Cold Spring Harbor Laboratory Press: Cold Spring Harbor, NY, USA, 2017 ISBN 9780879697709.
49. Neal, A.L.; Yule, A.B. The Effects of Dissolved Sugars Upon the Temporary Adhesion of Barnacle Cyprids. *J. Mar. Biol. Assoc. United Kingdom* 1996, *76*, 649–655, doi:10.1017/S0025315400031350.
50. Matsumura, K.; Nagano, M.; Fusetani, N. Purification of a Larval Settlement-Inducing Protein Complex (SIPC) of the Barnacle, *Balanus amphitrite*. *J. Exp. Zool.* 1998, *281*, doi:10.1002/(SICI)1097-010X(19980501)281:1<12::AID-JEZ3>3.0.CO;2-F.
51. Matsumura, K.; Mori, S.; Nagano, M.; Fusetani, N. Lentil Lectin Inhibits Adult Extract-Induced Settlement of the Barnacle,*Balanus amphitrite*. *J. Exp. Zool.* 1998, *280*, 213–219, doi:10.1002/(SICI)1097-010X(19980215)280:3<213::AID-JEZ2>3.0.CO;2-Q.
52. Sedanza, M.G.; Kim, H.-J.; Seposo, X.; Yoshida, A.; Yamaguchi, K.; Satuito, C.G. Regulatory Role of Sugars on the Settlement Inducing Activity of a Conspecific Cue in Pacific Oyster *Crassostrea gigas*. *Int. J. Mol. Sci.* 2021, *22*, 3273, doi:10.3390/ijms22063273.
53. Khandeparker, L.; Anil, A.C. Role of Conspecific Cues and Sugars in the Settlement of Cyprids of the Barnacle, *Balanus amphitrite*. *J. Zool.* 2011, *284*, 206–214, doi:10.1111/j.1469-7998.2011.00794.x.
54. Mitchell, R. Colonization by Higher Organisms. In *Microbial Adhesion and*

Aggregation; Marshall, K.C., Ed.; Springer Berlin Heidelberg: Berlin, Heidelberg, 1984; pp. 189–200.

Appendices

Appendix Text: Results (Chapter III)

A3. First Protein Separation Approach

A3.1 Fractionation by Ultrafiltration

Settlement percentages of *C. gigas* larval response to fractionated forms of CgSE with varying extract amounts are shown in Figure A3.1. Fractionation by ultrafiltration was done by separating molecular weight cut-off (MWCO) of the settlement inducing compound into (a) >100K and the <100K fraction was further sub-divided into (b) <50K and (c) >50K CgSE fractions. All fractions except <50K CgSE showed a significant amount of extract-dependent settlement inducing activity ($p < 0.05$). Moreover, quasi-binomial glm analysis showed that at 1 µg, all fractions elicited different larval settlement inducing responses that ranged from 0 to 46%, the highest of which came from >50K CgSE (46%). On the other hand, both crude CgSE and >50K CgSE with settlement percentages of 70% and 74%, respectively, were shown to elicit the highest settlement response at 100 µg amount of the extract ($p < 0.05$). The >50K CgSE was then chosen for the next purification step, Gel filtration chromatography.

A3.2 Gel filtration of >50K CgSE

The active fraction, >50K CgSE, was then applied to Superdex 200 10/300 GL column and eluted with 0.15 M NaCl, using the FPLC system. Most of the proteins were eluted in the high molecular mass fraction (F2) while some proteins were of low molecular mass (F3 to F6) as shown in Figures A3.2 and A3.3. Characterization of proteins that were visualized only by Stains-all following SDS-PAGE is found in Figure A3.3 A. This cationic dye binds to anionic sites within Ca2+ -binding proteins and stains dark blue or purple for highly acidic and phosphorylated proteins [1]. While proteins of other classes stain red or pink [1]. Stains-all has also been widely used to identify proteins that contain high levels of sulfated sugar residues, acidic proteins, and phosphoproteins in mineralized tissues such as those derived from shell, tooth, and bone [1,2]. In Figure A3.3 A, CgSE showed three prominent, blue-stained bands at 42, 44, and 48 kDa. While gel filtration eluted fraction F2 also showed 3 closely associated bands similar to CgSE at 42, 45, and 50 kDa, but with some 'polydisperse' smeared background around these positions. However, fraction F3 did not show any protein bands suggesting that this active fraction might be composed of low molecular weight proteins. Hence, this additional factor was considered in favor of fraction F2 which was chosen for the next purification step. In terms of the larval settlement bioassay results and as shown in Figure A3.3 B, no significant response was elicited at the void volume in all extract amounts (fraction, F1; $p > 0.05$). Fraction,

F2, with a molecular mass range of 41 to 118 kDa, showed a consistent high settlement inducing activity of 7%, 46%, and 52% at 1, 10, and 50 µg, respectively, and in comparison, to other eluted fractions ($p < 0.05$). It is interesting to note, that while F2 showed a consistent settlement inducing effect on the oyster larvae across all extract amounts, low molecular mass fractions F3, F4, and F5 elicited high inducing effect at 50 µg with settlement percentages at 55%, 55%, and 28%, respectively ($p < 0.05$).

A3.3 Separation of F2 proteins

To determine whether the polypeptides in the active fraction F2 could elicit a settlement inducing response in the oyster larvae, each polypeptide was isolated and subjected to a bioassay. The fraction, F2, was resolved in SDS-PAGE under nonreducing and reducing (in the presence of 2-mercaptoethanol) [3] conditions (Figure A3.4 A). At a high amount of sample (60 µg) loaded on each lane and under reducing conditions, only 3 prominent bands appeared at positions 45, 48, and 53 kDa. While under non-reducing conditions, more bands were observed such that, at positions 35, 38, 43, 48, and 60 kDa, made it possible for the isolation of these polypeptides. However, among these polypeptides, bands 43 and 48 kDa showed the most dominant signal. Bioassay results of active fraction F2 and its isolated polypeptides are found in Figure A3.4 B. Settlement inducing activity was not observed in all isolated polypeptides under reducing conditions. This suggests that disulfide bonds might be necessary for the folding and stabilization of the settlement inducing factors in this cue [4]. However, under nonreducing conditions, II, IV, and V polypeptides of active fraction F2 showed significant induced larval responses compared with the control, larvae immersed on the multi-well with filtered seawater only (C-FSW) ($p < 0.05$). Notably, the settlement percentages of the isolated polypeptides were lower compared to the active fraction (F2) contrary to other reported gregarious cues in other marine invertebrates when the usual effect of a purified cue was stronger [5]. Thus, these results imply that these isolated polypeptides may need to interact synergistically to create a strong and stable inducing signal as a group. Amino acid sequence analysis using a protein sequencer failed to identify these active polypeptide components most likely due to their N-termini being blocked.

B3 Molecular characterization of Gigasin-6 isoform X1

B3.1 Bioinformatic characterization of Gigasin-6 isoform X1 protein sequence

As shown in Appendix Figure A3.6, Gigasin-6 isoform X1 has a theoretical molecular mass of 64 kDa (563-aa-long residue) while isoform X2 is a 63 kDa (554-aa-long residue) protein. The disparity of molecular masses between the observed and theoretical mass may be due to the additional presence of post-translational modifications [6]. Gigasin-6 isoform X1 was chosen for further protein sequence characterization and qRT-PCR analyses due to its complete and longer length of amino acid residues that best represents the Gigasin-6 family of proteins. Removal of the first 33 amino acid signal peptide sequence results in a 530 amino acid protein with a calculated molecular weight of 59,920 Da. The amino acid composition of the mature form of this protein is rich in Leu (9.8%), Thr (9.1%), and Asp (8.1%). It has 71 negatively charged residues (Asp + Glu) and 65 positively charged residues, with a theoretical pI of 5.74. A SMART search for protein domains suggests that Gigasin-6 isoform X1 contains a Beta-lactamase domain but may be an inactive form of lactamase-related protein [7,8]. While InterProscan predicts this protein with a non-cytoplasmic domain and contains a transmembrane helix. However, further analysis by the Transmembrane Helices Hidden Markov Model (TMHHH) on the amino acid residues at positions 17 to 36, which was predicted as a transmembrane helix site, revealed that this site coincides within the signal peptide and was not considered to be a true transmembrane helix. Low Complexity Regions (LCR) were identified using IUPRED search and are illustrated in Figure 3.5A, Appendix Figure A3.6. Three LCR sites were shown to be distributed throughout the entire protein sequence. In this study, the mass spectrometry analysis identified the majority of the peptide sequences except those within the LCR and close to the signal peptide region. To determine post-translational modifications, Net*N*Glyc search predicted 5 putative *N*-glycosylation sites, of which two of these sites were identified by mass spectrometry analysis in this study. Other *N*-glycosylation sites were found within low complexity regions which do not contain any tryptic cleavage sites. Using another kind of enzyme or a combination of enzymes might enable mass spectrometry identification for these remaining *N*-glycosylation sites. Net*O*Glyc search suggests 12 possible *O*-glycosylation sites. Use of NetPhos search predicted 70 phosphorylation sites while motif analyses by ScanProsite suggested 24 out of the 70 identified phosphorylation sites might be related to signal transduction processes: 7 Protein Kinase C, 15 Casein Kinase II, 1 Tyrosine, and 1 cAMP and cGMP dependent protein kinase phosphorylation sites. Also, ScanProsite predicted 1 RGD cell attachment sequence at positions 177 to 179.

B3.2 Localization of N-glycosylation sites by Mass spectrometry analysis

N-glycosylation sites on Gigasin-6 isoform X1 and/or X2 were identified by comparing mass fingerprint patterns of the glycosylated and deglycosylated bands B5 and B7, respectively (Figure A3.7). Treatment of glycoproteins with PNGase F causes a 0.98 Da mass shift at the *N*-glycosylation site when this reaction converts *N*-linked asparagine residues into aspartic acid residues [9]. The deamination of an occupied asparagine residue by PNGase F combined with mass determination by MALDI MS is often sufficient to assign *N*-glycosylation sites [10]. Through this strategy, 4 newly appeared m/z signals were identified at 2397.12, 2525.21, 2238.23, and 2254.19 (Figure A3.7 B and Appendix Table 3.2). Figures A3.7 C and D present zoomed regions of the MALDI MS/MS fragmentation spectra of the putative *N*-glycosylation sites in the deglycosylated digest. The first two of these signals (Figure A3.7 C) were identified through an error-tolerant search while the latter two signals (Figure A3.7 D) by manual annotation of the reported spectra with the aid of the MS-Product bioinformatics tool. The peptide with an m/z signal of 2397.12 corresponds to the position of Asn521 carrying the putative *N*-linked (NST) site (^{521}**NST**YIEAFTVDKFDDAKFER^{540}). CID fragmentation spectrum of this ion shows intense signals at fragment ions y_5,y_8, y_9, y_{14}, and b_{19}, corresponding to the PNGase-F-generated aspartic acids of the peptide (Figure A3.7 C). The second m/z 2525.21 signal which showed the highest peak intensity was predicted to result from the same peptide (Figure A3.7 C). CID fragmentation spectrum of this ion shows intense signals at fragment ions y_5, y_6, y_9, y_{14}, b_{15}, and b_{19}. Also, Mascot search results indicate that this signal came from the protonated parent ion that gained a Glu (+ 129 Da) residue giving it an overall m/z shift from 2397.12 to 2525.21. Hence, using an error-tolerant search, this fragment ion was predicted as an actual modification on this peptide containing an *N*-glycosylation site. Interestingly, m/z 2525.21, has not been recorded in any current protein database. Therefore, this is the first time to report the existence of this actual m/z signal for this protein. This may be part of the actual post-translational modification of this protein. Moreover, the third and fourth signal at m/z 2238.23 and 2254.19 corresponds to the position of Asn298 carrying the putative *N*-linked (NGS) site from the same peptide (^{291}FMNYLLG**NGS**IPGTNDVLLAK^{309}) with the expected mass shift following the glycan moiety removal. Although the CID fragmentation pattern of these ion signals (Figure A3.7 D) was of low intensity to yield sufficient confirmation of its sequence through error tolerant search, their fragment ion y and b series patterns were consistent with the predicted fragment ion profiles of this sequence in MS-Product bioinformatic tool as shown in Figure A3.7 D. Using a different endopeptidase with a combination of trypsin with another endopeptidase might yield a stronger signal

intensity for this putative *N*-glycosylation site. However, the presence of these m/z values and the positive unitary mass shift that occurred after PNGase F treatment, combined with the absence of the same m/z value in the glycosylated sample, strongly support the assignment made [10,11].

C3. Molecular characterization of Stains-all stainable acidic proteins

By exploring the putative identity of the dominant and highly acidic, phosphorylated 48 kDa band, its biochemical properties were compared to other identified highly acidic and phosphorylated shell matrix proteins among the *Crassostrea* species. The 48 kDa band, which was demonstrated in this study to exhibit a settlement inducing activity, was putatively identified as a homolog of folian-cv1 from *Crassostrea virginica* based on their commonly observed biochemical characteristics [12]. Among these characteristics include (1) both were extracted from the EDTA-soluble shell matrix; (2) both are Asp-N rich, which is a common component reported in EDTA-soluble organic matrices [12,13]; (3) band position was also reported at 48 kDa in folian cv-1; (4) Folian cv-1 is moderately glycosylated and highly phosphorylated while the 48 kDa protein in this study did not show any detected glycosylation after Pro-Q Emerald staining or nor any observed band shift after PNGase F deglycosylation treatment (Figure 3.2). But it showed that just like folian cv-1, it was highly phosphorylated as demonstrated by the intense band signal following Pro-Q Diamond phosphoprotein staining as well as Stains-all staining methods (Figures 3.3, 3.4, and Figures A3.3, A3.4, respectively); (5) 48 kDa band in this study was endoprotease-resistant with trypsin, chymotrypsin, Glu-C which suggests that it may have an amino acid composition similar to those reported in folian cv-1; (6) Other shell organic matrix extracted along folian cv-1 showed common epitopes suggesting they are related [12]. Mass spectrometry analysis of the stains-all stainable acidic proteins in this study also shared similar m/z signal chemical signatures suggesting that they may contain similar chemical structures. Hence, a blast search from folian cv-1 and a multiple sequence alignment was performed on these homologs as shown in Appendix Figure A3.8.

Results of this analysis revealed dentin sialophosphoprotein-like in *C. gigas* has a 95.83% homology with Folian cv-1 (E-value = 0.006). Dentin sialophosphoprotein-like is a protein with 331-aa-long residue. Removal of the 18 amino acid signal peptide sequence reveals a mature protein with 313 amino acid residues and a calculated molecular weight of 31,992 Da as well as a theoretical pI of 2.26. It is predicted to be rich in Asp (36%), Ser (43.1%), and Glu (7.7%). It has 136 negatively charged residues (Asp + Glu), and only 1 positively charged residue (Arg + Lys). ScanProsite profile suggests that it is also Asp- and Ser-rich. There

were 149 putative phosphorylation sites and 4 *N*-myristoylation sites predicted. No putative *N*-glycosylation site was found in this protein sequence. This result agrees with the PNGase F deglycosylation result in this study wherein no band shift was observed after this treatment. However, Net*O*Glyc predicted 137 putative *O*-glycosylation sites. Likewise, this predicted amino acid composition also confirms the endoproteinase-resistant results in this study which shows susceptibility for missed or lack of cleavage sites: 1 Arg residue; absence of Lys, Phe, Trp, Leu; 10 Tyr and 24 Glu residues. The latter two residue types might be sandwiched by amino acid residues that could yield a missed cleavage as previously discussed in this study. To date, the functional role of this protein in *C. gigas* is unclear and its sequence was predicted only by automated computational analysis derived from a genomic sequence [14]. Nonetheless, even though CGDSP has high sequence homology to folian cv-1, it contains variations with some amino acid residues, sequence insertions, and deletions that are unique to itself (Supplemental Figure A3.8).

D3. Molecular characterization of Surface protein P12p-like

Peptide Mass Fingerprinting of the 17 kDa band was identified as the product(s) of two isoform genes of Surface protein P12p-like (CGS12P) (accession numbers: XP_034319257.1 and XP_034321529.1) as shown in Table 3.1. In Figure A3.9 A, a sequence alignment comparison of these two variants indicates they have identical amino acid compositions except for one substitution of an amino acid residue at position 7, where a Phe is replaced by Leu. These two isoform genes of Surface P12p-like in *C. gigas* are a 14,982 and 15,016 Da protein containing 128 amino acid residues. The difference between its observed and theoretical molecular mass could be attributed to the additional presence of post-translational modifications in this protein [6]. Removal of the first 20 amino acid peptide signal sequence results in a 108 amino acid protein with a calculated molecular weight of 12,279.91 Da. Its amino acid composition is rich in Asp (15.7%), Asn (13.9%), Gly (8.3%), and Glu (7.4%). It also has 25 negatively charged amino acid residues (Asp + Glu) and 13 positively charged residues. (Arg + Lys) and with a theoretical pI of 4.30. InterProscan search revealed that except for its signal peptide region, the entire sequence length of this protein contains a non-cytoplasmic domain, indicating that it may be located in the extracellular region. Half of the sequence length from the *N*-terminal region (position 23-73) was predicted to be an intrinsically disordered region or low complexity region (LCR). Notably, the LCR region is 28% Asn-, 23% Asp-, 17% Gly- rich and with several post-translational modifications (Figures 6A and 7). Mass spectrometry analysis in this study could identify half of the entire sequence length towards the

C-terminus but was unable to identify the location of the actual *N*-glycosylation modified sites and their corresponding peptide sequence due to the lack of tryptic cleavage sites on the *N*-terminal region. This protein showed a low-intensity band pattern in SDS-PAGE following Pro-Q Emerald (Figure 3.2) and Pro-Q Diamond staining (Figure 3.3) for glycoprotein and phosphoprotein, respectively. These results match with the predicted low-level yet diverse post-translational modifications on this protein. It contains 1 predicted *N*-glycosylation and 6 *O*-glycosylation sites, 2 *N*-myristoylation sites, 4 Casein kinase II phosphorylation sites (at positions 52-55, 53-56, 85-88, 87-90), 3 Protein kinase C phosphorylation site (at positions 78-80, 87-89, 118-120), and 5 other phosphorylation sites.

To compare the CGS12P with its homologs, a multiple sequence alignment was performed. Figure A3.10 shows CGS12P to have some variation and low homology with other proteins specific to *Crassostrea gigas*, but no homology was found in other organisms. The functional role of this gene in these organisms is still unclear. On the other hand, CGS12P was closely related to a pheromone-processing carboxypeptidase KEX1-like in *Crassostrea gigas* by 85% homology identity (E-value = 9×10^{-119}, NCBI Blastn). However, little is known about the functional significance of KEX1 in *C. gigas*. Notably, this is the first report implicating the actual presence of this protein as part of the adult shell matrix in *C. gigas.* In addition, this protein sequence that was archived in the NCBI database was predicted only by automated computational analysis derived from a genomic sequence [14].

Appendix Figures (Chapter III)

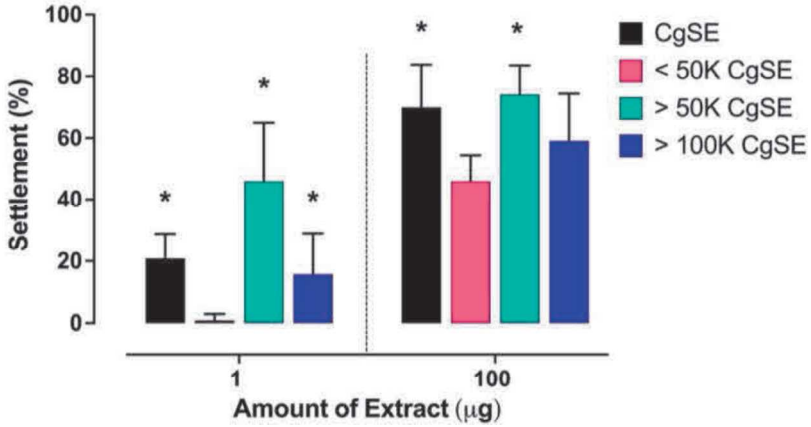

Figure A3.1. Settlement percentages of *C. gigas* larvae on different ultrafiltration fractions with varying amounts of CgSE after 24 h. Asterisks (*) denote significant differences in the amount coated from different fractionated forms of CgSE by ultrafiltration, determined via quasi-binomial glm ($p < 0.05$, n = 6, using different batches of larvae).

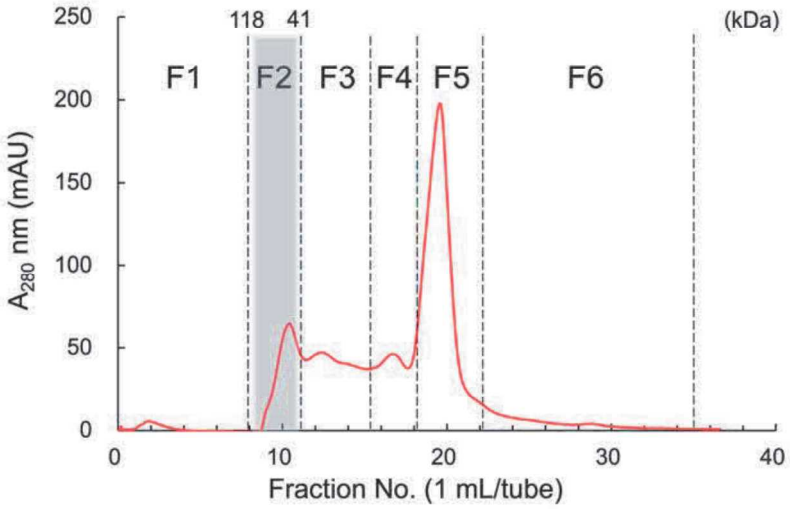

Figure A3.2. Chromatogram profile of fractions eluted from >50K CgSE gel filtration chromatography. Abbreviations: F1 to F6 = eluted pooled fractions.

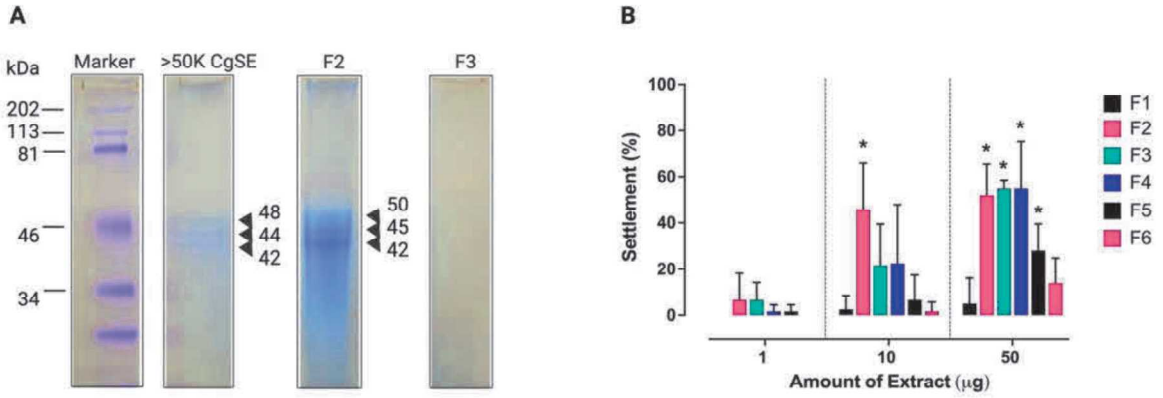

Figure A3.3. SDS-PAGE analysis and larval settlement bioassay after >50K CgSE Gel filtration chromatography. (A) Ten µg each of the active fractions >50K CgSE, F2, and F3 were resolved on a 10% polyacrylamide gel under reducing condition and stained with Stains-all. (B) Settlement Percentages of *C. gigas* larvae on different pooled fractions from >50K CgSE gel filtration chromatography with varying amounts of extract after 24 h. Asterisks (*) denote significantly settlement inducing groups, determined via quasi-binomial glm ($p < 0.05$).

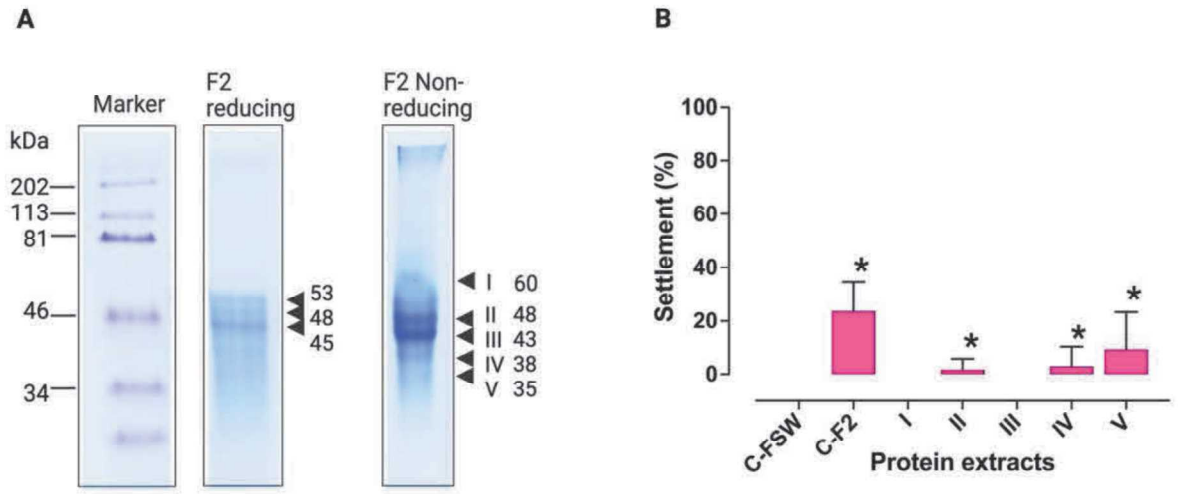

Figure A3.4. SDS-PAGE profile and Larval settlement bioassay of active fraction F2 and its isolated polypeptides. (**A**) Sixty μg of active fraction F2 was treated under reducing and nonreducing conditions and resolved on a 10% polyacrylamide gel. (**B**) Settlement Percentages of *C. gigas* larvae on different isolated polypeptides from active fraction F2 under a nonreducing condition with varying amounts of extract after 24 h. Asterisks (*) denote significant differences in larval response to the inducing cues, using C-FSW as the baseline, determined via quasi-binomial glm ($p < 0.05$, n = 6).

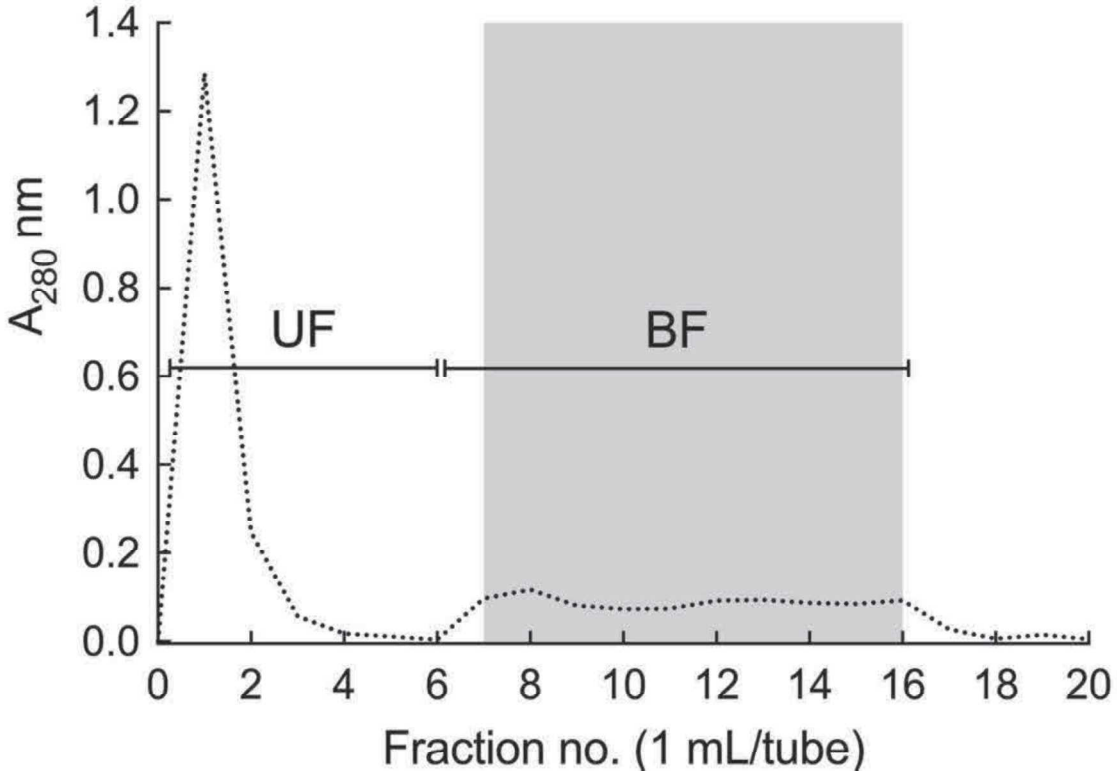

Figure A3.5. Chromatogram profile of fractions eluted after WGA affinity chromatography under 1.0 M NaCl buffer system. Abbreviations: UF = Unbound fraction component; BF = Bound fraction component; BS = Buffer System.

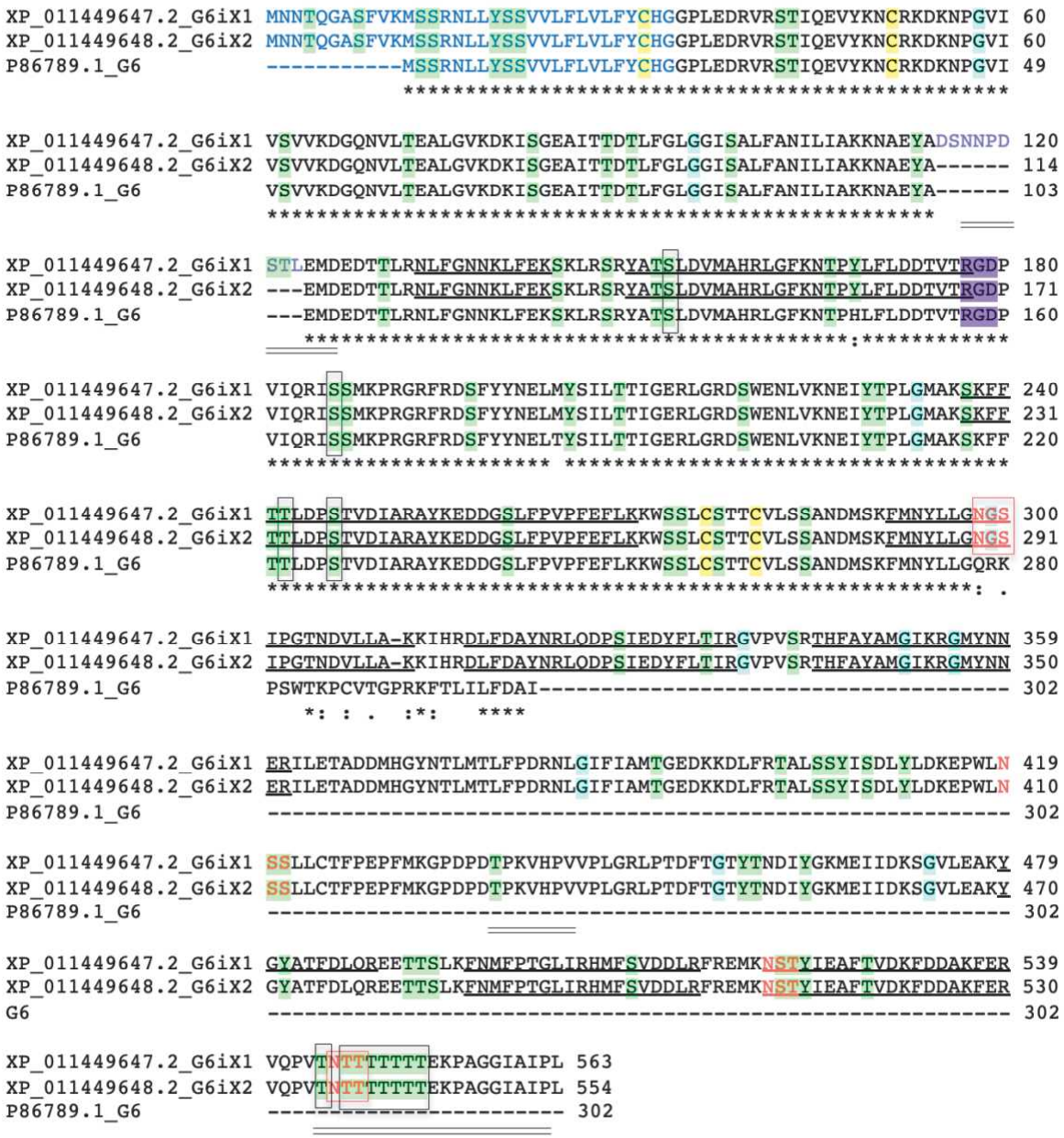

Figure A3.6. Multiple sequence alignment of Gigasin-6 and its isoforms. The predicted signal peptide (aa 1-33) is printed blue. Predicted *N*-glycosylation sites are printed in red while those enclosed in red boxes indicate identified actual *N*-glycosylation sites. Peptides sequenced by MS/MS containing the observed *N*-glycosylation sites are underlined. Predicted *O*-glycosylation sites are enclosed in black boxes. Potential Phosphorylation sites are shaded green while *N*-myristoylation sites are shaded blue. Cysteine-containing residues are shaded yellow while the cell attachment-containing sequence, RGD, is shaded pink. Gigasin 6-isoform X1 shows an extra and unique residue sequence printed in purple. Low complexity regions are doubly underlined and consist of T-rich (37%) residues. Dashes denote blanks or gaps. Asterisks, colons, and dots represent identical amino acids, conserved substitutions, and semi-conserved substitutions, respectively.

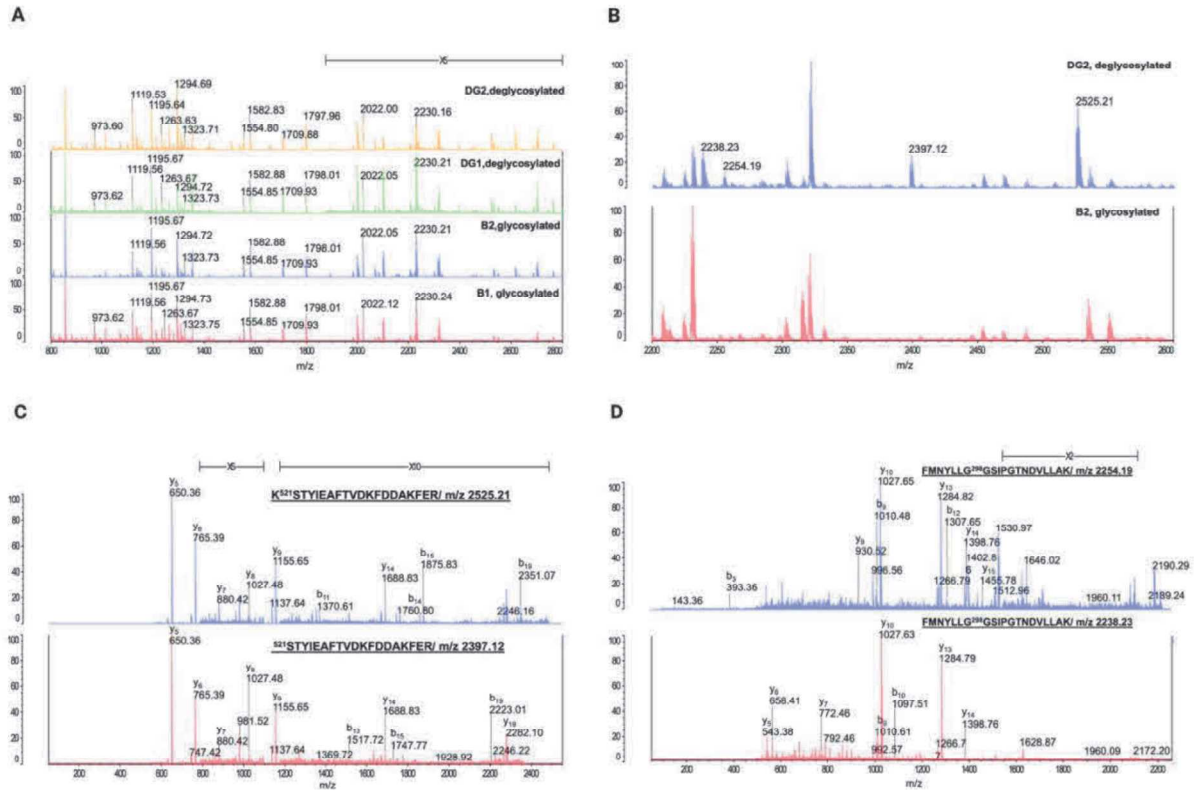

Figure A3.7. Mass spectrometry analysis on the *Crassostrea gigas* Gigasin-6 isoform X1 and/or X2 and its trimer. (**A**) Mass spectrometric profile of Gigasin-6 isoform X1 and/or X2 and its trimer show similar chemical signatures. (**B**) Localization of *N*-glycosylation sites in Gigasin-6 isoform X1 and/or X2 by PNGase F treatment and mass spectrometry analysis. (**C**) Zoomed region on the MS/MS spectra of identified *N*-glycosylation sites annotated by error-tolerant search. (**D**) Zoomed region on the MS/MS spectra of identified *N*-glycosylation sites by MS-Product bioinformatic tool and manual annotation.

```
XP_034310169.1_CGDSP     MKGLAILVFCALIAVGISYPVPDETTDEPDVTDAPDAGDYDVTD---GPDAYDE------       51
QIB98238.1_Foliancv1,2   MKGLAILVFCALIAVGVSYPVPDEADAGDAYEVADTTDVYDVDTDGYGPDAEEEEEGNGD       60
QIB98237.1_Foliancv1,1   MKGLAILVFCALIAVGVSYPVPDEADAGDAYDVADTTDVYDVDTDGYGPDAEEEEEGNGD       60
XP_022314946.1_CVUP,1    MKGLAILVFCALIAVGVSYPVPDEADAGD---VADTTDVYDVDTDGYGPDAEEEEEGNGE       57
XP_022289736.1_CVUP,2    MKGLAILVFCALIAVGVSYPVPDEADAGD---VADTTDVYDVDTDGYGPDAEEEEEGNGE       57
                         ****************:*******:        *  :. ***     **** :*

XP_034310169.1_CGDSP      ---EGDDDGSDGSESSDSSDSSDSSADDDSDDRSDGESGDDESGD--SDDSDGSDGSDSS      106
QIB98238.1_Foliancv1,2    EEGDGEDDENDDSDSSDSDERDDSSADDDSDDRSDDESGDDEDESDDSIDSEGSDGSDDS      120
QIB98237.1_Foliancv1,1    EEGDGEDDENDDSDSSDSDDSDDSSADDDSDDRSDDESGDDEDESDDSSDSEGSDGSDDS      120
XP_022314946.1_CVUP,1     EEGDGEDDGNDDSDSSDSDDSDDSSADDDSDDRSDDESGDDEDESDDSSDSEGSDASDDS      117
XP_022289736.1_CVUP,2     EEGDGEDDGNDDSDSSDSDDSDDSSADDDSDDRSDDESGDDEDESDDSSDSEGSDASDDS      117
                          :*:** .*.*:****.: .*************.******. .  * **:***.**.*

XP_034310169.1_CGDSP      DSSDSSDSSDSSDSSDSSDSSDSSDSDSYEDDSESDSSESDSGDDDSES-----------      155
QIB98238.1_Foliancv1,2    DDSDD------SDDSDDSDDSDDSDDDSYDDDSESDSESSSSDDDDSDSSSASSSSSSS      174
QIB98237.1_Foliancv1,1    DDSDD------SDDSDDSDDSDDSDDDSYDDDSESDSESSSSDDDDSDSSSASSSSSSS      174
XP_022314946.1_CVUP,1     DDSDD------SDDSDDSDDSDDSDDDSYDDDSESDSESSSSDDDDSSSSSSS----SSS      167
XP_022289736.1_CVUP,2     DDSDD------SDDSDDSDDSDDSDDDSYDDDSESDSESSSSDDDDSSSSSSS----SSS      167
                          *.**.      **.**.**.**.**.***:********..*.*.****.*

XP_034310169.1_CGDSP      DSDSDSESYSDSYSDSSDDDDSESSSDSDSYSDSDSSSDSESESESESESESSSSSESGDD      215
QIB98238.1_Foliancv1,2    SSSSESESYSDSYSDEEYSDESD------SYSDSEDSDDSDDSDDD--------------      214
QIB98237.1_Foliancv1,1    SSSSESESYSDSYSDEEYSDESD------SYSDSEDSDDSDDSDDD--------------      214
XP_022314946.1_CVUP,1     SSSSESESDSESYSDEEYSDESD------SSSDSDDSDDSDDSDDS-------DDDDDDD      214
XP_022289736.1_CVUP,2     SSSSESESDSESYSDEEYSDESD------SSSDSDDSDDSDDSDDS-------DD-----      209
                          .*.*:*** *:****.. .*:*:      * ***:.*.**:...:.

XP_034310169.1_CGDSP      DDSDSSDSDSSSSS-SSSSSSSSSDSDSDSSSGSDSDDDSDDDD-SSDSDSDSDSDSDSD      273
QIB98238.1_Foliancv1,2    --DDDDDSDSDSGSDSDSDSDSGSDSDSDSDSGSDSDSDSDDDDDSDDDDSDSDSDSDSD      272
QIB98237.1_Foliancv1,1    --DDDDDSDSDSGSDSDSDSDSGSDSDSDSDSGSDSDSDSDDDDDSDDDDSDSDSDSDSD      272
XP_022314946.1_CVUP,1     DDDDDDDSDSSSGSDSESDSDSDSDSDSDSDSGSDSDSDSDSGSDSDDDDSDSDSDSDSD      274
XP_022289736.1_CVUP,2     --DDDDDSDSSSGSDSESDSDSDSDSDSDSDSGSDSDSDSDSGSDSDDDDSDSDSDSDSD      267
                          .*..****.*.* *.*.*.*.*******.******.***... *.*.***********

XP_034310169.1_CGDSP      SDSYSEEDDDSSDDDDSSDDSYSDEYSDSSDSDSSDSSDSSDNDDSDDDDSSDSSDEN        331
QIB98238.1_Foliancv1,2    SDSDSEEDSDSD---DDDSDDYSEDYSESDSSYSDSESDSD-SDDDDDD-DDDDSDEY        325
QIB98237.1_Foliancv1,1    SDSDSEEDSDSD---DDDSDDYSEDYSESDSSYSDSESDSD-SDDDDDD-DDDDSDEY        325
XP_022314946.1_CVUP,1     SDSNSEEDSNSD---SDDDDSDSEDYSESDGSYSDSESDSD-SDDDDDDDDDDDSDEY        328
XP_022289736.1_CVUP,2     SDSNSEEDSNSD---SDDDDSDSDDYSESDDSYSYSESDSD-SDDDDDD-DDDDSDEY        320
                          *** ****.:*.   ....*. *::**:*..* * ..***. .**.*** ..*.***
```

Figure A3.8. Multiple sequence alignment of NCBI blast searched Folian cv1 homologs. The 48 kDa as the putative CGDSP shows close homology with Folian cv1 variants 1,2. Asterisks, colons, and dots represent identical amino acids, conserved substitutions, and semi-conserved substitutions, respectively. CGDSP, *C. gigas* Dentin sialophosphoprotein-like; Foliancv2, *C. virginica* Folian cv1 variant 2; Foliancv1, *C. virginica* Folian cv1 variant 1; CVUP1, *C. virginica* Uncharacterized Protein LOC111119246; CVUP2, *C. virginica* Uncharacterized Protein LOC111101504.

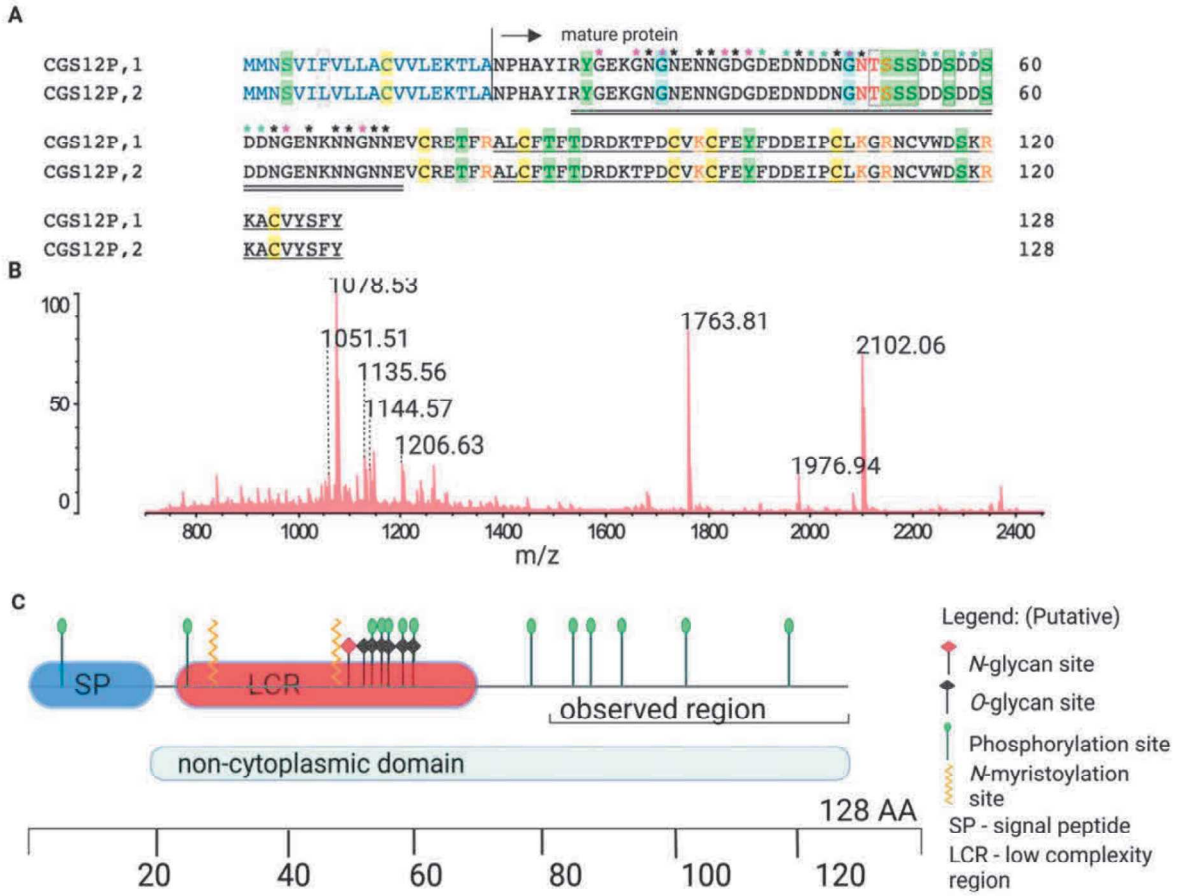

Figure A3.9. Molecular Characterization of *Crassostrea gigas* Surface Protein P12p-like. (**A**) Sequence alignment of Surface Protein P12p-like (CGS12P). Both isoforms contain the same amino acid sequence except for one substituted amino acid position as shown in the enclosed pink box. The Predicted signal peptide (aa 1-20) is in blue. Doubly-lined Low Complexity Region (LCR) domain (aa 23-73) is 28% N-, 23% D-, 17% G- rich and is indicated with * (asterisk) marks. No peptides from the LCR domain which contains *N*- and *O*-glycosylation-rich sites were identified due to a lack of tryptic cleavage sites. Sequences covered by MS spectra are underlined. Tryptic cleavage sites of MS spectra determined sequences are shaded in orange. The predicted *N*-glycosylated sequence is in red while *O*-glycosylation sites are enclosed in black boxes. Predicted phosphorylation sites are shaded green. It also contains several cysteines (yellow-shaded) interspersed along the entire sequence which may be stabilized by disulfide bonds. Predicted *N*-myristoylation sites are shaded in blue. (**B**) Mass spectrometry profile of surface protein P12p-like as represented in the observed region of Figure A10C. (**C**) Graphical representation of Surface protein P12p-like showing protein length, signal peptide (aa 1-20), domains, and post-translation modifications identified by our bioinformatics analysis. The LCR region contains several putative sites of post-translational modifications. Only those peptide fragments close to the C-terminal region were identified by mass spectrometry analysis in this study.

```
XP_034319014.1      MKMQVHVICKSVLHVYQTSSIFFLKTILAGAALFVGCHGYVPPPYPSHSDLTSGSLAAAA 60
XP_034319015.1      ------------------------------------------------------------ 0
XP_034321527.1      ---------------MMNSVILVLLACLVLEKTLANPHAYN-----G------------- 27
XP_034321534.1      ---------------MMNSVILVLLACLVLQKTLANPHAYN-----G------------- 27
XP_034319257.1      ---------------MMNSVIFVLLACVVLEKTLANPHAYI-----R------------- 27
XP_034321529.1      ---------------MMNSVILVLLACVVLEKTLANPHAYI-----R------------- 27


XP_034319014.1      GSIGIVSLAVSAALMSMSSLQNFKDPRKKTTTPEPAESCEDVLAAERQRFEEEKIAEKER 120
XP_034319015.1      --------------MSMSSLQNFKDPRKKTTTPEPAESCEDVLAAERQRFEEEKIAEKER 46
XP_034321527.1      ----------------------YG---EN-DGNENNGDCDVDNGDNENTSSSDDSDDS-- 59
XP_034321534.1      ----------------------YG---EN-DGNENNGDCDVDNGDNENTSSSDDSDDS-- 59
XP_034319257.1      ----------------------YG---EKGNGNENNGDGDEDNDDNGNTSSSDDSDDS-- 60
XP_034321529.1      ----------------------YG---EKGNGNENNGDGDEDNDDNGNTSSSDDSDDS-- 60
                                           :    ::    *   . :     : :  ..:.  :.


XP_034319014.1      FQEEKRDLEARIELSRQIIIGACEEFTAREPIKQPDCIKCFEFLDDQTSCLQA-GSPCAY 179
XP_034319015.1      FQEEKRDLEARIELSRQIIIGACEEFTAREPIKQPDCIKCFEFLDDQTSCLQA-GSPCAY 105
XP_034321527.1      DDREENKEKGNNEVCKNTIRATCDRYTDGD---FEECIECFEYFDDKTQCIQANKRRCLW 116
XP_034321534.1      DDREENKEKGNNEVCKNTIRATCDRYTDGD---FEECIECFEYFDDKTQCIQANKRRCLW 116
XP_034319257.1      DDNGENKNNGNNEVCRETFRALCFTFTDRD--KTPDCVKCFEYFDDEIPCLK--GRNCVW 116
XP_034321529.1      DDNGENKNNGNNEVCRETFRALCFTFTDRD--KTPDCVKCFEYFDDEIPCLK--GRNCVW 116
                     :. :.. :.. *:.:: : . *  :*  :     :*::***::**:  *::     * :


XP_034319014.1      ANFVGVCVYSP---- 190
XP_034319015.1      ANFVGVCVYSP---- 116
XP_034321527.1      DDDRRVCIYDLDIDT 131
XP_034321534.1      DDDRRVCIYDLDIDT 131
XP_034319257.1      DSKRKACVYSFY--- 128
XP_034321529.1      DSKRKACVYSFY--- 128
                     .   .*:*.
```

Figure A3.10. Multiple sequence alignment of *Crassostrea gigas* Surface Protein P12p-like with its homologs. CGS12P shows variation and low homology with other proteins specific to *Crassostrea gigas*, but no homology was found in other organisms. Dashes denote blanks or gaps. Asterisks, colons, and dots represent identical amino acids, conserved substitutions, and semi-conserved substitutions, respectively. Abbreviations: XP_034319014.1, uncharacterized protein LOC117687045 isoform X1 from *Crassostrea gigas*; XP_034319015.1, uncharacterized protein LOC117687045 isoform X2 from *Crassostrea gigas*; XP_034321527.1, pheromone-processing carboxypeptidase KEX1-like from *Crassostrea gigas*; XP_034321534.1, pheromone-processing carboxypeptidase KEX1-like from *Crassostrea gigas;* XP_034319257.1, surface protein P12p-like from *Crassostrea gigas;* XP_034321529.1, surface protein P12p-like from *Crassostrea gigas.*

Appendix Tables (Chapter III)

Table A3.1. Results of Peptide Mass Fingerprinting (PMF) ion search for protein identification of matrix proteins in the *Crassostrea gigas* EDTA-soluble shell extract.

Band No.	Protein ID	NCBI Accession No.	Sequence	m/z	Delta	Miss	Protein Score	Protein Expect
UB1	Gigasin-6 isoform X2	XP_011449648.2	^{124}NLFGNNKLFEK^{134}	1323.73	0.02	1	79	7.5E-04
			^{141}YATSLDVMAHR^{151}	1263.65	0.04	0	79	7.5E-04
			^{141}YATSLDVMAHR^{151} (Oxidation (M))	1279.65	0.04	0	79	7.5E-04
			^{156}NTPYLFLDDTVTR^{168}	1554.81	0.03	0	79	7.5E-04
			^{228}SKFFTTLDPSTVDIAR^{243}	1797.99	0.05	1	79	7.5E-04
			^{230}FFTTLDPSTVDIAR^{243}	1582.84	0.03	0	79	7.5E-04
			^{307}DLFDAYNR^{314}	1013.51	0.04	0	79	7.5E-04
			^{315}LQDPSIEDYFLTIR^{328}	1709.92	0.05	0	79	7.5E-04
			^{335}THFAYAMGIK^{344}	1138.60	0.03	0	79	7.5E-04
			^{335}THFAYAMGIKR^{345}	1294.71	0.04	1	79	7.5E-04
			^{335}THFAYAMGIKR^{345} (Oxidation (M))	1310.70	0.03	1	79	7.5E-04
			^{346}GMYNNER^{352}	883.40	0.03	0	79	7.5E-04
			^{470}YGYATFDLQR^{479}	1233.62	0.03	0	79	7.5E-04
			^{487}FNMFPTGLIR^{496}	1195.67	0.04	0	79	7.5E-04
			^{487}FNMFPTGLIR^{496} (Oxidation (M))	1211.66	0.04	0	79	7.5E-04
	Gigasin-6 isoform X1	XP_011449647.2	^{133}NLFGNNKLFEK^{143}	1323.73	0.02	1	78	1.2E-03
			^{150}YATSLDVMAHR^{160}	1263.65	0.04	0	78	1.2E-03
			^{150}YATSLDVMAHR^{160} (Oxidation (M))	1279.65	0.04	0	78	1.2E-03
			^{165}NTPYLFLDDTVTR^{177}	1554.81	0.03	0	78	1.2E-03
			^{237}SKFFTTLDPSTVDIAR^{252}	1797.99	0.05	1	78	1.2E-03
			^{239}FFTTLDPSTVDIAR^{252}	1582.84	0.03	0	78	1.2E-03
			^{316}DLFDAYNR^{323}	1013.51	0.04	0	78	1.2E-03
			^{324}LQDPSIEDYFLTIR^{337}	1709.92	0.05	0	78	1.2E-03
			^{344}THFAYAMGIK^{353}	1138.60	0.03	0	78	1.2E-03
			^{344}THFAYAMGIKR^{354}	1294.71	0.04	1	78	1.2E-03
			^{344}THFAYAMGIKR^{354} (Oxidation (M))	1310.70	0.03	1	78	1.2E-03
			^{355}GMYNNER^{361}	883.40	0.03	0	78	1.2E-03

			^{479}YGYATFDLQR^{488}	1233.62	0.03	0	78	1.2E-03
			^{496}FNMFPTGLIR^{505}	1195.67	0.04	0	78	1.2E-03
			^{496}FNMFPTGLIR^{505} (Oxidation (M))	1211.66	0.04	0	78	1.2E-03
UB2	Gigasin-6 isoform X2	XP_011449648.2	^{124}NLFGNNKLFEK^{134}	1323.73	0.02	1	85	2.1E-04
			^{152}LGFKNTPYLFLDDTVTR^{168}	2000.12	0.07	1	85	2.1E-04
			^{156}NTPYLFLDDTVTR^{168}	1554.81	0.03	0	85	2.1E-04
			^{228}SKFFTTLDPSTVDIAR^{243}	1798.01	0.07	1	85	2.1E-04
			^{230}FFTTLDPSTVDIAR^{243}	1582.84	0.03	0	85	2.1E-04
			^{244}AYKEDDGSLFPVPFEFLK^{261}	2102.11	0.06	1	85	2.1E-04
			^{244}AYKEDDGSLFPVPFEFLKK^{262}	2230.21	0.07	2	85	2.1E-04
			^{315}LQDPSIEDYFLTIR^{328}	1709.92	0.05	0	85	2.1E-04
			^{335}THFAYAMGIK^{344}	1138.61	0.04	0	85	2.1E-04
			^{335}THFAYAMGIKR^{345}	1294.71	0.04	1	85	2.1E-04
			^{335}THFAYAMGIKR^{345} (Oxidation (M))	1310.72	0.05	1	85	2.1E-04
			^{470}YGYATFDLQR^{479}	1233.62	0.03	0	85	2.1E-04
			^{487}FNMFPTGLIR^{496}	1195.67	0.04	0	85	2.1E-04
			^{487}FNMFPTGLIR^{496} (Oxidation (M))	1211.65	0.03	0	85	2.1E-04
			^{497}HMFSVDDLR^{505}	1119.56	0.03	0	85	2.1E-04
	Gigasin-6 isoform X1	XP_011449647.2	^{133}NLFGNNKLFEK^{143}	1323.73	0.02	1	84	2.8E-04
			^{161}LGFKNTPYLFLDDTVTR^{177}	2000.12	0.07	1	84	2.8E-04
			^{165}NTPYLFLDDTVTR^{177}	1554.81	0.03	0	84	2.8E-04
			^{237}SKFFTTLDPSTVDIAR^{252}	1798.01	0.07	1	84	2.8E-04
			^{239}FFTTLDPSTVDIAR^{252}	1582.84	0.03	0	84	2.8E-04
			^{253}AYKEDDGSLFPVPFEFLK^{270}	2102.11	0.06	1	84	2.8E-04
			^{253}AYKEDDGSLFPVPFEFLKK^{271}	2230.21	0.07	2	84	2.8E-04
			^{324}LQDPSIEDYFLTIR^{337}	1709.92	0.05	0	84	2.8E-04
			^{344}THFAYAMGIK^{353}	1138.61	0.04	0	84	2.8E-04
			^{344}THFAYAMGIKR^{354}	1294.71	0.04	1	84	2.8E-04
			^{344}THFAYAMGIKR^{354} (Oxidation (M))	1310.72	0.05	1	84	2.8E-04
			^{479}YGYATFDLQR^{488}	1233.62	0.03	0	84	2.8E-04
			^{496}FNMFPTGLIR^{505}	1195.67	0.04	0	84	2.8E-04
			^{496}FNMFPTGLIR^{505} (Oxidation (M))	1211.65	0.03	0	84	2.8E-04
			^{506}HMFSVDDLR^{514}	1119.56	0.03	0	84	2.8E-04

UB3	Surface protein P12p-like	XP_034321529.1	^{81}ALCFTFTDR^{89}	1144.57	0.02	0	80	5.9E-04
			^{81}ALCFTFTDRDKTPDCVK^{97}	2102.06	0.05	2	80	5.9E-04
			^{98}CFEYFDDEIPCLK^{110}	1763.81	0.04	0	80	5.9E-04
			^{98}CFEYFDDEIPCLKGR^{112}	1976.94	0.05	1	80	5.9E-04
			^{111}GRNCVWDSK^{119}	1135.56	0.03	1	80	5.9E-04
			^{113}NCVWDSKR^{120}	1078.53	0.02	1	80	5.9E-04
			^{113}NCVWDSKRK^{121}	1206.63	0.03	2	80	5.9E-04
			^{121}KACVYSFY^{128}	1051.51	0.02	1	80	5.9E-04
	Surface protein P12p-like	XP_034319527.1	Same as “XP_034321529.1”					
B1	Gigasin-6 isoform X1	XP_011449647.2	^{133}NLFGNNKLFEK^{143}	1323.75	0.04	1	97	1.3E-005
			^{150}YATSLDVMAHR^{160}	1263.67	0.05	0	97	1.3E-005
			^{150}YATSLDVMAHR^{160} (Oxidation (M))	1279.65	0.04	0	97	1.3E-005
			^{165}NTPYLFLDDTVTR^{177}	1554.85	0.07	0	97	1.3E-005
			^{237}SKFFTTLDPSTVDIAR^{252}	1798.01	0.07	1	97	1.3E-005
			^{239}FFTTLDPSTVDIAR^{252}	1582.88	0.07	0	97	1.3E-005
			^{316}DLFDAYNR^{323}	1013.51	0.04	0	97	1.3E-005
			^{324}LQDPSIEDYFLTIR^{337}	1709.93	0.05	0	97	1.3E-005
			^{344}THFAYAMGIK^{353}	1138.61	0.04	0	97	1.3E-005
			^{344}THFAYAMGIKR^{354}	1294.73	0.06	1	97	1.3E-005
			^{344}THFAYAMGIKR^{354} (Oxidation (M))	1310.72	0.05	1	97	1.3E-005
			^{441}VHPVVPLGR^{449}	973.62	0.03	0	97	1.3E-005
			^{479}YGYATFDLQR^{488}	1233.63	0.04	0	97	1.3E-005
			^{496}FNMFPTGLIR^{505}	1195.67	0.04	0	97	1.3E-005
			^{496}FNMFPTGLIR^{505} (Oxidation (M))	1211.67	0.05	0	97	1.3E-005
			^{506}HMFSVDDLR^{514}	1119.56	0.03	0	97	1.3E-005
			^{506}HMFSVDDLR^{514} (Oxidation (M))	1135.56	0.04	0	97	1.3E-005
	Gigasin-6 isoform X2	XP_011449648.2	^{124}NLFGNNKLFEK^{134}	1323.75	0.04	1	96	1.5E-005
			^{141}YATSLDVMAHR^{151}	1263.67	0.05	0	96	1.5E-005
			^{141}YATSLDVMAHR^{151} (Oxidation (M))	1279.65	0.04	0	96	1.5E-005
			^{156}NTPYLFLDDTVTR^{168}	1554.85	0.07	0	96	1.5E-005
			^{228}SKFFTTLDPSTVDIAR^{243}	1798.01	0.07	1	96	1.5E-005
			^{230}FFTTLDPSTVDIAR^{243}	1582.84	0.07	0	96	1.5E-005
			^{307}DLFDAYNR^{314}	1013.51	0.04	0	96	1.5E-005

			^{315}LQDPSIEDYFLTIR^{328}	1709.93	0.05	0	96	1.5E-005
			^{335}THFAYAMGIK^{344}	1138.61	0.04	0	96	1.5E-005
			^{335}THFAYAMGIKR^{345}	1294.73	0.06	1	96	1.5E-005
			^{335}THFAYAMGIKR^{345} (Oxidation (M))	1310.72	0.05	1	96	1.5E-005
			^{432}VHPVVPLGR^{440}	973.62	0.03	0	96	1.5E-005
			^{470}YGYATFDLQR^{479}	1233.63	0.04	0	96	1.5E-005
			^{487}FNMFPTGLIR^{496}	1195.67	0.04	0	96	1.5E-005
			^{487}FNMFPTGLIR^{496} (Oxidation (M))	1211.66	0.05	0	96	1.5E-005
			^{497}HMFSVDDLR^{505}	1119.56	0.03	0	96	1.5E-005
			^{497}HMFSVDDLR^{505} (Oxidation (M))	1135.56	0.04	0	96	1.5E-005
B2	Gigasin-6 isoform X1	XP_011449647.2	^{133}NLFGNNKLFEK^{143}	1323.73	0.02	1	63	3.1E-02
			^{165}NTPYLFLDDTVTR^{177}	1554.85	0.07	0	63	3.1E-02
			^{237}SKFFTTLDPSTVDIAR^{252}	1798.01	0.07	1	63	3.1E-02
			^{239}FFTTLDPSTVDIAR^{252}	1582.88	0.07	0	63	3.1E-02
			^{316}DLFDAYNR^{323}	1013.51	0.04	0	63	3.1E-02
			^{324}LQDPSIEDYFLTIR^{337}	1709.93	0.05	0	63	3.1E-02
			^{344}THFAYAMGIK^{353}	1138.61	0.04	0	63	3.1E-02
			^{344}THFAYAMGIKR^{354}	1294.72	0.05	1	63	3.1E-02
			^{344}THFAYAMGIKR^{354} (Oxidation (M))	1310.72	0.05	1	63	3.1E-02
			^{479}YGYATFDLQR^{488}	1233.63	0.04	0	63	3.1E-02
			^{496}FNMFPTGLIR^{505}	1195.67	0.04	0	63	3.1E-02
			^{496}FNMFPTGLIR^{505} (Oxidation (M))	1211.66	0.04	0	63	3.1E-02
			^{506}HMFSVDDLR^{514}	1119.56	0.03	0	63	3.1E-02
	Gigasin-6 isoform X2	XP_011449648.2	^{124}NLFGNNKLFEK^{134}	1323.73	0.02	1	63	3.1E-02
			^{156}NTPYLFLDDTVTR^{168}	1554.85	0.07	0	63	3.1E-02
			^{228}SKFFTTLDPSTVDIAR^{243}	1798.01	0.07	1	63	3.1E-02
			^{230}FFTTLDPSTVDIAR^{243}	1582.88	0.07	0	63	3.1E-02
			^{307}DLFDAYNR^{314}	1013.51	0.04	0	63	3.1E-02
			^{315}LQDPSIEDYFLTIR^{328}	1709.93	0.05	0	63	3.1E-02
			^{335}THFAYAMGIK^{344}	1138.61	0.04	0	63	3.1E-02
			^{335}THFAYAMGIKR^{345}	1294.72	0.05	1	63	3.1E-02
			^{335}THFAYAMGIKR^{345} (Oxidation (M))	1310.72	0.05	1	63	3.1E-02

			^{470}YGYATFDLQR^{479}	1233.63	0.04	0	63	3.1E-02
			^{487}FNMFPTGLIR^{496}	1195.67	0.04	0	63	3.1E-02
			^{487}FNMFPTGLIR^{496} (Oxidation (M))	1211.66	0.04	0	63	3.1E-02
			^{497}HMFSVDDLR^{505}	1119.56	0.03	0	63	3.1E-02
DG1	Gigasin-6 isoform X1	XP_011449647.2	^{133}NLFGNNKLFEK^{143}	1323.73	0.02	1	73	3.4E-03
			^{150}YATSLDVMAHR^{160}	1263.67	0.07	0	73	3.4E-03
			^{165}NTPYLFLDDTVTR^{177}	1554.85	0.07	0	73	3.4E-03
			^{237}SKFFTTLDPSTVDIAR^{252}	1798.01	0.07	1	73	3.4E-03
			^{239}FFTTLDPSTVDIAR^{252}	1582.88	0.07	0	73	3.4E-03
			^{316}DLFDAYNR^{323}	1013.51	0.04	0	73	3.4E-03
			^{324}LQDPSIEDYFLTIR^{337}	1709.93	0.05	0	73	3.4E-03
			^{344}THFAYAMGIKR^{354}	1294.72	0.05	1	73	3.4E-03
			^{344}THFAYAMGIKR^{354} (Oxidation (M))	1310.72	0.05	1	73	3.4E-03
			^{441}VHPVVPLGR^{449}	973.62	0.03	0	73	3.4E-03
			^{479}YGYATFDLQR^{488}	1233.63	0.04	0	73	3.4E-03
			^{496}FNMFPTGLIR^{505}	1195.67	0.04	0	73	3.4E-03
			^{496}FNMFPTGLIR^{505} (Oxidation (M))	1211.68	0.06	0	73	3.4E-03
			^{506}HMFSVDDLR^{514}	1119.56	0.03	0	73	3.4E-03
	Gigasin-6 isoform X2	XP_011449648.2	^{124}NLFGNNKLFEK^{134}	1323.73	0.02	1	72	4.3E-03
			^{141}YATSLDVMAHR^{151}	1263.68	0.07	0	72	4.3E-03
			^{156}NTPYLFLDDTVTR^{168}	1554.85	0.07	0	72	4.3E-03
			^{228}SKFFTTLDPSTVDIAR^{243}	1798.01	0.07	1	72	4.3E-03
			^{230}FFTTLDPSTVDIAR^{243}	1582.84	0.07	0	72	4.3E-03
			^{307}DLFDAYNR^{314}	1013.51	0.04	0	72	4.3E-03
			^{315}LQDPSIEDYFLTIR^{328}	1709.93	0.05	0	72	4.3E-03
			^{335}THFAYAMGIKR^{345}	1294.72	0.05	1	72	4.3E-03
			^{335}THFAYAMGIKR^{345} (Oxidation (M))	1310.72	0.05	1	72	4.3E-03
			^{432}VHPVVPLGR^{440}	973.62	0.03	0	72	4.3E-03
			^{470}YGYATFDLQR^{479}	1233.63	0.04	0	72	4.3E-03
			^{487}FNMFPTGLIR^{496}	1195.67	0.04	0	72	4.3E-03
			^{487}FNMFPTGLIR^{496} (Oxidation (M))	1211.66	0.05	0	72	4.3E-03
			^{497}HMFSVDDLR^{505}	1119.56	0.03	0	72	4.3E-03
DG2	Gigasin-6 isoform X1	XP_011449647.2	^{133}NLFGNNKLFEK^{143}	1323.71	0.00	1	91	5E-005

		^{150}YATSLDVMAHR^{160}	1263.63	0.02	0	91	5E-005
		^{165}NTPYLFLDDTVTR^{177}	1554.80	0.02	0	91	5E-005
		^{237}SKFFTTLDPSTVDIAR^{252}	1797.96	0.02	1	91	5E-005
		^{239}FFTTLDPSTVDIAR^{252}	1582.83	0.02	0	91	5E-005
		^{316}DLFDAYNR^{323}	1013.48	0.01	0	91	5E-005
		^{324}LQDPSIEDYFLTIR^{337}	1709.88	0.01	0	91	5E-005
		^{344}THFAYAMGIK^{353}	1138.57	-0.0	0	91	5E-005
		^{344}THFAYAMGIKR^{354}	1294.69	0.02	1	91	5E-005
		^{344}THFAYAMGIKR^{354} (Oxidation (M))	1310.68	0.01	1	91	5E-005
		^{441}VHPVVPLGR^{449}	973.60	0.01	0	91	5E-005
		^{479}YGYATFDLQR^{488}	1233.59	0.00	0	91	5E-005
		^{496}FNMFPTGLIR^{505}	1195.64	0.01	0	91	5E-005
		^{496}FNMFPTGLIR^{505} (Oxidation (M))	1211.64	0.02	0	91	5E-005
		^{506}HMFSVDDLR^{514}	1119.53	0.00	0	91	5E-005
		^{506}HMFSVDDLR^{514} (Oxidation (M))	1135.52	-0.0	0	91	5E-005
Gigasin-6 isoform X2	XP_011449648.2	^{124}NLFGNNKLFEK^{134}	1323.71	0.00	1	90	5.9E-005
		^{141}YATSLDVMAHR^{151}	1263.63	0.02	0	90	5.9E-005
		^{156}NTPYLFLDDTVTR^{168}	1554.80	0.02	0	90	5.9E-005
		^{228}SKFFTTLDPSTVDIAR^{243}	1797.96	0.02	1	90	5.9E-005
		^{230}FFTTLDPSTVDIAR243	1582.83	0.02	0	90	5.9E-005
		^{307}DLFDAYNR^{314}	1013.48	0.01	0	90	5.9E-005
		^{315}LQDPSIEDYFLTIR^{328}	1709.88	0.01	0	90	5.9E-005
		^{335}THFAYAMGIK^{344}	1138.57	-0.0	0	90	5.9E-005
		^{335}THFAYAMGIKR^{345}	1294.69	0.02	1	90	5.9E-005
		^{335}THFAYAMGIKR^{345} (Oxidation (M))	1310.68	0.01	1	90	5.9E-005
		^{432}VHPVVPLGR^{440}	973.60	0.01	0	90	5.9E-005
		^{470}YGYATFDLQR^{479}	1233.59	0.00	0	90	5.9E-005
		^{487}FNMFPTGLIR^{496}	1195.64	0.01	0	90	5.9E-005
		^{487}FNMFPTGLIR^{496} (Oxidation (M))	1211.64	0.02	0	90	5.9E-005
		^{497}HMFSVDDLR^{505}	1119.53	0.00	0	90	5.9E-005
		^{497}HMFSVDDLR^{505} (Oxidation (M))	1135.52	-0.0	0	90	5.9E-005

Table A3.2. MS/MS-based identification and characterization of the protein band containing the determined *N*-glycosylation sites of Gigasin-6 isoform X1 and/or X2 in Pacific oyster *Crassostrea gigas*.

Band No.	N-glycosylation Site	Peptide Sequence		Peptide m/z	*y* series ion m/z	*b* series ion m/z	Delta	Miss	Ion Score
Annotation by Error Tolerant Search									
DG2	N521	K^{521}STYIEAFTVDKFDDAKFER [+129.04 at E6]	+	2525.21	650.36	1370.611	0.0269	2	32
					765.39	1760.80			
					880.42	1875.83			
					1027.48	2351.07			
					1155.58				
					1688.83				
DG2	N521	^{521}STYIEAFTVDKFDDAKFER [+0.98 at N-term N]	+	2397.12	650.36	1517.72	-0.0044	2	45
					765.39	1747.77			
					880.42	2223.01			
					1027.48				
					1155.58				
					1688.83				
					2282.10				
Manual annotation by matching observed MS/MS ion fragments with MS-Product Program									
DG2	N298	FMNYLLG^{298}GSIPGTNDVLLAK		2254.19	930.52	393.36	N/A	N/A	N/A
					1027.65	1010.48			
					1284.82	1307.65			
					1398.76				
					1455.78				
DG2	N298	FMNYLLG^{298}GSIPGTNDVLLAK		2238.23	543.38	1010.61	N/A	N/A	N/A
					658.41	1097.51			
					772.46				
					1027.63				
					1284.79				
					1398.76				

Appendix Figures (Chapter IV)

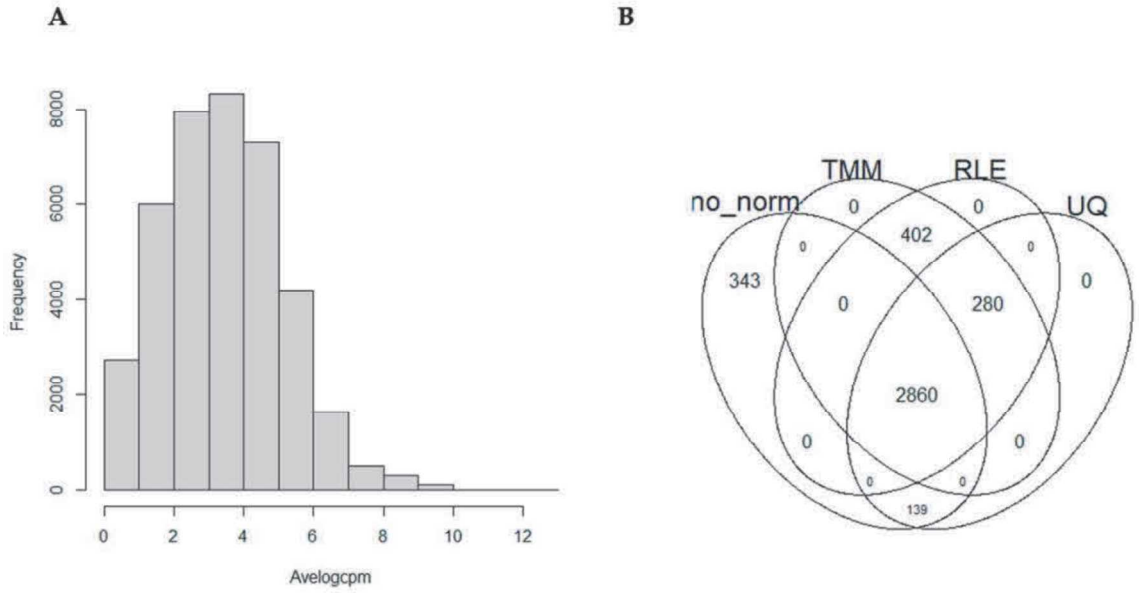

Figure A4.1. Histogram profile of the average log count-per-million (CPM) of the mapped transcript sequences (A) and the Venn diagram of transcripts with differential expression using various normalization models (B). Abbreviations: no_norm = no normalization; TMM = Trimmed Mean of M-values; RLE = Relative Log Expression; UQ = Upper Quartile.

Table A4.1. Full list of transcripts used for differential expression analysis in the Pacific oyster *Crassostrea gigas* transcriptome.

The table is provided as a separate **Table A4.1.xsls** online document that can be downloaded at https://www.mdpi.com/article/10.3390/d14070559/s1.

Appendix Tables (Chapter IV)

Table A4.2. Results of the Blastx homology search for the thirty differential transcript expression in the *Crassostrea gigas* transcriptome.

C. gigas Accession No. (Subject)	*C. gigas* Annotation	Homolog Protein ID (Query)	Species	Homolog Accession No.	Identities (Query/ Subject)	Homology (%)	BlastX E-value	Protein Length (Query)
LOC105335414	Abscisic acid and environmental stress-inducible protein	Abscisic acid and environmental stress-inducible protein (AESIP)	*Crassostrea gigas*	XP_011437564.2	84/84	100	0	126
LOC105341000	Coadhesin	Coadhesin	*Crassostrea gigas*	XP_034329440.1	831/831	100	0	831
LOC105324992	SCO-spondin	SCO-spondin	*Crassostrea gigas*	XP_034330544.1	3201/3201	100	0	3201
LOC105338148	Uncharacterized LOC105338148, Isoform X1	General transcriptional corepressor trfA-like,1 (trfA-like,1)	*Mercenaria mercenaria*	XP_045201952.1	96/290	33	3E-29	796
LOC105347520	A disintegrin and metalloproteinase with thrombospondin motifs adt-1	A disintegrin and metalloproteinase with thrombospondin motifs adt-1 (ADAM-1)	*Crassostrea gigas*	XP_034329445.1	595/595	100	0	595
LOC105324594	Uncharacterized LOC105324594	Calcium-dependent protein kinase 7 (CPK7)	*Crassostrea gigas*	XP_011427257.2	70/136	51	2E-37	133
LOC117688777	Spidroin-1-like	Spidroin-1-like,1	*Crassostrea gigas*	XP_034323067.1	46/46	100	2E-08	308
LOC105348066	Uncharacterized LOC105348066	Stereocilin	*Mizuhopecten yessoensis*	OWF40357.1	483/1126	43	0	1412
LOC105328424	Uncharacterized LOC105328424	IgGFc-binding protein	*Crassostrea gigas*	XP_011422608.2	267/678	39	6E-135	708
LOC105340987	Uncharacterized LOC105340987, isoform X1	SCO-spondin-like,1	*Crassostrea virginica*	XP_022339848.1	808/866	93	0	1957

LOC105340987	Uncharacterized LOC105340987, isoform X2	SCO-spondin-like,2	*Crassostrea virginica*	XP_022339848.1	808/866	93	0	1957
LOC109618369	Uncharacterized LOC109618369	General transcriptional corepressor trfA-like,2 (trfA-like,1)	*Mercenaria mercenaria*	XP_045201952.1	97/290	33	9E-31	796
LOC117688778	Spidroin-1-like	Spidroin-1-like,2	*Crassostrea gigas*	XP_034323070.1	63/63	100	7E-13	238
LOC117688748	Uncharacterized LOC117688748	Uncharacterized LOC117688748	*Crassostrea gigas*	XP_034322956.1	292/292	100	0	292
LOC117688749	Uncharacterized LOC117688749	Uncharacterized LOC117688749	*Crassostrea gigas*	XP_034322955.1	292/292	100	0	292
LOC105327237	Uncharacterized LOC105327237	SCO-spondin-like,3	*Crassostrea virginica*	XP_022339848.1	552/806	68	0	1957
LOC117684686	ADP-ribosylation factor 1-like	ADP-ribosylation factor 1-like	*Crassostrea gigas*	XP_034313686.1	189/189	100	3E-136	189
LOC105327534	Mediator of DNA damage checkpoint protein 1	Mediator of DNA damage checkpoint protein 1 (MDC1)	*Crassostrea gigas*	XP_011426386.2	590/590	100	0	590
LOC105344419	Glycine-rich cell wall structural protein 1.8 transcript var X2	Glycine-rich cell wall structural protein 1.8 transcript isoform X2 (GRP 1.8-X2)	*Crassostrea gigas*	XP_034323479.1	141/141	100	0	141
LOC105344419	Glycine-rich cell wall structural protein 1.8 transcript var X1	Glycine-rich cell wall structural protein 1.8 transcript isoform X1 (GRP 1.8-X1)	*Crassostrea gigas*	XP_011450501.3	155/155	100	0	155
LOC105329902	Uncharacterized LOC105329902	Tyrp-1	*Hyriopsis cumingii*	APC92582.1	282/731	39	2E-157	778
LOC105334609	Collagen alpha-2(I) chain	Collagen alpha-2(I) chain	*Crassostrea gigas*	XP_011436430.2	173/173	100	6E-37	173
LOC105324991	IgGFc-binding protein	IgGFc-binding protein	*Crassostrea gigas*	XP_011422608.2	687/687	100	0	708

LOC105345892	Leucine-rich repeat-containing protein 4C	Leucine-rich repeat-containing protein 4C (LRRC4C)	*Crassostrea gigas*	XP_011452538.2	391/391	100	0	391
LOC105322623	Uncharacterized LOC105322623	Uncharacterized LOC105322623	*Crassostrea gigas*	XP_011419747.2	400/400	100	0	400
LOC105330463	Spidroin-1-like	Spidroin-1-like,3	*Crassostrea gigas*	XP_034323071.1	30/30	100	3E-08	236
LOC105345888	Uncharacterized LOC105345888	CD283	*Mytilus edulis*	CAG2230452.1	131/372	35	3E-59	392
LOC105319426	Acanthoscurrin-2	Acanthoscurrin-2	*Crassostrea gigas*	XP_011415271.2	161/161	100	0	161
LOC105345889	Leucine-rich repeat LGI family member 2	Leucine-rich repeat LGI family member 2 (LGI2)	*Crassostrea gigas*	XP_011452536.2	195/195	100	0	195
LOC117688175	Uncharacterized LOC117688175	Serine/arginine repetitive matrix protein 1-like (SRRM1-like)	*Mizuhopecten yessoensis*	XP_021357253.1	108/218	50	9E-62	760

Table A4.3. Results of the Blastx homology search for the differential transcript expression involved in the Gene Ontology enriched processes of the *Crassostrea gigas* transcriptome.

C. gigas **Accession No. (Subject)**	**Protein ID (Query)**	**Species**	**Homolog Accession No.**	**Identities (Query/Subject)**	**Homology (%)**	**BlastX E-value**	**Protein Length (Query)**
Chitin binding (GO:0008061)							
LOC105326593	cgPif97	*Crassostrea gigas*	AFT63504.1	881/1024	86	0	438
LOC105317660	Peritrophin-44	*Crassostrea gigas*	XP_011412673.1	288/288	100	3E-98	138
LOC105346858	Lactadherin, 1	*Mizuhopecten yessoensis*	OWF43623.1	40/141	28	1E-06	565
LOC105346858	Lactadherin, 2	*Mizuhopecten yessoensis*	OWF43623.1	40/141	28	1E-06	565
LOC105331104	Collagen alpha 1(XII) chain	*Mytilus coruscus*	CAC5362153.1	373/693	54	0	797
LOC105336951	Peritrophin-1-like	*Mizuhopecten yessoensis*	XP_021368929.1	101/233	43	4E-53	239
LOC105329937	Peritrophin-1-like	*Gigantopelta aegis*	XP_041349873.1	67/144	47	9E-35	229
LOC105332576	Cleavage and polyadenylation factor I subunit 1	*Crassostrea gigas*	NP_001292250.1	472/472	100	0	472
LOC105332577	Putative chitinase 1 precursor	*Crassostrea gigas*	NP_001292252.1	475/475	100	0	475
LOC105332575	Putative chitinase	*Crassostrea gigas*	NP_001295799.1	471/471	100	0	471
Calcium ion binding (GO:0005509)							
LOC105323227	DNA ligase 1 Isoform X1	*Crassostrea gigas*	XP_011420555.2	843/843	100	0	843
	DNA ligase 1 Isoform X2	*Crassostrea gigas*	XP_011420556.2	843/843	100	0	843
	DNA ligase 1 Isoform X3	*Crassostrea gigas*	XP_011420558.2	811/811	100	0	811

	DNA ligase 1 Isoform X4	*Crassostrea gigas*	XP_011420559	811/811	100	0	811
	DNA ligase 1 Isoform X5	*Crassostrea gigas*	XP_011420561	795/795	100	0	795
	DNA ligase 1 Isoform X6	*Crassostrea gigas*	XP_034305628	766/766	100	0	766
LOC105345764	EF-hand domain-containing protein 1	*Crassostrea gigas*	XP_011452353.1	641/641	100	0	641
LOC105331896	Calmodulin-4 transcript var X1	*Crassostrea gigas*	XP_011432565.2	178/178	100	4E-122	178
	Calmodulin-4 transcript var X2	*Crassostrea gigas*	XP_011432571.2	176/176	100	1E-122	176
	Calmodulin-4 transcript var X3	*Crassostrea gigas*	XP_011432579.2	175/175	100	8E-122	175
LOC105337218	Calcium-binding protein E63-1-like	*Haliotis rubra*	XP_046557623.1	26/70	37	0.002	144
LOC105328000	Calmodulin-A isoform X1	*Crassostrea gigas*	XP_011427010.1	159/159	100	4E-95	210
	Calmodulin-A isoform X2	*Crassostrea gigas*	XP_011427011.1	153/153	100	4E-90	204
LOC105320395	Leucine-rich repeat-containing protein 74B isoform X1	*Crassostrea gigas*	XP_011416624.1	601/601	100	0	601
	Leucine-rich repeat-containing protein 74B isoform X2	*Crassostrea gigas*	XP_011416625.1	601/601	100	0	601
	Leucine-rich repeat-containing protein 74B isoform X3	*Crassostrea gigas*	XP_034323188.1	601/601	100	0	601
	Leucine-rich repeat-containing protein 74B isoform X4	*Crassostrea gigas*	XP_011416626.1	574/574	100	0	574
	Leucine-rich repeat-containing protein 74B isoform X5	*Crassostrea gigas*	XP_011416627.1	563/563	100	0	563
LOC105332695	Sarcoplasmic calcium-binding protein	*Crassostrea gigas*	XP_011433684.1	179/179	100	0	179
LOC105338472	Myosin	*Crassostrea gigas*	XP_011441922.1	162/162	100	0	162
LOC105330202	EGF-like repeat and discoidin I-like domain-containing protein 3	*Mercenaria mercenaria*	XP_045201990.1	150/321	65	9E-89	322
LOC105341555	Mammalian ependymin-related protein 1	*Crassostrea gigas*	XP_011446436.2	205/205	100	0	205

LOC105327335	EF-hand domain-containing family member B	*Crassostrea gigas*	XP_011426045.1	602/602	100	0	602
LOC105327434	Neurocalcin homolog	*Crassostrea gigas*	XP_011426213.2	185/185	100	0	185
LOC105334536	Sarcoplasmic calcium-binding protein isoform X2	*Crassostrea gigas*	XP_011436355.1	178/178	100	0	178
LOC105335092	Neurocalcin homolog	*Crassostrea gigas*	XP_011437085.1	185/185	100	0	185
LOC105336644	Regucalcin isoform X1	*Crassostrea gigas*	XP_011439353.1	308/308	100	0	308
	Regucalcin isoform X2	*Crassostrea gigas*	XP_019926340.2	308/308	100	0	308
LOC105336648	Regucalcin transcript variant X1	*Crassostrea gigas*	XP_011439356.2	300/300	100	0	300
	Regucalcin variant X3	*Crassostrea gigas*	XP_011439357.2	300/300	100	0	300
LOC105335791	Annexin A4	*Crassostrea gigas*	XP_011438182.2	311/311	100	0	311
Extracellular region (GO:0005576)							
LOC105321645	Poly(3-hydroxyalkanoate) depolymerase C	*Orchesella cincta*	ODN05040.1	145/328	44	4E-83	350
	Poly(3-hydroxyalkanoate) depolymerase C	*Orchesella cincta*	ODN05040.1	106/233	45	6E-65	350
LOC105326593	Pif97	*Crassostrea gigas*	AFT63504.1	438/438	100	0	438
LOC105342849	Trithorax group protein osa-like (for transcript variant X1)	*Crassostrea virginica*	XP_022308844.1	368/697	53	4E-158	1637
	Trithorax group protein osa-like (for transcript variant X2)	*Crassostrea virginica*	XP_022308844.1	368/697	53	3E-158	1637
LOC105340688	NPC intracellular cholesterol transporter 2	*Crassostrea gigas*	XP_011445166.2	144/144	100	0	144
LOC105341555	Mammalian ependymin-related protein 1	*Crassostrea gigas*	XP_011446436.2	205/205	100	0	205
LOC105317660	Peritrophin-44	*Crassostrea gigas*	XP_011412673.1	288/288	100	3E-98	138
LOC105346858	Lactadherin, 1	*Mizuhopecten*	OWF43623.1	40/141	28	1E-06	565

		yessoensis					
	Lactadherin, 2	*Mizuhopecten yessoensis*	OWF43623.1	40/141	28	1E-06	565
LOC105318009	Golgi-associated plant pathogenesis-related protein 1	*Crassostrea gigas*	XP_011413171.2	236/236	100	0	236
LOC105331104	Collagen alpha 1(XII) chain	*Mytilus coruscus*	CAC5362153.1	373/693	54	0	797
LOC105345005	Dermatopontin isoform X1	*Crassostrea gigas*	XP_011451270.1	192/192	100	0	192
	Dermatopontin isoform X2	*Crassostrea gigas*	XP_034301840.1	182/182	100	0	182
LOC105338558	Metalloproteinase inhibitor 3	*Crassostrea gigas*	XP_011442039.2	251/251	100	0	251
LOC105336951	Peritrophin-1-like	*Mizuhopecten yessoensis*	XP_021368929.1	101/233	43	4E-53	239
LOC105329937	Peritrophin-1-like	*Gigantopelta aegis*	XP_041349873.1	67/144	47	9E-35	229
LOC105345185	Inactive pancreatic lipase-related protein 1	*Crassostrea gigas*	XP_011451564.2	357/357	100	0	357
LOC105331423	Pancreatic lipase-related protein 2	*Crassostrea gigas*	XP_011431904.2	518/518	100	0	158
LOC105343616	Adenosine deaminase AGSA isoform X1, X2	*Crassostrea gigas*	XP_011449339.2	612/612	100	0	645
LOC105327922	von Willebrand factor C domain-containing protein 2-like	*Crassostrea gigas*	XP_011426902.2	137/137	100	0	137