Characterisation of Disulfide-Rich Peptides

Exploring Potential Wound Healing Properties

Rozita Takjoo

M.Sc. Analytical Chemistry

B.Sc. Chemistry

James Cook University

For the degree of Doctor of Philosophy in Medical and Molecular Sciences

College of Public Health, Medical and Veterinary Sciences

February 2022

Acknowledgements

PhD has been a life-changing experience for me, and this project would have been impossible without the support and guidance that I received from the people whose enthusiasm and assistance helped to drive this work for four years. Since I started my PhD on February 26th, 2018, I have felt at home at the Australian Institute for Tropical Health and Medicine (AITHM). I have been given unique opportunities and taken advantage of being in a friendly atmosphere. I am indebted to all my friends in the AITHM who were so helpful in many ways. I am grateful to all priceless contributions, assistance, and support I have received over the years.

Firstly, I would like to express my deepest appreciation to my advisor, Professor Norelle L. Daly, for all the support and encouragement she provided me. Without her guidance and constant advice, this PhD would not have been achievable. She has supported me not only by providing me a research assistantship over last few years, but also by academically and emotionally supporting me through the journey to finish this thesis. I would also like to extend my sincere thanks to my secondary advisor, Dr. Michael J. Smout, for his invaluable advice on my research, and for always being so patient and supportive of my work. I would like to extend my gratitude to Dr. David T. Wilson, my other secondary advisor, who made it possible for me to obtain great insights into HPLC analysis and mass spectroscopy. He was always willing to offer advice, always willing to solve problems, and always available.

I greatly appreciate the support received throughout the completion of this work from Dr. Paramjit Bansal, who taught me peptide chemistry. He has been a true friend with his never-ending encouragement since I started my PhD. I would like to thank Prof. Alex Loukas for his invaluable suggestions on all manuscripts. I am also grateful to AITHM lab manager, Mr Phill Walsh, for his invaluable laboratory advice, Ms. Doris Pierce for her kind advice and support, and my lovely Colombian friend Ms. Linda Hernandez Duran for her friendship, encouraging smiles, and words of support.

I would like to thank the PhD students in Daly's group for their help and support. Specifically, I would like to thank Dr. Claudia Cobos Caceres, Dr. Athena Andreosso, and Ms. Silvia Saggiomo for their support and for always being ready to help. A special thank you to my lovely friend, Ms. Casey A. Schmidt, for her amazing

"There's real poetry in the real world. Science is the poetry of reality."

Richard Dawkins

Statement of the Contribution of Others

During my PhD candidature my thesis work has included:

1) Peptide (synthesis, oxidation, purification, mass analysis)

2) Structural analysis of peptides (Nuclear Magnetic Resonance Spectroscopy)

3) Biological assay (Cell culture, xCELLigence cell proliferation assay)

4) Thesis write-up

I have had a JCU scholarship, travel, and research grants. I have attached a list outlining the contributions from others in the following table.

Nature of Assistance	Contribution	Name and Affiliations
Intellectual support	Project Plan and Development	Prof. Norelle L. Daly, AITHM
	Data Analysis Assistance	Prof. Norelle L. Daly, AITHM Dr. Michael J. Smout, AITHM Dr. David T. Wilson, AITHM
	Editorial Support	Prof. Norelle L. Daly, AITHM Dr. Michael J. Smout, AITHM Dr. David T. Wilson, AITHM Prof. Alex Loukas, AITHM
	Research Costs	• Australian Research Council (LE160100218) • Australian Research Council Future Fellowship (110100226) • US National Institutes of Health Award (R01CA164719)
	Stipend	• James Cook University Postgraduate Research Scholarship
	Conference Travel Assistance	• Travel grant for 13th Australian Peptide Conference, 2019 • College of Public Health, Medical and Veterinary Science, James Cook University – Student Allowance

	Publication Assistance	• College of Public Health, Medical and Veterinary Sciences Higher Degree Research Enhancement Scheme Grant–Round 1, 2020
	Doctoral Completion Grant	• College of Public Health, Medical and Veterinary Science, James Cook University
Research Assistant	Peptide Synthesis	Dr. Paramjit Bansal, AITHM
	Purification, Mass Spectrometry	Dr. David T. Wilson, AITHM
	xCELLIigence Cell Proliferation Assay	Dr. Michael J. Smout, AITHM
	NMR Spectroscopy	Prof. Norelle L. Daly, AITHM
	Structural Evolution Study	Dr. Kartik Sunagar, Evolutionary Venomics Lab, India Naeem Y. Shaikh, Evolutionary Venomics Lab, India

Published Work by the Author Incorporated into the Thesis

Chapter 3: Folding of truncated granulin peptides. **Takjoo, R**., Wilson, D., Bansal, P.S., Loukas, A., Smout, M. J., and Daly, N.L. *Biomolecules*, 2020, Doi: 10.3390/biom10081152.

Author contributions:

NLD and **RT** designed the study. **RT** carried out peptide synthesis, oxidation, and purification of the N-terminal truncated peptides, and performed structural analysis studies. **RT** carried out the cell proliferation assay. NLD, MJS and DW supervised the project. **RT** and NLD wrote the paper. All authors have read and agreed to the published version of the manuscript.

Unpublished Works by the Author Incorporated into the Thesis

The following chapters will be submitted for publication shortly.

Chapter 2: Identification of an inter-cysteine loop potentially involved in the activity of *Ov*-GRN-1. **Takjoo, R**., Wilson, D. T., Bansal, P.S., Loukas, A., Smout, M. J., and Daly, N. L.

Author contributions:

NLD designed the study. **RT** carried out peptide synthesis, oxidation, and purification. **RT** and NLD performed structural analysis studies. **RT** and MJS carried out the cell proliferation assay. **RT** and NLD wrote the paper. All authors discussed the results and commented on the manuscript.

Chapter 4: The granulin disulfide framework: Distinctions between the N- and C-terminal regions. **Takjoo, R**., Wilson, D. T., Loukas, A., Smout, M. J., and Daly, N.L.

Author contributions:

NLD and **RT** designed the study. **RT** carried out peptide synthesis, oxidation, and purification. RT performed all the cell assays and structural analysis studies. **RT** and NLD wrote the paper. All authors discussed the results and commented on the manuscript.

Chapter 5: Structural Analysis of an *Asterias rubens* peptide indicates the presence of a disulfide directed ß-hairpin. Takjoo, R., Schmidt, C. A., Wilson, D. T., Shaikh, N. Y., Sunagar K., Loukas A., Smout, M. J., and Daly, N. L.

Author contributions:

NLD and **RT** conceived and designed the study. **RT** carried out peptide synthesis, oxidation, and purification. **RT** caried out the cell proliferation assay and structural analysis studies. **RT** and NLD wrote the paper. All authors discussed the results and commented on the manuscript.

Abstract

Naturally occurring disulfide-rich peptides are being explored as drug leads for a wide range of applications because of their potent bioactivities and relative stability compared to non-constrained peptides. However, there are major gaps in our understanding of the folding and evolution of these peptides. The thesis focusses on exploring the potential of disulfide-rich peptides as wound healing agents based on two main sources: (i) a family of growth factor proteins and (ii) an organism with remarkable regenerative properties. Furthermore, it provides fundamental information regarding the folding of a growth-factor family and the evolution of a disulfide structural motif.

Chapter 2 identified an inter-cysteine loop important for bioactivity in *Ov*-GRN-1. *Ov*-GRN-1 is a protein belonging to the granulin family in the excretory/secretory products of the carcinogenic liver fluke, *Opisthorchis viverrini*, that is involved in cell proliferation and wound healing. Although granulins generally contain 12 cysteine residues, *Ov*-GRN-1 has been previously shown to fold independently with a construct comprising only the first six cysteine residues. This truncated form, referred to as *Ov*-GRN$_{12\text{-}35\_3s}$, contains 24 residues, a non-native disulfide bond with respect to the full-length granulin structure, and stimulates cell proliferation and *in vivo* wound healing. This structure has been referred to as a "mini-granulin" fold. To explore whether it is possible to further truncate *Ov*-GRN-1, a series of small fragments of *Ov*-GRN$_{12\text{-}35\_3s}$, referred to as GRN-L1, GRN-L2, and GRN-L3, were chemically synthesised using solid-phase peptide synthesis. Nuclear magnetic resonance (NMR) spectroscopy analysis indicates that GRN-L1, GRN-L2, and GRN-L3 are unstructured in solution; however, GRN-L2 promotes cell proliferation significantly in a human normal fibroblast cell line. The results gained for this chapter demonstrate that a small unstructured peptide, GRN-L2, might be responsible, in part, for the bioactivity of *Ov*-GRN-1 as a wound healing agent. Our results show that well-defined structures are not crucial for bioactivity of granulin peptides, which has implications for further drug design studies related to wound healing.

Chapters 3 and 4 further explored the structures and functions of the N- and C-terminal regions of granulins. The significant sequence and structural variability in the granulin family led to the question of whether or not the discrete N-terminal domain identified for *Ov*-GRN-1 is a common feature of granulins. Furthermore, the C-terminal half of granulins is often poorly structured in full-length granulins, and when this study began there were no reports on the folding or activity of truncated C-terminal granulin regions. Chapter 3 studied granulin peptides from the zebrafish model organism and showed that the mini-granulin fold is present in a naturally occurring paragranulin. This mini-granulin fold is also present in a truncated form of granulin AaE and appears to stabilize the fold and enable the peptide to form a β-hairpin structure. The results obtained for this chapter showed that formation of the non-native disulfide bond into the N-terminal region of granulin might be a common feature across the family, but that bioactivity is highly dependent on the primary structure. Moreover, folding of the paragranulin-derived peptide is more efficient *in vitro* compared to the truncated version. Chapter 4 showed the C-terminal region of zebrafish AaE does not fold independently into a well-defined structure, and has no significant inhibitory effect on cell growth. These results highlight differences between the N- and C-terminal regions of granulins and has implications for the evolution of the domain.

Chapter 5 examined the structure and bioactivity of a peptide (KASH2) with injury-related effects identified in the common sea star, *Asterias rubens*. Sea stars including *Asterias rubens* have remarkable regenerative properties, and several proteins/peptides from the coelomic fluid of this sea star have recently been shown to be potentially involved in wound healing. Despite the potential of these peptides, there is no structural information on them, and no bioassays have been carried out on the isolated peptides. Interestingly, KASH2 has no effect on the cell growth of normal human skin fibroblast cells; but structural analysis demonstrates the presence of an ancestral protein fold termed the disulfide directed β-hairpin (DDH). The DDH fold was originally proposed to be restricted to arachnids but here we show it is more widespread and is likely the result of convergent evolution. Elucidating the role of KASH2 within the sea star requires further study.

In summary, this thesis provides a greater understanding of the folding and structure/function relationships of disulfide rich peptides with wound healing potential. The results highlight the potential of the "miniaturisation" approach allowing us to understand the functional significance of peptides as potential drug leads and also provided insight into the evolution of a particular class of disulfide-rich peptides.

Table of Contents

List of Tables

List of Figures

Chapter 1. Introduction

1.1 Development of disulfide rich peptides as potential drug leads

Since the introduction of insulin in the 1920s for the treatment of type 1 diabetes [1], peptide therapeutics have become important in medicine [2]. There are around 80 peptide drugs on the market for a wide range of diseases, including diabetes, cancer, osteoporosis, multiple sclerosis, HIV infection, and chronic pain [1]. Examples of these drugs include dulaglutide for diabetes, leuprolide for cancer, glatiramer for multiple sclerosis, carfilzomib for acromegaly, and vasopressin for central diabetes insipidus [1, 2].

Peptides have shown they can have high specificity, low systemic toxicity [3], and few side effects [4], making them attractive as drug leads. Along with remarkable potential of peptides as therapeutics, some drawbacks related to their use include limited oral bioavailability, poor plasma stability and short circulation time [1, 5-7]. Several approaches have been developed to overcome this limitation, including chemical modifications such as incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization [8]. The advantage of the stability conferred by multiple disulfide bonds has also been popular in peptide-based drug development. For example, the disulfide-rich peptides ziconotide (PRIALT®) and linaclotide, have been approved for treating severe and chronic pain [9], and irritable bowel syndrome respectively [10]. Several disulfide-rich peptides are in pre-clinical and clinical trials, including the spider venom peptide, Hi1a, which has promise for the development of therapeutics to protect the brain from ischemic injury [11], and a modified cyclotide [T20K] kalata B1, with immunomodulatory properties, that is on its way to clinical trials as a promising therapeutic for multiple sclerosis [12]. Cyclotides are exceptionally stable peptides that naturally combine both backbone cyclization and three disulfide bonds [13, 14].

Although disulfide-rich peptides are appropriate for a range of therapeutic applications, this thesis focuses on their wound-healing potential. Wound healing treatments are well suited to topical application, and consequently the limited oral bioavailability often associated with peptides, might not be an issue in this field. Two promising approaches for the development of new wound-healing agents take inspiration from nature, relating to host/parasite interactions and organisms known to have remarkable regenerative properties. For example, the study of host/parasite

interactions led to the discovery of a granulin peptide, *Ov*-GRN-1, from the parasitic worm *Opisthorchis viverrini* [15]. A recombinant version of *Ov*-GRN-1 stimulates rapid healing of wounds and allows the liver fluke parasite to repeatedly harvest new epithelial tissue [16]. *Ov*-GRN-1, is unique in its ability to facilitate this cycle because it regulates several wound-healing mechanisms in parallel. A downsizing, and mutational approach has been used to develop potent wound-healing peptides [17, 18] based on *Ov*-GRN-1. The aim of downsizing proteins is to significantly reduce their molecular weight, while maintaining bioactivity and increasing stability. Smaller peptides are also generally cheaper to manufacture than proteins, and less-immunogenic [19]. The most potent of the truncated *Ov*-GRN-1 peptides discovered, referred to as GRN$_{P4A}$, contains 24 residues, three disulfide bonds, and has improved wound-healing properties and reduced immunogenic potential compared to the parent protein [18]. There is limited information on the structure/function relationships of the truncated granulin peptides, and it is not clear how small the fragments can be while still maintaining their wound-healing properties.

Other organisms with remarkable regenerative properties include the sea star *Asterias rubens*. A recent study has identified several proteins in this sea star that are upregulated in response to injuries including puncture wounds and blood loss, and consequently are thought to have potential as wound healing agents [20]. However, the structures and bioactivities of the isolated proteins or peptides have not been explored.

Synthetic versions of granulin and selected sea star peptides were studied in this thesis. To put this work into context a brief background to granulins and sea star peptides is provided in sections 1.2 and 1.3.

1.2   Granulins

Granulins are a family of protein growth factors involved in several processes including proliferation/cell growth, wound healing, inflammation, and tumour growth [21]. They are cysteine-rich peptides originally identified from leukocytes, and have molecular weights of approximately 6 kDa [22]. Granulins are found in a range of species including humans [23], fishes [24, 25], oysters [26], insects [27], parasites [15], ascidians [28], mollusks [29], and plants [30]. They are characterised by a conserved

cysteine framework as shown in Figure 1-1. Despite the conserved cysteine framework, there is significant sequence and bioactivity variation amongst granulin peptides [21]. The differences in structure and activity across the granulin peptides indicate that the sequence diversity has a significant impact on the structure/function relationships.
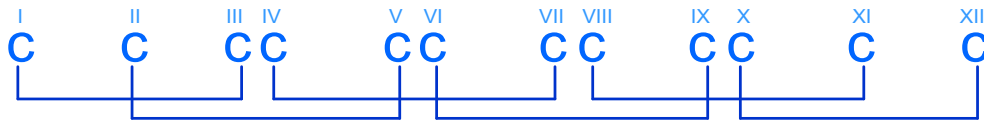
Figure 1-1 Schematic representation of cysteine framework and disulfide bond pairing present in the granulin family. The cysteine residues are numbered using Roman numerals (I-XII).

Granulin peptides are often derived from a large precursor protein referred to as progranulin (PGRN). Granulin precursors are also known as acrogranin, granulin/epithelin precursor, proepithelin, PC cell-derived growth factor, and 88-kDa glycoprotein. These names were derived from different groups and sources including specific tissues or cells [31]. The human granulin precursor was discovered in human bone marrow in 1992 [23]. It contains a signal peptide and tandem repeats of seven-and-a-half granulin motifs arranged in the order P-G-F-B-A-C-D-E, in which A-G are full repeats and P is the half-motif, as shown in Figure 1-2 [21, 32]. While the majority of granulins contain 12 cysteine residues, granulin G only contains ten cysteine residues, and the half granulin unit, paragranulin, comprises only six cysteine residues [33].

The human granulin precursor gene includes 13 exons, located on the long arm of chromosome 17. The seven granulin domains, including P-G-F-B-A-C-D-E are created by 12 out of the 13 exons of the *GRN* gene [34, 35]. The human granulin precursor is a secreted growth factor which is associated with many physiological and pathological activities similar to granulins, including embryonic development, host defence, wound repair [36], tumorigenesis, and inflammation [32].
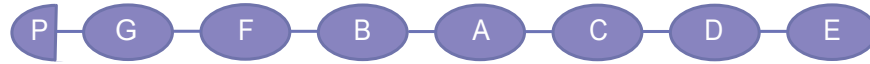
Figure 1-2 A schematic illustration for human progranulin. The order of the 7.5 granulin motifs in human progranulin, including A to G and a half-granulin motif named paragranulin (P).

An overview of the structural, engineering and bioactivity studies that have been carried out on granulins are outlined below.

1.2.1   Structures and bioactivities of full-length granulins

The three-dimensional structures of granulins from different organisms such as carp, zebrafish, and human have been previously studied. These structures have used nuclear magnetic resonance (NMR) spectroscopy, which is a powerful technique for determining the structures and dynamics of peptides and small proteins in solution [37].

Initial structural studies on granulin peptides focussed on carp granulin-1. The three-dimensional structure of carp granulin-1, extracted from spleen and head kidney of *Cyprinus carpio*, [24, 38] revealed a unique fold whereby the six disulfide bonds form a laddered arrangement, which brace the stacking of β-hairpins [38]. The structure shows partial similarity with the epithermal growth factor protein family [39, 40].

The sequences and structures of zebrafish granulins have also been determined. In zebrafish granulins, four separate precursors encoding 19 granulins were found. Two of the precursors, granulin A and granulin B, contain nine and eight granulins, respectively. The other two, progranulin l and progranulin 2, separately originate from two precursors including one and one-half granulins as shown in Figure 1-3 [41]. The structural determination of all four zebrafish granulin precursors produced using bacterial expression [42] demonstrated that granulin AaE was the most well-structured containing four β-hairpins (Figure 1-4) and displays a similar structure to the carp-granulin-1 [41, 42]. Zebrafish granulin AaE was shown to promote the survival of neuronal cells [42]. Amongst all zebrafish progranulins, PGRN-A is the orthologue of human PGRN which is involved in embryonic liver morphogenesis [43] and also muscle growth and regeneration [44, 45].

A) Zebrafish progranulin A (zPGRN-A)
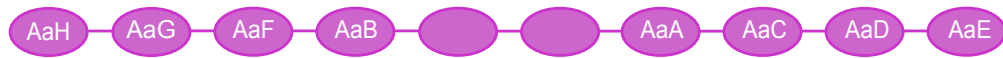
B) Zebrafish progranulin B (zPGRN-B)

C) Zebrafish progranulin 1(zPGRN-1)

D) Zebrafish progranulin 2(zPGRN-2)

Figure 1-3 Schematic illustration for some of granulin modules in zebrafish progranulins. A) Schematic representation of zebrafish progranulin A. Granulin modules with no label were missing from the partial clone [42]. B) Schematic representation of zebrafish progranulin B. C) Granulin modules from zebrafish progranulin 1. D) Granulin modules from zebrafish progranulin 2. Both [C, D] include one full-length, and a half-granulin module named paragranulin (P) that comprises six cysteines of the full length granulin [42].
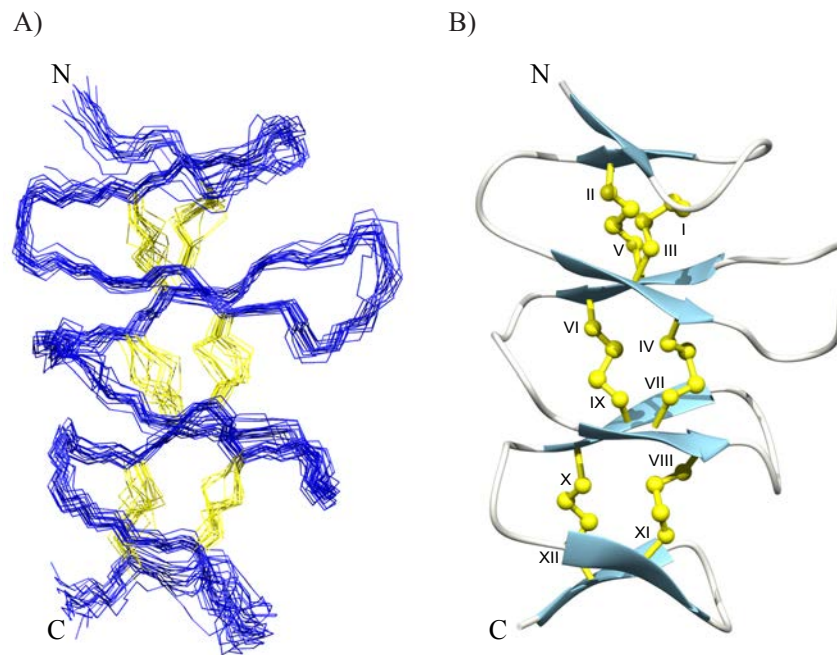
Figure 1-4 Three-dimensional structure of zebrafish granulin AaE. A) An overlay of the sixteen lowest energy structures of full-length zebrafish granulin AaE (PDB ID: 6cku). B) Ribbon representation of the secondary structure present in a single structure of full-length zebrafish granulin AaE (PDB ID: 6cku). The disulfide bonds are shown in yellow, and the β-hairpins are presented as arrows.

In contrast to the carp and zebrafish granulin structures, which show relatively well-defined structures, several of the human granulin peptides have disordered structures. The human granulins A, C, and F have well-defined N-terminal regions, but disordered C-terminal regions based on NMR studies. By contrast, human granulins B, D, and E do not have well-folded structures at all [36]. Structural analysis of a reduced form of human granulin B using an experimental and computational approach indicated that the protein was intrinsically disordered, but demonstrated biological activity inducing an inflammatory response in human neuroblastoma cells [46, 47]. Among the three partially well-folded human granulins, human granulin A inhibits proliferation of a breast cancer cell line, whereas human granulin F stimulates cell proliferation. Human granulin C has weak inhibitory activity of the proliferation of a breast cancer cell line [36].

The structural and functional differences across the granulin family highlight the complexity of the structure/function relationships. This complexity is reflected in the sequence variability as there are very few conserved residues [21]. Sequence alignment of selected granulins is shown in Table 1-1, highlighting the lack of conserved residues. With the exception of the cysteine residues, only six residues are conserved in the selected peptides.

Table 1- 1 Sequences of selected granulin peptides.

Peptide	Sequence
Ov-GRN-1*	..SPSCPDPVYTCRPGQTCCRGLHG-YGCCPMDSATCCSDLLHCCPHGTACTAYG--LCVR..
hGRNA	DVKC-DMEVSCPDGYTCCRLQSGAWGCCPFTQAVCCEDHIHCCPAGFTCDTQKGT-CEQ
Zebrafish AaE	DVQCGGGFS-CHDGETCCPTSQTTWGCCPSPKAVCCDDMQHCCPAGYKCGPGGT--CIS
Carp granulin-1	VIHCDAATI-CPDGTTCCLSPYGVWYCCPFSMGQCCRDGIHCCRHGYHCDSTSTH-CLR

Hyphen (-) indicates a sequence gap. Cysteine residues are shown in red. The sequence similarities of the granulin peptides are highlighted in cyan (Gly), green (Thr), yellow (Pro), pink (Asp), and grey (His). *Residues 27-82 of the *Ov*-GRN-1 sequence are shown in this table, (..) indicates the sequence that is not presented here.

The differences in function, particularly for human granulin A which inhibits cell growth rather than promoting cell growth as happens with other human granulins, indicates that there could be multiple biological targets for granulins. Recent studies have shown that human granulin A can interact with enolase 1 (ENO1), a glycolytic enzyme, suggesting it is involved in the anti-cancer activity reported for this peptide [48, 49]. Interestingly, it has been shown that granulin A can act synergistically with cisplatin to inhibit the growth of hepatocellular carcinoma (HCC) *in vitro* and *in vivo* [48]. Overexpression of ENO1 reduced the synergistic effect, highlighting the role of this protein in the mechanism of action of granulin A. The application of cisplatin as a chemotherapy drug for cancer treatment is restricted to HCC treatment because of resistance and toxicity, but the enhanced potency when used in combination with granulin A might allow novel treatment regimens to be developed [48].

1.2.2 Structures and bioactivities of truncated granulin precursors and peptides

Truncation of granulin precursors and peptides has led to some interesting discoveries and developments regarding the structures, bioactivity, and therapeutic potential of this family of proteins. Human progranulin has been truncated into an engineered protein termed Atsttrin, which comprises three PGRN fragments (A, C, and F), and displays selective tumour necrosis factor receptor (TNFR) binding. PGRN and Atsttrin prevent inflammation in multiple arthritis mouse models and inhibit TNFα-activated intracellular interaction [50]. Compared to PGRN, Atsttrin showed better anti-inflammatory effects with higher binding affinity to TNFR and a high stability with a longer half-life of about 120 hours [50].

Truncation of the granulin peptides themselves has also provided some interesting data. Truncation of carp granulin-1 indicates that the N-terminal region can fold independently with two disulfide-bonds (Cys I-Cys III, Cys II-Cys V, based on the numbering in the full-length protein), and the first two residues of the N-terminus are engaged in hydrophobic interactions and have a pivotal role in folding [51]. Specifically, removing the first two residues prevents formation of the first β-hairpin, most likely through interruption of the hydrophobic interactions involving the side chains of residues of Val1, His3, and Ile9 [52, 53]. Similarly, truncation of human granulin A, which also contained two disulfide bonds, revealed that both hydrophobic and turn-stabilizing interactions are crucial for having the correct disulfide bond and two β-hairpins [36, 53].

The parasitic liver-fluke granulin, *Ov*-GRN-1, has significant sequence differences to carp, human, and zebrafish granulins [17, 54]. Despite the recombinant version of *Ov*-GRN-1 being shown to accelerate wound repair in mice, making it of interest in the design of novel wound healing agents, the yield is low and it requires complex refolding of denatured protein [18], restricting its application as a therapeutic. To address these limitations, truncated analogues of *Ov*-GRN-1 have been developed [17, 18]. Analysis of the structures of truncated versions of *Ov*-GRN-1 peptides indicates distinct differences between the carp granulin-1 truncated analogues. Truncated versions of *Ov*-GRN-1 peptides containing disulfide bonds between Cys I-

Cys III and Cys II-Cys V (analogous to the carp granulin-1 and human granulin A truncated peptides) do not contain the characteristic β-hairpins present in granulins. The structures are relatively well-defined but contain irregular secondary structure [17]. Following on from these studies, an additional truncated Ov-GRN-1 peptide was synthesised that contains two additional cysteine residues (Cys IV and Cys VI). The structure of this peptide was well-defined with a β-hairpin motif, indicating that the N-terminal region of Ov-GRN-1 can fold independently [17, 18]. Interestingly, this peptide appears to contain a non-native disulfide bond between Cys IV-Cys VI, in addition to the disulfide bonds between Cys I-Cys III and Cys II-Cys V.

A range of proline to alanine substitutions were synthesised and, excitingly, these truncated forms of Ov-GRN-1 maintain or exceed the potent wound-healing activity observed for the full-length protein in a mouse model [18] and are significantly easier to produce, making them more viable candidates for drug development. The mutant containing the proline 4 to alanine substitution (GRN$_{P4A}$), appears to be the most potent *in vivo* [18]. Treatment with GRN$_{P4A}$ resulted in 20% faster wound healing compared to Regranex, the standard biological therapeutic for the treatment of chronic wounds. In addition, GRN$_{P4A}$ promoted 43% faster wound healing compared to a control peptide, and 16% faster wound healing compared to the full length Ov-GRN-1 protein [18]. Further characterisation of the structure/function relationships of granulin peptides might aid in their development as drug leads for wound healing.

1.3 Sea star peptides

1.3.1 Bioactive compounds from sea stars

Sea stars are marine invertebrates (phylum Echinodermata; class Asteroidea) that have yielded a range of isolated compounds including steroids, glycosides, alkaloids, phospholipids, peptides [55], and ceramide derivatives like cerebrosides [55] that show a broad spectrum of biological activities including antimicrobial [56], antitumour [57, 58], antimutagenic [59], cytotoxic [60-62], hemolytic [63, 64], antifungal [65], and wound healing [66-69] activities. For example, cerebrosides derived from the sea star *Asterias amurensis* have *in vivo* and *in vitro* anti-tumour activities [57]. Cerebrosides are one of the simplest lipid classes of glycosphingolipids [57], and seem to be important components of a wide variety of tissues and organs in biological systems [70]. Additionally, two steroidal glycosides isolated from the sea star *Anasterias minuta* have demonstrated antifungal activity against the plant pathogenic fungus *Cladosporium cucumerinum* [65].

One of the most extensively studied processes in echinoderms, especially sea stars, is arm regeneration [71], and several findings have shown these creatures are effective models for wound healing and regeneration. For example, the aqueous-methanol extract of sea star *Pentaceraster regulus*, containing polar compounds; reguloside-A, reguloside-B, and two unidentified steroidal saponins, have a positive effect on wound healing in guinea pigs after topical application of the extract [66]. The effects of bioactive compounds from normal hexane fractions (NHFs) and methanol-water fractions (MWFs) of the arms of *Archaster typicus* sea star have been studied in *in vitro* wound healing and *in vivo* lower jaw regeneration of zebrafish assays. The results reveal that bioactive compounds from the sea star arms stimulate blastema formation and remodeling, respectively, and inhibit tissue overgrowth [69].

1.3.2 Bioactive peptides and proteins from sea stars

A range of bioactive peptides have also been isolated and characterised from sea stars. These peptides include a muscle relaxant neuropeptide comprising 16 amino acids, which was purified from the whole body of the sea star *Patiria pectinifera* and identified as a member of the PP/OK-type neuropeptide family [72]. Another myorelaxant neuropeptide isotype containing 16 amino acids has also been isolated from the pyloric caeca of sea star *Patiria pectinifera* [73]. Neuropeptides have been identified as chemical signals, and are biologically active compounds that contribute to various mechanisms including metabolism regulation, memory, pain perception, maintenance of homeostasis, and immune processes [73]. Several studies have suggested that marine echinoderms are a potential source for the discovery of novel antibiotics [74-76], and a novel cysteine-rich peptide from a coelomic epithelium extract of the sea star *Patiria pectinifera* containing 38 residues has been isolated and shown to have antimicrobial activity [56].

A relaxin-like gonad-stimulating peptide (RGP) isolated from the sea star *Aphelasterias japonica* is the first-identified invertebrate gonadotropin responsible for final gamete maturation [77]. The new RGP ortholog, AjaRGP, induces gamete spawning and oocyte maturation in ovarian fragments of *Aphelasterias japonica* [77]. Another RGP has been recognised from the sea star *Astropecten scoparius* [78]. The effect of synthetic non-amidated (AscRGP-GR) and amidated (AscRGP-NH$_2$) peptides on gamete shedding activity has been examined using *Astropecten scoparius* ovaries. Both AscRGP-GR and AscRGP-NH$_2$ provoked oocyte maturation and ovulation in similar dose-dependent manners, and act as a natural gonadotropic hormone in this sea star [78].

Bioactive proteins have also been isolated and identified from sea stars. For example, a protein isolated from the coelomic fluid of sea star *Astropecten indicus* termed AiP1, with a molecular weight of 66 kDa, stimulates wound healing and thrombolytic activities *in vitro* [79]. Different body parts, including gastrointestinal organs and eggs, of the common sea star *Asterias rubens* have been shown to have relatively high antibacterial activity against pathogenic bacteria [75]. Lysozyme might be responsible for some of the observed antibacterial activity [75].

A recent proteomic study of the cell-free coelomic (biofluid) of the common sea star, *Asterias rubens*, highlighted a range of upregulated proteins potentially involved in defence, cell migration, wound healing, and regeneration [20]. This study looked at proteomic changes in response to injury including puncture wounds and blood loss. Although more than 90 proteins/peptides were identified in this study, three uncharacterised secretory 'orphan' soluble proteins termed CHUPA (CHU), KARTESH (KASH), and WASSP were identified that were upregulated in response to injury. The CHU and KASH peptides are present in two and three isoforms respectively, that can imply some functional divergence [20]. The CHU sequence has close homologs in sea stars (Asteroidea) and sea urchins (Echinoidea), whereas KASH and WASSP appear to be specific to Asteroidea. Both the CHU and KASH sequences have highly conserved four-cysteine residue motifs with different cysteine spacing. Analysis of the structures of these peptides might provide insight into their functions and their roles in response to injury.

1.4   Solid-phase peptide synthesis

To be able to characterise peptides as drug leads, they have to be produced in adequate yields. There is often insufficient amounts of material available from natural sources, making it necessary to produce the peptides/proteins using other methods and to allow the production of analogues. Recombinant expression and solid-phase peptide synthesis (SPPS) are the main methods used for the production of peptides and proteins, with the latter being useful in the production of the granulin truncated peptides [17, 18, 51, 52].

Peptide synthesis involves the formation of a peptide bond between two amino acids. The first coupling of two amino acids was successfully done by Emil Fischer [80]. Several methods have been used for peptide synthesis but the solid-phase method introduced by R. B. Merrifield [80, 81] is the most commonly used. In SPPS, two main protocols are commonly used utilizing either Boc (tert-butyloxycarbonyl) or Fmoc (fluorenylmethyloxycarbonyl) chemistry [82]. Boc chemistry is very effective for producing peptides but generally requires the use of trifluoroacetic acid (TFA) for deprotection of the Boc-protected amino acid, and cleavage of the peptide from the resin requires hydrogen fluoride. The use of TFA during the deprotection cycle is problematic for automation, and the hazards associated with the use of hydrogen

fluoride can be restrictive. By contrast, the Fmoc method protects the α-amino function, acid-labile side-chain protecting groups and solid-phase resin linkers using the base-labile N-Fmoc group. The use of moderate acidic conditions for final deprotection and cleavage of the peptide is the paramount advantage of this method [82]. Furthermore, by removing the use of TFA in the peptide assembly, the Fmoc method can be easily automated [83]. Fmoc chemistry has been used in the synthesis of carp granulin-1 [51] and in the synthesis of the *Ov*-GRN-1 truncated analogues [17, 18].

Following peptide synthesis and purification, the process of oxidative folding involves the formation of disulfide bonds in proteins/peptides, and the mechanism comprises a series of thiol/ disulfide exchange reactions between cysteine thiolates and an oxidising disulfide [84]. Disulfide bonds stabilise proteins and peptides and generally improves their function [85, 86]. Under appropriate conditions, peptides can fold into their native conformation. However, it can be challenging to form the native disulfide bonds of peptides and proteins with multiple disulfide bonds *in vitro*. Chemical methods for the regioselective formation of disulfide bonds using a combination of two or three types of thiol protecting groups, such as acetamidomethyl (Acm) and 4-methylbenzyl (MeBzl) groups, can be used for the stepwise formation of disulfide bonds [87, 88]. The pathways involved in formation of the disulfide bonds can influence the effectiveness of this strategy as formation of a particular disulfide bond prior to other bonds can potentially prevent the correct folding.

1.5 Clinical targets for peptides

Peptides represent an exceptional class of therapeutic compounds, occupying a niche between small molecules and proteins [2]. There are many advantages in using peptides as therapeutics, such as standard synthetic protocols, high selectivity and potency, shorter time to market, and predictable metabolism [89]. The potential bioactivity identified in granulin peptides such as *Ov*-GRN-1 [15] provides opportunities to design new treatment agents. Given the enormous variety among granulin compounds, they have been implicated in a range of therapeutic applications. The most important targets for researchers are diseases that cause huge problems and health burdens on society, such as diabetes.

Diabetes is a 'silent pandemic' that represents one of the most challenging public health and worldwide health problems of the 21st century. Diabetes is associated with nerve damage and poor circulation in the lower limbs, either of which may lead to foot ulcers and infections, and eventually to amputations. Almost 15% of people with diabetes suffer from diabetic foot ulcers which can lead to chronic wounds [90]. Few active treatments have been approved for chronic wounds. Regranex, a recombinant human platelet–derived growth factor, is currently the only biological wound-healing agent approved by the FDA for chronic ulcers [91, 92]. As diabetes is a huge challenge to our nation's health and economy, new wound healing agents are required for the treatment of chronic wounds from diabetic ulcers, and much more research is still needed to continue improving outcomes and reducing the burden on individuals and families. Exploring the structure/function relationships of peptides with wound healing potential might help to produce better therapeutics and alleviate this burden.

1.6 Scope of the thesis

The overall aim of this study is to expand the knowledge of the structure/function relationships of disulfide-rich peptides with potential as wound healing agents. Understanding how peptides fold and elucidating structure/function relationships can aid in the development of peptides as drug leads. The importance of these types of studies has already been shown with the *Ov*-GRN-1 truncated peptides, whereby analogues with higher wound healing potency and higher folding yields have been developed.

This thesis has provided fundamental information regarding the structural and functional properties of peptides from a parasite, zebrafish, and the sea star *Asterias rubens*. Although peptide folding is a key aspect in biology, the processes involved are still not fully understood. This study has addressed some of the knowledge gaps in predicting the folds of peptides and engineering novel drug leads, as well as insight into evolution of a structural fold. The design of the chapters is based on published manuscripts or manuscripts to be published.

Chapter 1 provides a background to the disulfide-rich peptides highlighted in this project including the granulin family, which is a major focus of this work, and bioactive peptides from sea stars.

Chapter 2 examines the role of the inter-cysteine loops in *Ov*-GRN-1 in cell proliferation activity, and how small the truncation can be while still conserving wound healing properties. It is generally believed that well-defined structures are essential for bioactivity of the peptides, but this chapter explored the idea that a small unstructured peptide might be involved in bioactivity.

Chapter 3 provides insight into the folding of the N-terminal region of granulins, including a naturally occurring paragranulin. The results gained from this chapter illustrated that, while the non-native disulfide bond in the N-terminal region of truncated granulin might be a common feature across the family, the bioactivity is highly dependent on the primary structure. This study is published in *Biomolecules*, doi: 10.3390/biom10081152 [93].

For further investigation on zebrafish granulin AaE, Chapter 4 explored the structure/activity relationships of the C-terminal region of zebrafish granulin AaE. The results showed that peptide corresponding to the C-terminal region of zebrafish AaE do not fold independently into well-defined structures and have no significant inhibitory effect on cell growth.

Chapter 5 explores the structural analysis of a peptide from the common sea star, *Asterias rubens*. This work expands the scope of the project by including an additional peptide class with potential in wound healing. The coelomic proteome of this sea star has previously been shown to produce several proteins and peptides potentially involved in wound healing. In the current study we have characterised a peptide from *A. rubens*, which led to the identification of an ancestral protein fold termed the disulfide directed β-hairpin (DDH) that was originally proposed to be restricted to arachnids. This study also highlights the challenges of using proteomic changes to identify potential wound healing peptides.

Chapter 6 provides a summary of the outcomes from the thesis and the potential future directions.

Chapters 2, 4 and 5 are unpublished, written in a paper format, that will be submitted shortly.

1.7 References

1. Muttenthaler, M., et al., *Trends in peptide drug discovery.* Nat. Rev. Drug Discov., 2021. 20(4): p. 309-325.

2. Lau, J.L. and Dunn, M.K., *Therapeutic peptides: Historical perspectives, current development trends, and future directions.* Bioorg. Med. Chem., 2018. 26(10): p. 2700-2707.

3. Loffet, A., *Peptides as drugs: Is there a market?* J. Pept. Sci., 2002. 8(1): p. 1-7.

4. Yin, N., et al., *Enhancing the oral bioavailability of peptidedrugs by using chemical modification and other approaches.* Med. Chem., 2014. 4(12).

5. Sachdeva, S., *Peptides as 'Drugs': The journey so far.* Int. J. Pept. Res. Ther., 2017. 23(1): p. 49-60.

6. Apostolopoulos, V., et al., *A global review on short peptides: Frontiers and perspectives.* Molecules (Basel, Switzerland), 2021. 26(2): p. 430.

7. Peigneur, S., et al., *Small cyclic sodium channel inhibitors.* Biochem. Pharmacol., 2021. 183: p. 114291.

8. Gentilucci, L., et al., *Chemical modifications designed to improve peptide stability: incorporation of non-natural amino acids, pseudo-peptide bonds, and cyclization.* Curr. Pharm. Des., 2010. 16(28): p. 3185-203.

9. Miljanich, G.P., *Ziconotide: Neuronal calcium channel blocker for treating severe chronic pain.* Curr. Med. Chem., 2004. 11(23): p. 3029-40.

10. Busby, R.W., et al., *Pharmacologic properties, metabolism, and disposition of linaclotide, a novel therapeutic peptide approved for the treatment of irritable bowel syndrome with constipation and chronic idiopathic constipation.* J. Pharmacol. Exp. Ther., 2013. 344(1): p. 196-206.

11. Chassagnon, I.R., et al., *Potent neuroprotection after stroke afforded by a double-knot spider-venom peptide that inhibits acid-sensing ion channel 1a.* Proc. Natl. Acad. Sci., 2017. 114(14): p. 3750-3755.

12. Gründemann, C., et al., *T20K: An immunomodulatory cyclotide on its way to the clinic.* Int. J. Pept. Res. Ther., 2019. 25(1): p. 9-13.

13. Craik, D.J. and Conibear, A.C., *The chemistry of cyclotides.* J. Org. Chem., 2011. 76(12): p. 4805-4817.

14. de Veer, S.J., et al., *Cyclotides: From structure to function.* Chem. Rev., 2019. 119(24): p. 12375-12421.

15. Smout, M.J., et al., *A granulin-like growth factor secreted by the carcinogenic liver fluke, Opisthorchis viverrini, promotes proliferation of host cells.* PLoS Pathog., 2009. 5(10): p. e1000611.

16. Smout, M.J., et al., *Carcinogenic parasite secretes growth factor that accelerates wound healing and potentially promotes neoplasia.* PLoS Pathog., 2015. 11(10): p. e1005209.

17. Bansal, P.S., et al., *Development of a potent wound healing agent based on the liver fluke granulin structural fold.* J. Med. Chem., 2017. 60(10): p. 4258-4266.

18. Dastpeyman, M., et al., *Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing.* J. Med. Chem., 2018. 61(19): p. 8746-8753.

19. Reid, R.C., et al., *Downsizing a human inflammatory protein to a small molecule with equal potency and functionality.* Nat. Commun., 2013. 4: p. 2802.

20. Shabelnikov, S.V., et al., *Injury affects coelomic fluid proteome of the common starfish, Asterias rubens.* J. Exp. Biol., 2019. 222(6): p. jeb198556.

21. Dastpeyman, M., et al., *Folding of granulin domains.* Pept. Sci., 2018. 110(3).

22. Bateman, A., et al., *Granulins, a novel class of peptide from leukocytes.* Biochem. Biophys. Res. Commun., 1990. 173(3): p. 1161-1168.

23. Bhandari, V., et al., *Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains.* Proc. Natl. Acad. Sci., 1992. 89(5): p. 1715-1719.

24. Belcourt, D.R., et al., *Isolation and primary structure of the three major forms of granulin-like peptides from hematopoietic tissues of a teleost fish (Cyprinus carpio).* J. Biol. Chem., 1993. 268(13): p. 9230-9237.

25. Cadieux, B., et al., *The zebrafish progranulin gene family and antisense transcripts.* BMC Genom., 2005. 6(1): p. 156-156.

26. Zhao, M., et al., *Functional characterization and molecular mechanism exploration of three granulin epithelin precursor splice variants in biomineralization of the pearl oyster Pinctada fucata.* Mol. Genet. Genom., 2016. 291(1): p. 399-409.

27. Hanington, P.C., et al., *Molecular and functional characterization of granulin-like molecules of insects.* Insect Biochem. Mol. Biol., 2008. 38(5): p. 596-603.

28. Zhao, Y.-P., et al., *Progranulin deficiency exaggerates, whereas progranulin-derived Atsttrin attenuates, severity of dermatitis in mice.* FEBS Lett., 2013. 587(12): p. 1805-1810.

29. Nara, K., et al., *Granulin-like peptide in the mid-gut gland of the bivalve mollusk, Patinopecten yessoensis.* Biochim. Biophys. Acta., 2004. 1675(1-3): p. 147.

30. Tolkatchev, D., et al., *A peptide derived from the c-terminal part of a plant cysteine protease folds into a stack of two β-hairpins, a scaffold present in the emerging family of granulin-like growth factors.* J. Pept. Res., 2001. 57(3): p. 227-233.

31. Arechavaleta-Velasco, F., et al., *Progranulin and its biological effects in cancer.* Med. Oncol., 2017. 34(12): p. 1-11.

32. Palfree, R.G., et al., *The evolution of the secreted regulatory protein progranulin.* PloS one, 2015. 10(8): p. e0133749.

33. Ong, C.H.P. and Bateman, A., *Progranulin (granulin-epithelin precursor, PC-cell derived growth factor, acrogranin) in proliferation and tumorigenesis.* Histol. Histopathol., 2003. 18(4): p. 1275-1288.

34. Bhandari, V. and Bateman, A., *Structure and chromosomal location of the human granulin gene.* Biochem. Biophys. Res. Commun., 1992. 188(1): p. 57-63.

35. Baba, T., et al., *Exon/intron organization of the gene encoding the mouse epithelin/granulin precursor (acrogranin).* FEBS Lett., 1993. 322(2): p. 89-94.

36. Tolkatchev, D., et al., *Structure dissection of human progranulin identifies well-folded granulin/epithelin modules with unique functional activities.* Protein Sci., 2008. 17(4): p. 711-724.

37. Wuthrich, K., *NMR of proteins and nucleic acids.* 1986, New York: Wiley.

38. Hrabal, R., et al., *The hairpin stack fold, a novel protein architecture for a new family of protein growth factors.* Nat. Struct. Biol., 1996. 3(9): p. 747-751.

39. Kohda, D. and Inagaki, F., *3D Structures of EGF and TGF -Alpha determined by NMR.* Anal. Sci., 1991. 7(Supple): p. 853-856.

40. Bateman, A. and Bennett, H.P., *Granulins: the structure and function of an emerging family of growth factors.* J. Endocrinol., 1998. 158(2): p. 145-151.

41. Wang, P.F., *Structure genomics of zebrafish granulins.* 2004, McGill University, Montreal.

42. Wang, P., et al., *Structure dissection of zebrafish progranulins identifies a well-folded granulin/epithelin module protein with pro-cell survival activities.* Protein Sci., 2018. 27(8): p. 1476-1490.

43. Li, Y.-H., et al., *Progranulin A-mediated MET signaling is essential for liver morphogenesis in zebrafish.* J. Biol. Chem., 2010. 285(52): p. 41001-41009.

44. Li, Y.-H., et al., *Progranulin regulates zebrafish muscle growth and regeneration through maintaining the pool of myogenic progenitor cells.* Sci. Rep., 2013. 3(1): p. 1176.

45. Wu, J.-L., et al., *A potent tilapia secreted granulin peptide enhances the survival of transgenic zebrafish infected by Vibrio vulnificus via modulation of innate immunity.* Fish Shellfish Immunol., 2018. 75: p. 74-90.

46. Ghag, G., et al., *Fully reduced granulin-B is intrinsically disordered and displays concentration-dependent dynamics.* Protein Eng. Des. Sel., 2016. 29(5): p. 177-186.

47. Ghag, G., et al., *Disulfide bonds and disorder in granulin-3: An unusual handshake between structural stability and plasticity.* Protein Sci., 2017. 26(9): p. 1759-1772.

48. Qiao, G., et al., *Granulin A synergizes with cisplatin to inhibit the growth of human hepatocellular carcinoma.* Int. J. Mol. Sci., 2018. 19(10): p. 3060.

49. Chen, X.L., et al., *Interaction between granulin A and enolase 1 attenuates the migration and invasion of human hepatoma cells.* Oncotarget, 2017. 8(18): p. 30305-30316.

50. Tang, W., et al., *The growth factor progranulin binds to TNF receptors and is therapeutic against inflammatory arthritis in mice.* Science, 2011. 332(6028): p. 478-484.

51. Vranken, W.F., et al., *Solution structures of a 30-residue amino-terminal domain of the carp granulin-1 protein and its amino-terminally truncated 3-30 subfragment: Implications for the conformational stability of the stack of two β-hairpins.* Proteins, 2002. 47(1): p. 14-24.

52. Vranken, W.F., et al., *A 30-residue fragment of the carp granulin-1 protein folds into a stack of two β-hairpins similar to that found in the native protein.* J. Pept. Res., 1999. 53(5): p. 590-597.

53. Tolkatchev, D., et al., *Design and solution structure of a well-folded stack of two β-hairpins based on the amino-terminal fragment of human granulin A.* Biochemistry, 2000. 39(11): p. 2878-2886.

54. Smout, M.J., et al., *Expression, refolding and purification of Ov-GRN-1, a granulin-like growth factor from the carcinogenic liver fluke, that causes proliferation of mammalian host cells.* Protein Expr. Purif., 2011. 79(2): p. 263-270.

55. Dong, G., et al., *Chemical constituents and bioactivities of starfish.* Chem. Biodivers., 2011. 8(5): p. 740-91.

56. Kim, C.H., et al., *Identification of a novel antimicrobial peptide from the sea star Patiria pectinifera.* Dev. Comp. Immunol., 2018. 86: p. 203-213.

57. Du, L., et al., *The anti-tumor activities of cerebrosides derived from sea cucumber Acaudina molpadioides and starfish Asterias amurensis in vitro and in vivo.* J. Oleo. Sci., 2012. 61(6): p. 321-30.

58. Lazzara, V., et al., *Bright spots in the darkness of cancer: A review of starfishes-derived compounds and their anti-tumor action.* Mar. Drugs, 2019. 17(11).

59. Han, Y.H., et al., *Antimutagenic activity of 5alpha-cholest-7-en-3beta-ol, a new component from the starfish asterina pectinifera.* Biol. Pharm. Bull., 2000. 23(10): p. 1247-9.

60. Kang, J.X., et al., *Three new cytotoxic polyhydroxysteroidal glycosides from starfish Craspidaster hesperus.* Mar. Drugs, 2016. 14(10).

61. Kicha, A.A., et al., *Polar steroid compounds from the Arctic starfish Asterias microdiscus and their cytotoxic properties against normal and tumor cells in vitro.* Nat. Prod. Res., 2020: p. 1-8.

62. Ha, D.T., et al., *Asterosaponins from the tropical starfish Acanthaster planci and their cytotoxic and anticancer activities in vitro.* Nat. Prod. Res., 2021. 35(4): p. 548-555.

63. Ivanchina, N.V., et al., *Hemolytic polar steroidal constituents of the starfish Aphelasterias japonica.* J. Nat. Prod., 2000. 63(8): p. 1178-81.

64. Lee, K.S., et al., *Starfish polysaccharides downregulate metastatic activity through the MAPK signaling pathway in MCF-7 human breast cancer cells.* Mol. Biol. Rep., 2013. 40(10): p. 5959-66.

65. Chludil, H.D., et al., *Antifungal steroidal glycosides from the patagonian starfish anasteriasminuta: Structure-activity correlations.* J. Nat. Prod., 2002. 65(2): p. 153-7.

66. Gupta, A., et al., *Wound healing in guinea pigs after topical application of starfish Pentaceraster regulus extract.* J. Wound Care, 2008. 17(10): p. 441-4.

67. Franco, C.F., et al., *Radial nerve cord protein phosphorylation dynamics during starfish arm tip wound healing events.* Electophoresis, 2012. 33(24): p. 3764-3778.

68. Ben Khadra, Y., et al., *Re-growth, morphogenesis, and differentiation during starfish arm regeneration.* Wound Repair Regen., 2015. 23(4): p. 623-634.

69. Dai, Y., et al., *Tissue extract fractions from starfish undergoing regeneration promote wound healing and lower jaw blastema regeneration of zebrafish.* Sci. Rep., 2016. 6: p. 38693.

70. Tan, R.X. and Chen, J.H., *The cerebrosides.* Nat. Prod. Rep., 2003. 20(5): p. 509-34.

71. Ferrario, C., et al., *Fundamental aspects of arm repair phase in two echinoderm models.* Dev. Biol., 2018. 433(2): p. 297-309.

72. Kim, C.H., et al., *Identification of a novel starfish neuropeptide that acts as a muscle relaxant.* J. Neurochem., 2016. 137(1): p. 33-45.

73. Kubarova, A., et al., *Isolation of a starfish myorelaxant peptide (SMP) isotype from the pyloric caeca of Patiria pectinifera.* Fish. Aquat. Sci., 2021. 24(4): p. 163-170.

74. Choi, D.H., et al., *Characterization of antimicrobial agents extracted from Asterina pectinifera.* Int. J. Antimicrob. Agents, 1999. 11(1): p. 65-68.

75. Haug, T., et al., *Antibacterial activity in Strongylocentrotus droebachiensis (Echinoidea), Cucumaria frondosa (Holothuroidea), and Asterias rubens (Asteroidea).* J. Invertebr. Pathol., 2002. 81(2): p. 94-102.

76. Stabili, L., et al., *Screening of three echinoderm species as new opportunity for drug discovery: Their bioactivities and antimicrobial properties.* Evid. Based Complementary Altern. Med., 2018. 2018: p. 7891748-8.

77. Mita, M. and Katayama, H., *A relaxin-like gonad-stimulating peptide from the starfish Aphelasterias japonica.* Gen. Comp. Endocrinol., 2016. 229: p. 56-61.

78. Mita, M., et al., *A relaxin-like gonad-stimulating peptide identified from the starfish Astropecten scoparius.* Mol. Reprod. Dev., 2021. 88(1): p. 34-42.

79. Baveja, M., et al., *AiP1, a protein from the coelomic fluid of sea star Astropecten indicus promotes wound healing and fibrinogenolysis in vitro.* J. Basic Appl. Zool., 2019. 80(1): p. 1-13.

80. Coin, I., et al., *Solid-phase peptide synthesis: From standard procedures to the synthesis of difficult sequences.* Nat. Protoc., 2007. 2(12): p. 3247-3256.

81.   Jaradat, D.s.M.M., *Thirteen decades of peptide synthesis: Key developments in solid phase peptide synthesis and amide bond formation utilized in peptide ligation.* Amino Acids, 2018. 50(1): p. 39-68.

82.   Amblard, M., et al., *Methods and protocols of modern solid phase peptide synthesis.* Mol. Biotechnol., 2006. 33(3): p. 239-254.

83.   Dryland, A. and Sheppard, R.C., *Peptide synthesis. Part 8. A system for solid-phase synthesis under low pressure continuous flow conditions.* J. Chem. Soc., Perkin trans. 1, 1986: p. 125.

84.   Bulaj, G., *Formation of disulfide bonds in proteins and peptides.* Biotechnol. Adv., 2005. 23(1): p. 87-92.

85.   Postma, T.M. and Albericio, F., *Disulfide formation strategies in peptide synthesis.* Eur. J. Org. Chem., 2014. 2014(17): p. 3519-3530.

86.   Fischer, M., et al., *Protein import and oxidative folding in the mitochondrial intermembrane space of intact mammalian cells.* Mol. Biol. Cell., 2013. 24(14): p. 2160-2170.

87.   Okumura, M., et al., *A chemical method for investigating disulfide-coupled peptide and protein folding.* FEBS J., 2012. 279(13): p. 2283-2295.

88.   Hidaka, Y. and Shimamoto, S., *Folding of peptides and proteins: Role of disulfide bonds, recent developments.* Biomol. Concepts, 2013. 4(6): p. 597-604.

89.   Fosgerau, K. and Hoffmann, T., *Peptide therapeutics: Current status and future directions.* Drug Discov. Today, 2015. 20(1): p. 122-128.

90.   Brem, H. and Tomic-Canic, M., *Cellular and molecular basis of wound healing in diabetes.* J. Clin. Investig., 2007. 117(5): p. 1219-1222.

91.   Niezgoda, J.A., et al., *Randomized clinical trial comparing OASIS Wound Matrix to Regranex Gel for diabetic ulcers.* Adv. Skin Wound Care, 2005. 18(5 Pt 1): p. 258-266.

92.   Chan, R.K., et al., *Effect of recombinant platelet-derived growth factor (Regranex®) on wound closure in genetically diabetic mice.* J. Burn Care Res., 2006. 27(2): p. 202-205.

93.   Takjoo, R., et al., *Folding of truncated granulin peptides.* Biomolecules, 2020. 10(8): p. 1152.

Chapter 2. Identification of an Inter-Cysteine Loop Potentially Involved in the Activity of *Ov*-GRN-1

2.1 Abstract

Identification of small bioactive regions in proteins and peptides can be useful information in drug design studies. Here we have dissected the N-terminal domain of *Ov*-GRN-1, a granulin protein from the flatworm liver fluke *Opisthorchis viverrini* which has potent wound healing properties, and have shown that an inter-cysteine loop maintains bioactivity. Peptides corresponding to the three inter-cysteine loops of the N-terminal domain were produced using synthetic chemistry, and their structures and bioactivities analysed using NMR spectroscopy and cell proliferation assays respectively. As expected for such small peptides, NMR analysis indicated that the peptides were unstructured in solution. However, a seven-residue peptide corresponding to loop 2 (GRN-L2) promoted cell proliferation to the same extent as the N-terminal domain, in contrast to the other fragments. Consequently, our results suggest that GRN-L2 might be responsible, in part, for the bioactivity of *Ov*-GRN-1, and might be a useful lead molecule for subsequent wound healing studies.

2.2 Introduction

Identification of small bioactive regions in proteins and peptides can be useful in drug design studies, including peptide mimetic development [1, 2] and grafting studies [3-7]. Peptides and peptidomimetics are good candidates to target protein–protein interactions because of their specificity and ability to mimic a protein surface to effectively compete for binding. Peptides can be rationally designed based on the natural sequences that mediate protein–protein interactions in the proteins, and therefore can cover a critical part of the binding surface; furthermore, peptidomimetics can be chemically modified to stabilize the bioactive conformation, mimicking the three-dimensional structure of the binding surface of the protein [8].

Grafting bioactive peptide sequences into a template with high conformational and proteolytic stability can provide one potential solution for overcoming poor stability and bioavailability of peptide epitopes [9]. An ideal scaffolding protein is typically small, monomeric, free of internal disulfide bonds and post-translational modifications, and thermostable. These properties can confer enhanced stability and cellular longevity to bioactive peptide sequences [10].

Such bioactive peptides might also be drug leads in their own right, as they are likely to be less immunogenic than larger proteins, and easier and cheaper to manufacture [11, 12]. One of the unique approaches to producing bioactive peptides is a process referred to as protein "downsizing" [12], where small regions of proteins are produced in isolation and display bioactivity. Although downsizing will not be effective where dis-continuous regions are involved in the bioactivity, there are now several examples where this approach has been effective. For example, downsizing C3a, a 77-residue human inflammatory protein, resulted in the identification of a tripeptide with the same high potency, functional profile, and specificity of action as the full-length C3a protein, but with higher plasma stability and bioavailability [12].

The downsizing approach has also been effectively applied to the granulin protein *Ov*-GRN-1 from the parasitic human liver fluke, *Opisthorchis viverrini*, which is a 6 kDa protein that induces angiogenesis and accelerates wound repair [13]. Although the *Ov*-GRN-1 protein has potent wound-healing properties, low yields in recombinant expression have limited its development as a potential wound-healing

agent [14], and prompted studies aimed at producing smaller versions with synthetic methods. A truncated form (GRN$_{12-35\_3s}$) corresponding to residues 12-35 of *Ov*-GRN-1 was produced using solid-phase peptide synthesis and folded into a relatively well-defined structure containing a β-hairpin with three disulfide bonds [15]. This "mini-granulin" fold maintains potent *in vivo* wound healing properties [15], and incorporates a non-native disulfide bond, with respect to the full-length granulin module, which was essential for the regular secondary structure formation in the truncated peptides derived from liver fluke [16].

GRN$_{12-35\_3s}$ is still a relatively large/structurally complex peptide that comprises three inter-cysteine loops. To determine if *Ov*-GRN-1 can be downsized even further, the three inter-cysteine loops of *Ov*-GRN$_{12-35\_3s}$, referred to as GRN-L1, GRN-L2, and GRN-L3, were individually synthesised using solid-phase peptide synthesis. Cysteine residues were included at the N- and C-termini to allow formation of a disulfide bond to bring the termini in close proximity and form a loop structure. The structures and cell proliferation effects of these peptides were examined.

2.3 Experimental section

2.3.1   Peptide synthesis and purification

Truncated granulin peptides were synthesised manually by standard solid-phase methods on 2-chlorotrityl chloride resin (Auspep, Australia). HCTU (*O*-(1*H*-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate)/DIPEA was used to activate the amino acids following methods previously described [16]. Peptides were deprotected and cleaved from the resin using the following cleavage cocktail: 95% TFA: 2.5% TIPS: 2.5% H_2O (v/v/v). The peptides were then precipitated and washed several times with cold diethyl ether, dissolved in 50% acetonitrile:50% H_2O, lyophilized (at least twice to remove TFA), and stored as lyophilized powder. The resulting crude peptides were purified by reversed-phase high performance liquid chromatography (RP-HPLC) on a C_{18} preparative column (Phenomenex Jupiter C_{18}, 10 μm, 300 Å, 250 mm × 21.2 mm). Gradients of Solvent B (90% acetonitrile: 10% H_2O: 0.045% TFA) and Solvent A (99.95% H_2O: 0.05% TFA) were used, and the eluent absorbance was monitored at 214 nm. Peptides were oxidised by dissolving linear purified and lyophilized peptide in 100 mM ammonium bicarbonate (pH 8.2) for 24 h at room temperature. The pH was checked after dissolution to confirm the solutions were still at pH 8. Oxidised peptides were purified using RP-HPLC on a C_{18} preparative column (Phenomenex Jupiter C_{18}, 10 μm, 300 Å, 250 mm × 21.2 mm). The purity of the peptides was assessed using analytical RP-HPLC with a Phenomenex Jupiter 4 μm Proteo column (C_{12}, 150 × 2.00 mm, 10 μm, 90 Å) using a gradient of 0–60% solvent B (Solvent A: 99.95% H_2O: 0.05% TFA; Solvent B: 90% acetonitrile: 10% H_2O: 0.045% TFA) over 60 min with a flow rate of 0.4 mL/min. The eluent absorbance was monitored at 214 nm.

2.3.2   NMR spectroscopy and structure calculations

Samples were prepared from lyophilized peptides at a concentration of ~0.2 mM in 90% H_2O/10% D_2O (v/v) with 4,4-dimethyl-4-silapentane-1-sulfonic acid (DSS; Cambridge Isotope Laboratories) as a reference. The pH of the NMR sample solutions was ~4.5. All NMR spectra were recorded on a 600-MHz AVANCE III NMR spectrometer (Bruker, Karlsruhe, Germany) at 290 K. Two-dimensional spectra including TOCSY, NOESY, DQF-COSY, and HSQC were recorded with an interscan

delay of 1 s. NOESY spectra were acquired with mixing times of 200-300 ms, and TOCSY spectra were acquired with isotropic mixing periods of 80 ms. All spectra were processed using TOPSPIN software (Bruker, Billerica, MA, USA), and were assigned using CCPNMR [17] based on the approach described by Wüthrich [18]. The three-dimensional structures of peptides were calculated using the CYANA program, based on automated assignment of the NOEs [19]. Torsion-angle restraints predicted by TALOS-N were used in the structure calculations [20]. The disulfide bond connectivity of GRN-L3 (Cys 2−Cys 11) was included in the calculations. Structures were visualized using MOLMOL [21].

2.3.3 Cell culture

The human skin normal fibroblast cell line, 1BR.3.GN, was obtained from the European Collection of Authenticated Cell Cultures (ECACC). 1BR.3.GN cells were cultured and maintained in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) (Life Technologies, Australia) containing $1 \times$ antibiotic/antimycotic, including penicillin, streptomycin, and amphotericin B, and $1 \times$ GlutaMAX, supplemented with 10% fetal bovine serum (FBS) (Gibco, Scotland) at 37 °C and 5% CO_2. Phosphate Buffered Saline (PBS), purchased from ThermoFisher was used for negative vehicle control. Cell proliferation assays were performed with DMEM/F12 low nutrient media supplemented with 0.5% FBS.

2.3.4 Cell proliferation monitoring in real time using xCELLigence

Cells were seeded at 5000 cells per well in 150 μL of complete media in E-plates (ACEA Biosciences) and grown overnight while monitored with an xCELLigence SP system (ACEA Biosciences), which monitors cellular events in real time by measuring electrical impedance across gold microelectrodes integrated into the base of tissue culture plates. Cells were washed three times with low-nutrient media prior to addition of 150 μL of low-nutrient media and incubated for a minimum of 6 h before further treatment. Peptides were quantified based on absorbance measurements at 214 nm and predicted extinction coefficients based on the amino acid compositions [22, 23]. Treatments were prepared at $8.5 \times$ concentration and added to each well in 20 μL, for a final total volume of 170 μL. The pH was checked after dissolution to confirm the solutions were still at pH 7. The xCELLigence system recorded cell

indexes at intervals of 1 h for 5−6 days following treatment. Readings for the cell indices were normalized before treatment, and cell proliferation ratios characterised the relative numbers of cells compared to control cells. Cell proliferation rates were compared between treatment and control wells, and a one-way ANOVA test was used for multiple comparisons with Holm-Sidak's correction using GraphPad Prism 9.0.

2.4 Results

2.4.1 Design and synthesis of Ov-GRN$_{12\text{-}35\_3s}$ fragments

Three Ov-GRN$_{12\text{-}35\_3s}$ fragments (GRN-L1, -L2, and -L3) were chemically synthesised on a 0.1 mmole scale using Fmoc solid-phase peptide synthesis. In the truncated analogue GRN-L3, Cys III and Cys V were replaced with alanine residues to prevent disulfide bond formation between these residues and enable formation of the Cys IV to Cys VI disulfide bond present in Ov-GRN$_{12\text{-}35\_3s}$. The folding yields for GRN-7, GRN-8 and GRN-11 were 33%, 25% and 30% respectively. The sequences of the full-length Ov-GRN$_{12\text{-}35\_3s}$ and synthetic fragments, and the location of the fragments in the Ov-GRN$_{12\text{-}35\_3s}$ structure are shown in Figure 2-1. The crude peptides were purified using RP-HPLC and mass analysis carried out using MALDI TOF/TOF mass spectrometry.
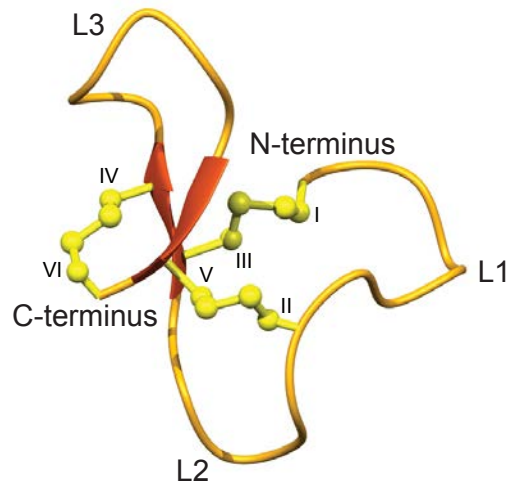
Peptide	Sequence
Ov-GRN₁₂₋₃₅_₃ₛ	**C**PDPVYT**C**RPGQT**CC**RGLHGYG**CC**
GRN-L1	**C**PDPVYT**C**
GRN-L2	**C**RPGQT**C**
GRN-L3	**AC**RGLHGYG**AC**

Figure 2-1 Three-dimensional representation of *Ov*-GRN₁₂₋₃₅_₃ₛ. The sequences of *Ov*-GRN₁₂₋₃₅_₃ₛ (PDB ID: 5UJG) and the fragments are given below the structure. The cysteine residues are numbered using Roman numerals (I-VI).

2.4.2   Structural analysis with NMR spectroscopy

The structures of the purified peptides were analysed using NMR spectroscopy. The one-dimensional spectra of the Ov-GRN$_{12-35\_3s}$ fragments did not have significant dispersion in the amide region (Figure 2-2). Further analysis of the TOCSY and NOESY spectra for the individual fragments allowed assignment of the resonances and determination of the secondary shifts, which indicated that the GRN-L1 and -L3 peptides had no consecutive positive or negative secondary chemical shifts, consistent with the peptides being unstructured in solution (Figure 2-3). Although GRN-L2 had consecutive positive secondary shifts for residues 4-7, which are consistent with the presence of β-sheet structure, there was a relatively low number of non-intra-residue or sequential NOEs in the NOESY spectrum, which prevented determination of a well-defined structure.
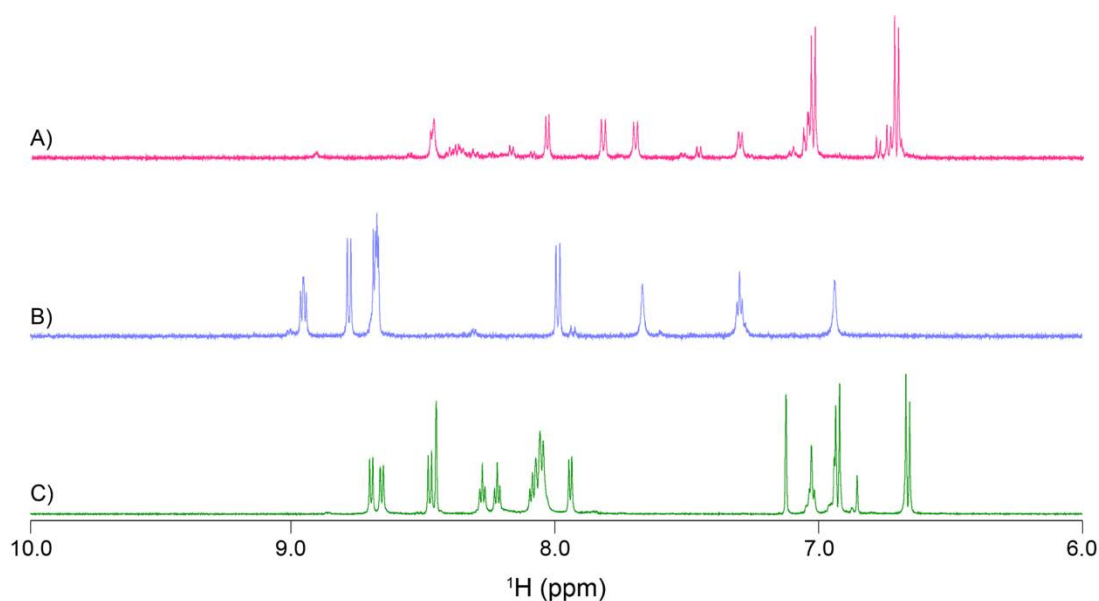
Figure 2-2 One-dimensional proton NMR spectra. A) GRN-L1, B) GRN-L2, and C) GRN-L3. Spectra are shown from 6.5 to 9.5 ppm; this region contains primarily the amide and the aromatic ring protons.
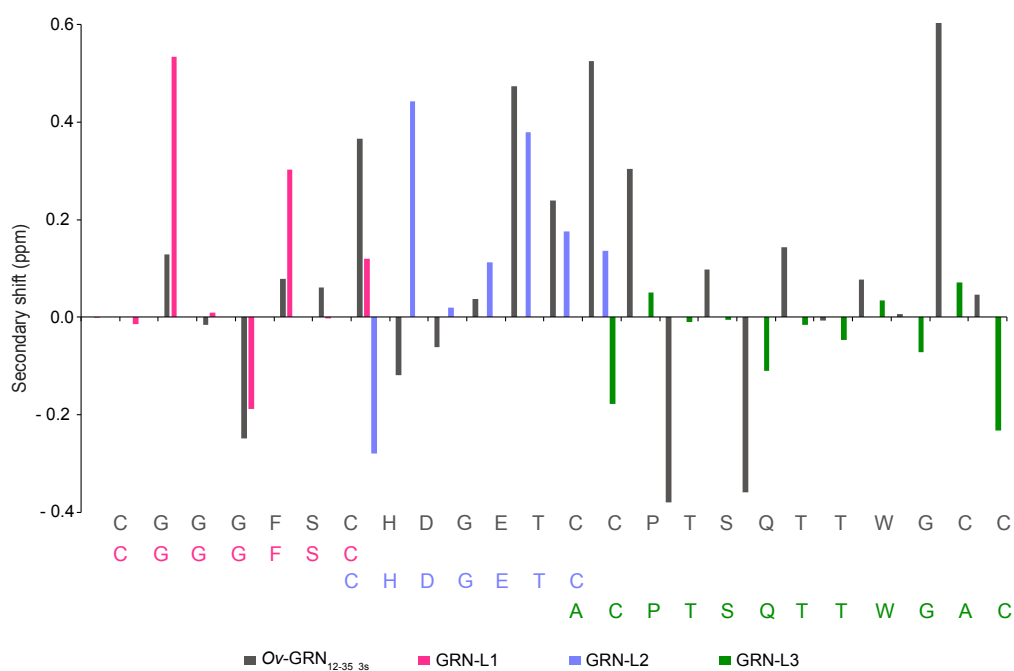
Figure 2-3 αH Secondary-shifts for *Ov*-GRN$_{12-35\_3s}$, GRN-L1, GRN-L2, and GRN-L3. The αH secondary shift was calculated by subtracting the random coil ^{1}H NMR chemical shifts previously reported by Wishart et al. [24] from the experimental αH chemical shifts. The sequence of the peptides is given at the bottom of the diagram. *Ov*-GRN$_{12-35\_3s}$ (BMRB code: 30232) is provided as a comparison.

2.4.3 Cell proliferation assay

The influence of the *Ov*-GRN$_{12-35\_3s}$ fragments on the proliferation of a human normal fibroblast cell line in real time was assessed using xCELLigence technology, and dose response curves were determined for the peptides. GRN-L2 resulted in a 13.6% and 13.7% increase ($p < 0.01$) in cell growth compared to the negative control peptide (a 21-residue tropomyosin derived peptide with the sequence of RAETGESKIVELEEELRVVGN) when applied to cells at concentrations of 200 nM and 1 μM respectively. A tropomyosin fragment, previously shown to have no activity in a cell proliferation assay was used as a negative control. Importantly, this peptide was produced in the laboratory using the same conditions to the *Ov*-GRN$_{12-35\_3s}$ fragments. GRN-L1 and GRN-L3 peptides tested at concentrations of 200 nM and 1 μM did not display statistically significant cell proliferation compared to the peptide control (Figure 2-4).
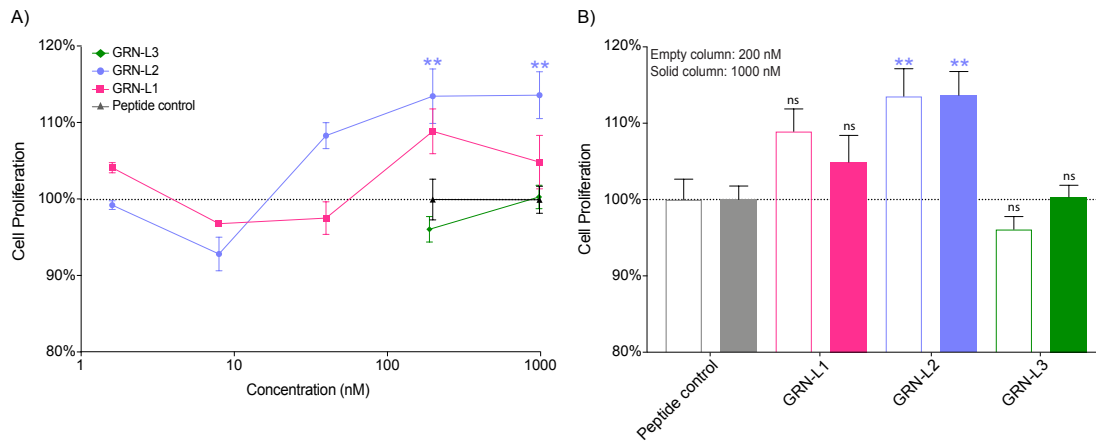
Figure 2-4 Cell proliferation assay of GRN-L1, GRN-L2, and GRN-L3. A) The average of six replicates is plotted with SEM bars. Cell index was measured four days after treatment. B) Day 4 relative proliferation from panel A. Empty columns and solid columns represent concentrations of 200 nM and 1 µM, respectively. Both panels: The proliferation rates relative to peptide control are plotted as mean ± SEM bars. Data were analysed by one-way ANOVA against peptide control (Not significant = ns, ** = P<0.01).

2.5 Discussion

Truncated forms of *Ov*-GRN-1 have significant potential as wound-healing agents but the key features for bioactivity are still poorly understood. Here we showed that a peptide corresponding to the second inter-cysteine loop of *Ov*-GRN-1 (GRN-L2) promotes cell proliferation of human fibroblasts. This sequence is unique to *Ov*-GRN-1 in the granulin family based on a BLAST search (https://blast.ncbi.nlm.nih.gov/Blast.cgi), indicating that this loop could be responsible, in part, for the potent wound-healing activity of *Ov*-GRN-1. The rate of proliferation enhancement is less than the levels observed in previous studies on *Ov*-GRN-1 and truncated *Ov*-GRN-1 peptides [15]. However, further studies are required to determine if this rate of proliferation correlates with wound healing in a mouse model, or ultimately with therapeutic potential in humans.

Based on NMR spectroscopy GRN-L2 appears to be unstructured in solution, suggesting that a well-defined structure is not critical for bioactivity. There are now numerous examples of peptides that lack well-defined structure, but still display bioactivity. These examples include disulfide-rich peptides, unstructured because of reduction or modification of the cysteine residues, such as human granulin B and the scorpion venom peptide chlorotoxin (CTX). The structural analysis of fully reduced human granulin B indicated that the peptide was intrinsically disordered, but it was able to induce an inflammatory response in SY-SH5Y human neuroblastoma cells by activation of NF-κB in a concentration dependent manner [25]. Furthermore, CTX contains 36 amino acids and inhibits glioma cell migration [26]. Despite the full-length peptide having a well-defined structure braced by four disulfide bonds, the reduced form of CTX adopts a random coil structure but inhibits cell migration with a similar potency as CTX [27].

Synthetic fragments of CTX with no disulfide bonds have been studied and a region corresponding to residues 29 to 36 of CTX, which is not structured in solution, inhibited cell migration, albeit at relatively high concentrations [26]. In a parallel study, this region of CTX was used to design blood-brain barrier shuttles, which included a peptide referred to as miniCTX3, a monocyclic lactam-bridge peptidomimetic containing residues 29–32 of CTX. This truncated analogue of CTX was capable of transporting nanoparticles across endothelial cells [28]. This study highlights the potential of identifying small bioactive regions of disulfide-rich peptides in drug design.

2.6 Conclusions

Overall, this study provides insight into understanding the important sequence requirements for *Ov*-GRN-1 bioactivity, and indicates that a small region of *Ov*-GRN-1, which is not structured in solution, has a significant influence on cell proliferation. It will be of interest to determine if the cell proliferation effects on human fibroblasts observed for GRN-L2 correlate with *in vivo* wound healing, as has previously been observed for two and three disulfide bond truncates of *Ov*-GRN-1.

2.7 Funding

2.8 Acknowledgements

2.9 References

1.	Recio, C., et al., *The potential therapeutic application of peptides and peptidomimetics in cardiovascular disease.* Front. Pharmacol., 2016. 7: p. 526-526.

2.	Takata, K., et al., *The ApoA-I mimetic peptide FAMP promotes recovery from hindlimb ischemia through a nitric oxide (NO)-related pathway.* Int. J. Cardiol., 2016. 207: p. 317-325.

3.	Swedberg, J.E., et al., *Substrate-guided design of a potent and selective kallikrein-related peptidase inhibitor for kallikrein 4.* Chem. Biol., 2009. 16(6): p. 633-43.

4.	Chan, L.Y., et al., *Engineering pro-angiogenic peptides using stable, disulfide-rich cyclic scaffolds.* Blood, 2011. 118(25): p. 6709-17.

5.	Poth, A.G., et al., *Cyclotides as grafting frameworks for protein engineering and drug design applications.* Biopolymers, 2013. 100(5): p. 480-91.

6.	Wang, C.K., et al., *Molecular grafting onto a stable framework yields novel cyclic peptides for the treatment of multiple sclerosis.* ACS Chem. Biol., 2014. 9(1): p. 156-63.

7.	Conibear, A.C., et al., *The cyclic cystine ladder of theta-defensins as a stable, bifunctional scaffold: A proof-of-concept study using the integrin-binding RGD motif.* ChemBioChem., 2014. 15(3): p. 451-9.

8.	Cunningham, A.D., et al., *Peptides and peptidomimetics as regulators of protein–protein interactions.* Curr. Opin. Struct. Biol., 2017. 44: p. 59-66.

9.	Chittoor, B., et al., *The single disulfide-directed beta-hairpin fold. Dynamics, stability, and engineering.* Biochemistry, 2017. 56(19): p. 2455-2466.

10.	Chee, S.M.Q., et al., *Functional display of bioactive peptides on the vGFP scaffold.* Sci. Rep., 2021. 11(1): p. 10127-10127.

11.	Hamley, I.W., *Small bioactive peptides for biomaterials design and therapeutics.* Chem. Rev., 2017. 117(24): p. 14015-14041.

12.	Reid, R.C., et al., *Downsizing a human inflammatory protein to a small molecule with equal potency and functionality.* Nat. Commun., 2013. 4: p. 2802.

13.	Smout, M.J., et al., *Carcinogenic parasite secretes growth factor that accelerates wound healing and potentially promotes neoplasia.* PLoS Pathog., 2015. 11(10): p. e1005209.

14. Papatpremsiri, A., et al., *Suppression of Ov-GRN-1 encoding granulin of opisthorchis viverrini inhibits proliferation of biliary epithelial cells.* Exp. Parasitol., 2015.

15. Bansal, P.S., et al., *Development of a potent wound healing agent based on the liver fluke granulin structural fold.* J. Med. Chem., 2017. 60(10): p. 4258-4266.

16. Takjoo, R., et al., *Folding of truncated granulin peptides.* Biomolecules, 2020. 10(8): p. 1152.

17. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: Development of a software pipeline.* Proteins, 2005. 59(4): p. 687-696.

18. Wüthrich, K., *NMR studies of structure and function of biological macromolecules (Nobel Lecture).* J. Biomol. NMR, 2003. 27(1): p. 13-39.

19. Güntert, P., *Automated NMR structure calculation with CYANA.* Methods Mol. Biol., 2004. 278: p. 353-378.

20. Shen, Y. and Bax, A., *Protein structural information derived from NMR chemical shift with the neural network program TALOS-N.* Methods Mol. Biol., 2015. 1260: p. 17-32.

21. Koradi, R., et al., *MOLMOL: A program for display and analysis of macromolecular structures.* J. Mol. Graph., 1996. 14(1): p. 51-55.

22. Gill, S.C. and von Hippel, P.H., *Calculation of protein extinction coefficients from amino acid sequence data.* Anal. Biochem., 1989. 182(2): p. 319-26.

23. Hilario, E.C., et al., *An improved method of predicting extinction coefficients for the determination of protein concentration.* PDA J. Pharm. Sci. Technol., 2017. 71(2): p. 127-135.

24. Wishart, D.S., et al., *^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J. Biomol. NMR, 1995. 5(1): p. 67-81.

25. Ghag, G., et al., *Fully reduced granulin-B is intrinsically disordered and displays concentration-dependent dynamics.* Protein Eng. Des. Sel., 2016. 29(5): p. 177-186.

26. Dastpeyman, M., et al., *A C-Terminal fragment of chlorotoxin retains bioactivity and inhibits cell migration.* Front. Pharmacol., 2019. 10: p. 250.

27. Ojeda, P.G., et al., *The role of disulfide bonds in structure and activity of chlorotoxin.* Future Med. Chem., 2014. 6(15): p. 1617-28.

28. Díaz-Perlas, C., et al., *From venoms to BBB-shuttles. MiniCTX3: a molecular vector derived from scorpion venom.* Chem. Comm., 2018. 54(90): p. 12738-12741.

Chapter 3. Folding of Truncated Granulin Peptides

3.1 Abstract

Granulins are a family of unique protein growth factors which are found in a range of species and have several bioactivities that include cell proliferation and wound healing. They typically contain six disulfide bonds, but the sequences, structures and bioactivities vary significantly. We have previously shown that an N-terminally truncated version of a granulin from the human liver fluke, *Opisthorchis viverrini*, can fold independently into a "mini-granulin" structure and has potent wound healing properties *in vivo*. The incorporation of a non-native third disulfide bond, with respect to the full-length granulin module, was critical for the formation of regular secondary structure in the liver fluke derived peptide. By contrast, this third disulfide bond is not required for a carp granulin-1 truncated peptide to fold independently. This distinction led us to explore granulins from the zebrafish model organism. Here we show that the mini-granulin fold occurs in a naturally occurring paragranulin (half-domain) from zebrafish, and is also present in a truncated form of a full-length zebrafish granulin, suggesting this structure might be a common property in either naturally occurring or engineered N-terminally truncated granulins and the carp granulin-1 folding is an anomaly. The *in vitro* folding yield is significantly higher in the naturally occurring paragranulin, but only the truncated zebrafish granulin peptide promoted the proliferation of fibroblasts consistent with a growth factor function, and therefore the function of the paragranulin remains unknown. These findings provide insight into the folding and evolution of granulin domains and might be useful in the elucidation of the structural features important for bioactivity to aid the design of more potent and stable analogues for the development of novel wound healing agents.

3.2 Introduction

Granulin proteins are prevalent throughout nature, being found in a wide variety of organisms [1]. They have potential as a novel class of therapeutic agents for treating a range of diseases and conditions, including chronic wounds [2-11]. Most granulins contain twelve cysteine residues, that form six disulfide bonds (Figure 3-1A, -1B), and are often expressed as progranulins containing multiple granulin modules. For example, the human and other mammalian granulin precursors (PGRN) contain a signal peptide and tandem repeats of seven-and-a-half granulin modules, as shown in Figure 3-1C for the human precursor [1, 12]. Granulin G contains only ten cysteine residues and the half granulin module, paragranulin, comprises only the first six cysteine residues [13].

Significant sequence, structural and bioactivity variations exist across the granulin family [1, 12]. The sequence conservation is primarily limited to the cysteine framework [12], with significant variation in the inter-cysteine loops. Structural diversity has been observed through the analysis of fish and human derived granulins. In contrast to the carp granulin-1 and zebrafish granulin AaE, which display relatively well-defined structures containing a four ß-hairpin stack [14, 15], several of the human granulin peptides display disordered regions. Human granulins A, C and F, for example, have well-defined N-terminal regions, but disordered C-terminal regions when produced recombinantly as single units, but it is unknown if this reflects the native granulin structures in biological systems. The bioactivity of the human granulins also varies, with granulin A shown to inhibit the proliferation of a breast cancer cell line, whereas human granulin F stimulates cell proliferation [16].

Figure 3-1 Granulin cysteine framework and precursor organization. A) Schematic representation of the highly conserved cysteine framework and disulfide bond pairing present in the granulin family. The cysteine residues are numbered using Roman numerals (I-XII). B) The solution structure of full-length zebrafish granulin AaE (PDB ID: 6cku). Disulfide bonds are shown in yellow and cysteine residues are labelled using Roman numerals. C) The order of the 7.5 granulin motifs in human progranulins, including A to G and a half-granulin motif named paragranulin (P) that includes the N-terminal six cysteines of the full length granulin. D) Schematic representation of zebrafish progranulin A. Granulin modules with no label were missing from the partial clone [15]. E) Granulin modules from zebrafish progranulin 1 including one full-length, and a half-granulin module named paragranulin (P) that comprises six cysteines of the full length granulin [15]. Target peptides of interest in the current study are shown with dashed circles.

Truncated forms of granulins can fold into well-defined structures with only two or three disulfide bonds and, in some cases, have potent bioactivities. Peptides corresponding to the N-terminal region of carp granulin-1, an analogue of the N-terminal region of human granulin A (hGRNA) and a plant cysteine protease (oryzain ß) analogue, can all fold independently into ß-sheet structures with only two disulfide bonds [17-20]. However, we have recently shown that analogous peptides from the N-terminus of a granulin derived from the human liver fluke Opisthorchis viverrini (*Ov*-GRN-1) do not form a regular secondary structure [2, 3]. By contrast, an *Ov*-GRN-1 derived peptide, (*Ov*-GRN$_{12-35\_3s}$), that contains three disulfide bonds, including a disulfide bond (Cys IV-Cys VI) not present in the full length granulins, can fold independently with a ß-sheet structure. Interestingly, both the two and three-disulfide bond containing *Ov*-GRN-1 derived peptides display potent *in vivo* wound healing properties, similar to the full-length *Ov*-GRN-1 [2, 3]. Indeed, an analogue of the *Ov*-GRN-1 peptide with three disulfide bonds has significant potential as a wound healing agent and is as potent as Regranex [2, 3], a recombinant human platelet-derived growth factor, but can be made with synthetic methods rather than recombinant technologies. Regranex is currently the only approved biologic on the market for the treatment of chronic wounds [21, 22].

The determination of whether granulins with sequences that differ significantly from *Ov*-GRN-1 can also fold independently while incorporating the Cys IV-Cys VI disulfide bond, is a critical step in advancing the understanding of the folding of this structural motif. Such an advance might have significant implications in the field of disulfide-rich peptide/protein structure and folding pathway predictions using molecular simulations [23]. On a more applied note, further understanding of the structure/function relationships of granulin-derived peptides might be useful in the design of wound healing agents for treating chronic wounds, such as diabetic foot ulcers.

To address the question regarding the folding of the N-terminal regions of granulins, we have analysed the structures of two truncated forms of zebrafish AaE, as well as a paragranulin from zebrafish progranulin 1. The peptides of interest in this study are highlighted with dashed circles in Figure 3-1D, E. In addition to a well-defined three-dimensional structure, zebrafish granulin AaE [15, 24] promotes the

survival of neuronal cells [15]. However, the structure or bioactivity of the half-module granulin from zebrafish has not been determined. Indeed, there have been limited studies on the naturally occurring paragranulins (half-modules). We show that incorporating a Cys IV-Cys VI disulfide bond into the N-terminal region of granulin might be a common feature across the family, but that the proliferative bioactivity of human cell lines is highly dependent on the primary structure.

3.3 Experimental section

3.3.1   Peptide synthesis

The synthesis of truncated granulin peptides was undertaken by manual solid-phase peptide synthesis using standard 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry. All peptides were assembled on 2-chlorotrityl chloride resin (Auspep, Tullamarine, Australia). The Fmoc amino acids (Auspep, Tullamarine, Australia) were activated using *O*-(1*H*-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, HCTU, (Iris Biotech GMBH, Marktredwitz, Germany) and coupled on the resin with DIPEA/DMF. Peptides were cleaved by a cleavage cocktail, including 95% trifluoracetic acid (TFA)/2.5% triisopropylsilane (TIPS)/2.5% H_2O (v/v/v). TFA was removed with a stream of nitrogen gas. Peptides were then precipitated in ice-cold diethyl ether and dissolved in 50% acetonitrile: 50% H_2O:0.1% TFA (v/v) and subsequently lyophilised. The peptides were synthesised without isotope labelling and the subsequent NMR spectra were recorded with natural abundance.

3.3.2   Purification

The crude peptides were purified using reversed-phase high performance liquid chromatography (RP-HPLC) on a C_{18} preparative column (Phenomenex Jupiter 250 × 21.2 mm, 10 μm, 300 Å) with gradients of solvent B (90% acetonitrile/10% H_2O/0.045% TFA (v/v/v) and solvent A (H_2O/0.05% TFA (v/v)). The masses of collected fractions were determined using a 5800 MALDI TOF/TOF mass spectrometer (SCIEX, Foster City, CA, USA).

3.3.3 Disulfide bond formation

The disulfide bonds of the truncated peptides were formed using non-regioselective oxidation in a solution of peptide in 0.1 M ammonium bicarbonate buffer (pH 8–8.2) and 5 mM reduced glutathione at room temperature. The oxidation reaction was left for 48 h prior to acidification with TFA (30–50 µL TFA to a 10 mL oxidation reaction) and loaded on a C_{18} preparative HPLC column with a flow rate of 6 mL/min. Fractions were collected and the peptide mass was analysed using a SCIEX 5800 MALDI TOF/TOF spectrometer [2, 3] (Appendix 1).

3.3.4 NMR spectroscopy

NMR samples were prepared from unlabelled purified peptide (0.2 mM) in 90% H_2O/10% D_2O. All NMR spectra were recorded on a 600 MHz AVANCE III NMR spectrometer (Bruker, Karlsruhe, Germany), equipped with a 5 mm TCI cryoprobe. Two-dimensional ^{1}H-^{1}H TOCSY, ^{1}H-^{1}H NOESY, ^{1}H-^{1}H DQF-COSY and ^{1}H-^{13}C HSQC spectra were acquired at 290 K. Spectra were recorded using an interscan delay of 1 s. NOESY spectra were acquired with mixing times of 200 ms, and TOCSY spectra were acquired with isotropic mixing periods of 80 ms. All spectra were processed using Bruker TopSpin (Version 3.5pl7) and assigned using CCPNMR analysis 2.1, based on the approach described in Wüthrich *et al.* [25, 26]. More than 90% of the protons were assigned and 70% of the C_α and C_β atoms were unambiguously assigned. Amide temperature coefficients were calculated based on TOCSY spectra directly referenced to DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) and recorded at temperatures ranging from 290 K to 305 K using the method outlined in Cierpicki *et al.* [27]. Hydrogen bond restraints (Appendix 2) were included in the calculations based on hydrogen bonds identified in preliminary structures and amide protons with temperature coefficients more positive than – 4 ppb/K.

The αH secondary shifts were determined by subtracting random coil ^{1}H NMR chemical shifts from the experimental αH chemical shifts [28]. Root mean square deviation (RMSD) values were calculated relative to a mean structure using MOLMOL [29]. The structures and chemical shifts have been deposited into the Protein Data Bank and the Biological Magnetic Resonance Data Bank (ZF-N$_{24\_3s}$— PDB ID: 7JIA, BMRB ID: 30780; ZF-para$_{\_3s}$—PDB ID: 7JIY, BMRB ID: 30781).

3.3.5 Structure calculations

The three-dimensional structures of peptides were calculated using the CYANA program, based on automated assignment of the NOEs [30]. Torsion-angle restraints predicted by TALOS-N were used in the structure calculations [31]. Structures were visualised using MOLMOL [29].

3.3.6 Mammalian cell culture

The 1BR.3.GN human skin normal fibroblast cell line was obtained from a European Collection of Authenticated Cell Cultures (ECACC, Porton Down, UK). The 1BR.3.GN cells were grown and maintained in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) (Life Technologies, Melbourne, Australia) containing $1 \times$ antibiotic/antimycotic and $1 \times$ GlutaMAX, supplemented with 10% foetal bovine serum (FBS) (Gibco, Glasgow, Scotland) at 37 °C and 5% CO_2. Cell proliferation assays were performed with DMEM/F12 media, supplemented with 10% FBS.

3.3.7 Cell proliferation monitoring in real time using xCELLigence

Cells were seeded at 5000 cells/well in 170 µL of complete media in E-plates (ACEA Biosciences, San Diego, CA, United States) and grown overnight and monitored with an xCELLigence SP system (ACEA Biosciences), which monitors cellular events in real time by measuring electrical impedance across gold microelectrodes integrated into the base of tissue culture plates. Cells were washed three times with PBS prior to addition of 150 µL of low nutrient media and incubated for a minimum of 6 h before further treatment. Treatments were prepared at $8.5 \times$ concentration and added to each well in a total volume of 20 µL. The xCELLigence system recorded cell indexes at intervals of 1 h for 5-6 days following treatment. Readings for the cell index were normalized prior to treatment, and cell proliferation ratios represent the relative numbers of cells compared to control cells at day 4.

Comparisons of induction of cell proliferation in response to treatments were accomplished using the Two-Way ANOVA test with Dunnett's multiple comparison correction, using GraphPad Prism 8.0.

3.4 Results

3.4.1 Design and synthesis of zebrafish granulin peptides

To analyse the folding of zebrafish granulins, three peptides were designed: ZF-N$_{24\_2s}$, ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$. The first peptide contains two disulfide bonds, and the latter two peptides contain three disulfide bonds. The sequences and sources of the synthetic peptides are shown in Table 3-1. An Ov-GRN-1 derived peptide is given in the table to highlight the differences in the inter-cysteine loop sequences.

Table 3- 1 Granulin-derived peptide sequences.

Peptide	Source	Sequence
ZF-N$_{24\_3s}$	Zebrafish granulin AaE	**C**GGGF–S**C**HDGET**CC**PTSQTTWG**CC**
ZF-N$_{24\_2s}$	Zebrafish granulin AaE	**C**GGGF–S**C**HDGET**C**APTSQTTWG**C**A
ZF-para$_{\_3s}$	Zebrafish paragranulin	**C**EGNFY–**C**PAEKF**CC**KTRTGQWG**CC**
Ov-GRN$_{12-35\_3s}$#	Ov-GRN-1	**C**PDPVYT**C**RPGQT**CC**RGLHG–YG**CC**

Hyphen (–) indicates a sequence gap. Cysteine residues are shown in bold and red. # The sequence of the previously studied Ov-GRN$_{12-35\_3s}$ [2], is cited accordingly and presented in this table for comparison purposes. The sequences of the peptides synthesised in the current study are highlighted in blue. The cysteine residues involved in the non-native (Cys IV-Cys VI) disulfide bond in zebrafish granulin AaE, and by homology in Ov-GRN-1 (the 3D structure of Ov-GRN-1 has not been determined), have been underlined. The sequences are derived from the precursor proteins of zebrafish granulin A (Q90ZD0), zebrafish progranulin-1 (Q8QGN9) and Ov-GRN-1 (B8XSI4).

Peptides were synthesised using Fmoc chemistry on 2-chlorotrityl resin without selective protection of the cysteine residues, and the disulfide bonds formed by air oxidation in ammonium bicarbonate (0.1 M) and reduced glutathione (5 mM) at room temperature. The analytical RP-HPLC trace of all three oxidation reactions contain several peaks, including a sharp peak that eluted earlier than the other peaks and corresponds to a fully oxidised isomer, based on mass spectrometry (Appendix 3). The early eluting peaks were purified (>95% purity based on analytical RP-HPLC) and the structures analysed using NMR spectroscopy. The yields of the early eluting peaks varied amongst the peptides. A comparison of the folding of the peptides is shown in Figure 3-2, highlighting the higher yield obtained for the naturally occurring paragranulin compared to the truncated version of zebrafish granulin AaE (ZF-N$_{24\_3s}$).

3.4.2 Structural analysis with NMR spectroscopy

The one-dimensional proton NMR spectra of ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ display large chemical shift dispersion in the amide region, and analysis of the secondary shifts indicates that the peptides contain ß-sheet structure based on the consecutive positive shifts (Figure 3-3). The three-dimensional structures of both peptides were calculated using CYANA based on distance and dihedral angle restraints derived from one- and two-dimensional homonuclear and heteronuclear NMR experiments. Calculation of the structures without disulfide bond restraints indicated that the most likely connectivity was Cys I-Cys III, Cys II-Cys V and Cys IV-Cys VI, consistent with the previous study on *Ov*-GRN-1 peptides [2, 3]. Calculation of the structures with all 15 possible disulfide bond connectivities also indicated that this connectivity had the lowest target function (Appendix 4). Subsequent structures were calculated using restraints for this disulfide connectivity. Hydrogen bond restraints were included based on analysis of the preliminary structures and temperature coefficients (Appendix 5 and 6). The refinement statistics are given in Table 3-2.
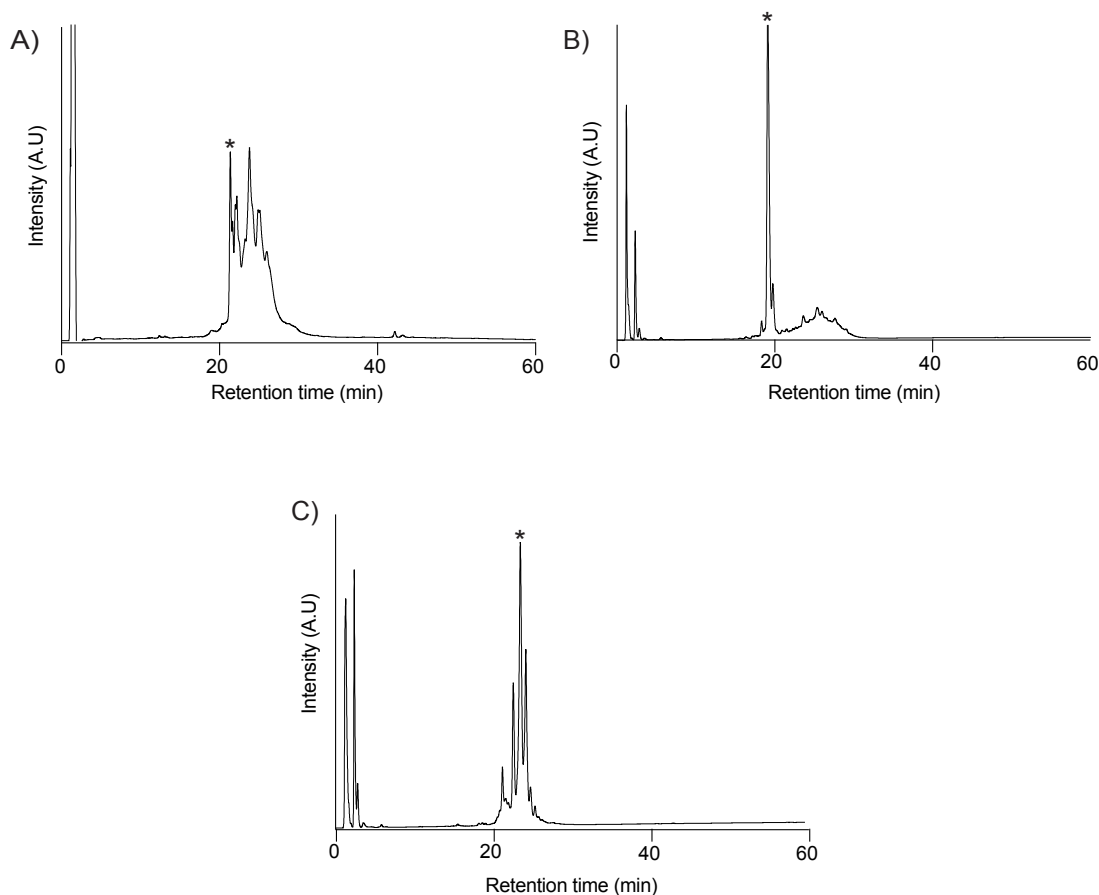
Figure 3-2 RP-HPLC analysis of the oxidation reaction of granulin N-terminal truncated analogues. A) ZF-N$_{24\_3s}$, B) ZF-para$_{\_3s}$ and C) ZF-N$_{24\_2s}$. Analytical RP-HPLC was carried out using a Phenomenex Jupiter 4 μm Proteo column (C12, 150 × 2.00 mm, 10 μm, 90 Å) using a gradient of 0–60% solvent B (Solvent A: 99.95% H$_2$O: 0.05% TFA; Solvent B: 90% acetonitrile: 10% H$_2$O: 0.045% TFA) over 60 min with a flow rate of 0.4 mL/min. The absorbance was monitored at 214 nm. The early eluting sharp peaks (highlighted by asterisks (*)) were purified for subsequent analyses.

In contrast to the three-disulfide bond containing peptides, ZF-N$_{24\_2s}$ did not display significant dispersion in the amide region of the one-dimensional proton spectrum. Despite this lack of dispersion in the amide region, the backbone protons and most of the side chain protons could be assigned based on two-dimensional spectra. Analysis of the secondary shifts (Figure 3-3) indicates that the peptide does not contain regular secondary structure, as the majority of the shifts are within 0.1 ppm of the random coil values.

The major element of secondary structure in ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ is a ß-hairpin comprising residues 13–23. An NOE between the α-protons of Cys 13 and Cys 23 is clearly present in both spectra, and represents a key restraint in defining the alignment of the ß-strands. The structures appear to be stabilised by hydrogen bonds consistent with amide protons, having temperature coefficients more positive than − 4.6 ppb/K [27]. The disulfide bonds have the connectivity Cys I-Cys III, Cys II-Cys V and Cys IV-Cys VI; the two former disulfide bonds are equivalent to those present in the full-length structures of the well characterised granulins [14, 18]. In the full-length structures Cys IV is bonded to Cys VII and Cys VI is bonded to Cys IX (refer to Figure 3-1A for disulfide bond connectivity) [15, 16, 32].
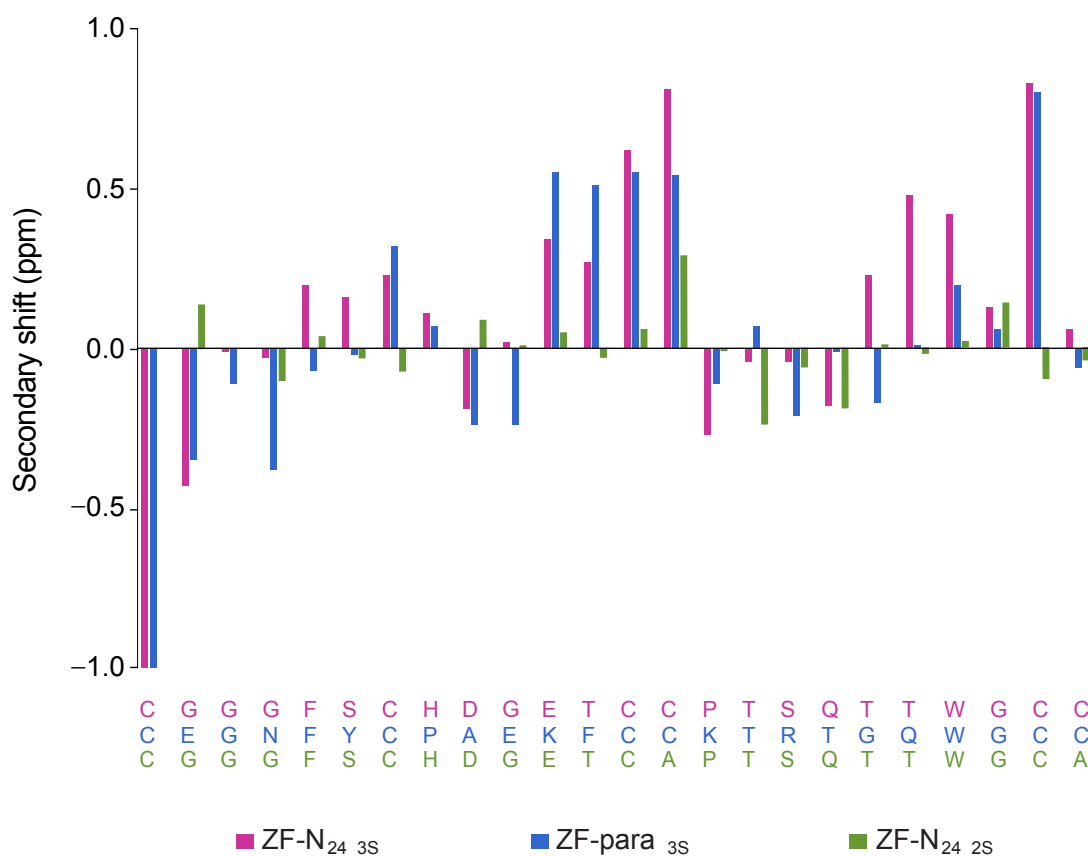
Figure 3-3 αH secondary-shift comparison for truncated granulin analogues. The αH secondary shifts were calculated by subtracting the random coil ^{1}H NMR chemical shifts previously reported by Wishart et al. [28] from the experimental αH chemical shifts. The sequences of ZF-N$_{24\_3s}$ (violet), ZFpara$_{\_3s}$ (blue) and ZF-N$_{24\_2s}$ (green) are given at the bottom of the diagram and are colour coded with respect to the graph.

Table 3- 2 Structural statistics.

Experimental Restraints	ZF-N$_{24\_3s}$	ZF-para$_{\_3s}$
Interproton distance restraints		
Intra-residue, $\|i - j\| = 0$	41	72
Sequential, $\|i - j\| = 1$	65	88
Medium range, $1 < \|i - j\| < 5$	12	36
Long range, $\|i - j\| >= 5$	31	64
Disulfide-bond restraints (3 restraints per bond)	9	9
Dihedral-angle restraints	27	32
Hydrogen bond restraints (2 restraints per bond)	4	8
Root Mean Square Deviations from Mean Coordinate Structure (Å)		
Backbone atoms	1.29 ± 0.38	0.63 ± 0.20
All heavy atoms	1.81 ± 0.41	1.31 ± 0.35
Ramachandran Statistics		
% In most favoured region	80.6	96.4
% Residues in additionally allowed regions	19.4	3.6

Comparison of the structures of ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ with the structure of a truncated form of *Ov*-GRN-1 (*Ov*-GRN$_{12-35\_3s}$) [2, 3], also containing the first six cysteine residues, indicates the peptides have similar overall folds despite sequence differences (37.5% sequence identity to *Ov*-GRN$_{12-35\_3s}$; the sequence identity is primarily related to the conserved cysteine residues) (Figure 3-4). A superposition of ZF-N$_{24\_3s}$ with the structure of zebrafish granulin AaE is given in Figure 3-5 and

highlights the similarity between the truncated and full-length versions, despite the differences in disulfide bonds.
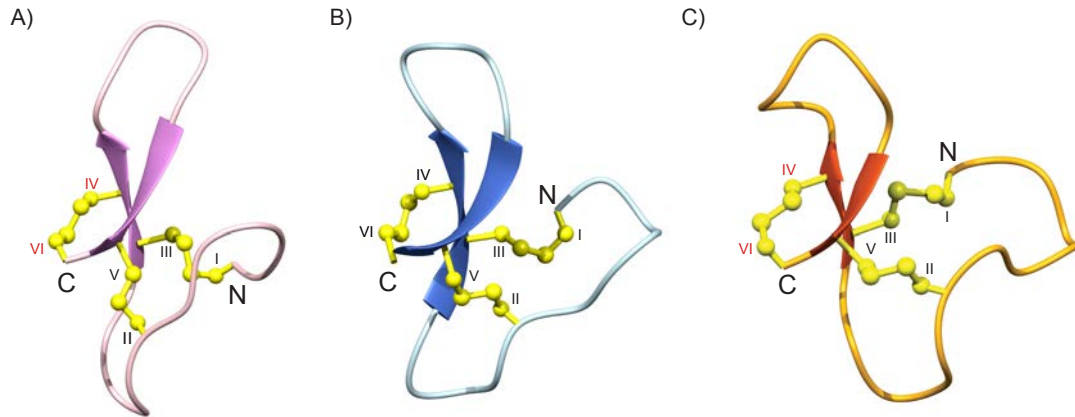
Figure 3-4 Three-dimensional structures of A) ZF-N$_{24\_3s}$, B) ZF-para$_{\_3s}$ and C) *Ov*-GRN$_{12\text{-}35\_3s}$ N-terminal truncated peptides. The ß-hairpins are shown as arrows, and ZF-N$_{24\_3s}$ is shown in violet, ZF-para$_{\_3s}$ in blue and *Ov*-GRN$_{12\text{-}35\_3s}$ (PDB: 5UJG) in orange. Disulfide bonds are shown in yellow and cysteine residues are labelled using Roman numerals. The cysteine residues involved in forming the disulfide bond (IV-VI) not present in the full-length granulin structures of zebrafish granulin AaE and by homology *Ov*-GRN-1, are labelled with red letters (A, C). The figure was prepared using MOLMOL [29].
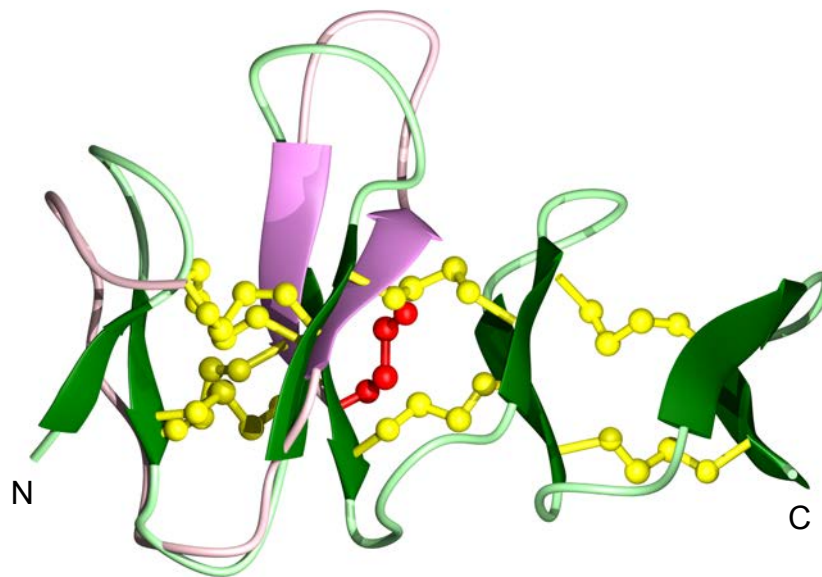
Figure 3-5 Superposition of the structures of the ZF-N$_{24\text{-}3s}$ and zebrafish AaE.
Superposition of the structures of the ZF-N$_{24\text{-}3s}$ and zebrafish AaE. The ß-hairpins are shown as arrows—in violet for ZF-N$_{24\text{-}3s}$ (PDB ID: 7JIA), and dark green for zebrafish AaE (PDB ID: 6cku). Disulfide bonds are shown in yellow. The cysteine residues involved in forming the non-native disulfide bond (IV-VI) in ZF-N$_{24\text{-}3s}$ is shown in red. The figure was prepared using MOLMOL [29].

3.1.1 Cell proliferation assay

The effect of the truncated peptides on the growth of 1BR.3.GN fibroblasts was evaluated using an xCELLigence system. The 1BR.3.GN is a human fibroblast skin cell line derived from transformed normal fibroblasts. Cells were cultured with two concentrations of the peptides, 200 nM and 1 µM. ZF-N$_{24\_3s}$ significantly promoted the proliferation of fibroblasts compared to the negative control peptide (a 19-residue peptide from tropomyosin) (Figure 3-6). The proliferation rate for ZF-N$_{24\_3s}$ was 113.6% ($p < 0.05$), relative to the control peptide at day 4 after treatment. By contrast, the effect of ZF-para$_{\text{-}3s}$ or ZF-N$_{24\_2S}$ on fibroblast cell proliferation was not statistically significant at either 200 nM or 1 µM (Figure 3-6).
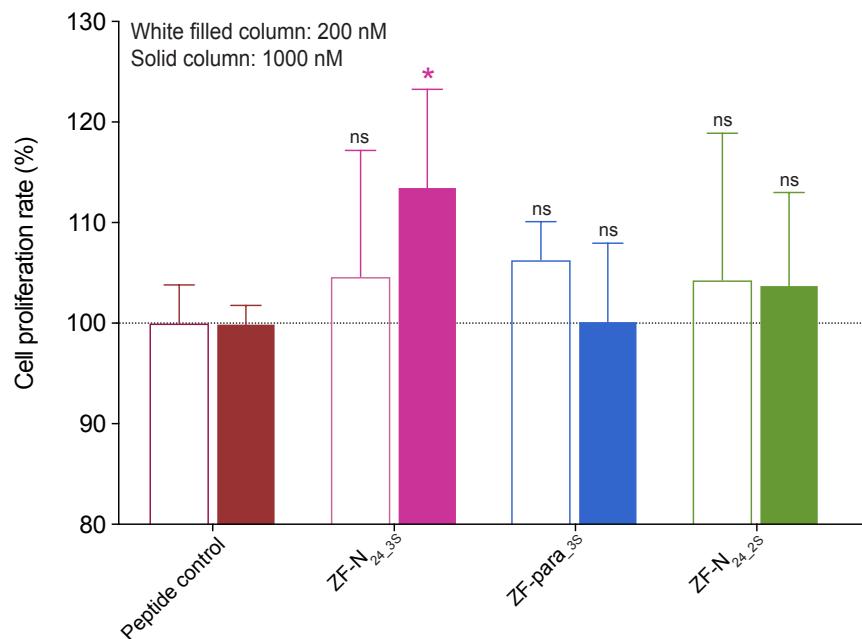
Figure 3-6 Cell proliferation analysis of granulin peptides using xCELLigence technology. Fibroblast cells were treated with peptides at concentrations of 200 nM and 1 μM. Data represent mean $\pm$ SD of three independent experiments. Data were analysed by one-way ANOVA against peptide control: ns—not significant; *—$p <$ 0.05. Cell index was measured at day 4 after treatment. White-filled columns and solid columns represent concentrations of 200 nM and 1 μM, respectively.

3.5 Discussion

Folding of disulfide rich peptides still represents one of the great challenges in structure prediction and simulation [33, 34]. Experimental analysis of the folding processes has to underlie these simulations, and analysis of the granulin structural framework present in the vast majority of organisms is likely to provide valuable insight for future computational studies. The potential as a wound healing agent might also be enhanced as a result of greater understanding of the structure/function relationships. In the current study we show that the N-terminal half of a range of granulins can fold independently of the C-terminal region via the incorporation of a Cys IV-Cys VI disulfide bond, indicating this is a common feature in this family. It appears likely that the naturally occurring half-motifs are optimised for folding compared to engineered versions derived from full-length granulin sequences, and that the Cys IV-Cys VI disulfide bond likely constitutes a native disulfide bond. However,

the structure of a paragranulin on native material will be required to confirm this suggestion.

The secondary shifts of the granulin peptides ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ were similar (Figure 3-3), indicating they have similar structures, and this was subsequently confirmed by the determination of the three-dimensional structures of the peptides (Figure 3-4). Both peptides contain a ß-hairpin and similar arrangement of the disulfide bonds to *Ov*-GRN$_{12-35\_3s}$ [2, 3], the first granulin peptide shown to adopt this conformation with three disulfide bonds. The common fold we have identified for the half-granulin domain could represent an evolutionary ancestor of the full-length scaffold, hence the propensity for a range of granulins to adopt this fold in the truncated form even though the disulfide connectivity differs in the full-length proteins.

The more efficient *in vitro* folding of the paragranulin-derived peptide, ZF-para$_{\_3s}$, relative to the engineered version (ZF-N$_{24\_3s}$), poses questions regarding the sequence differences between the peptides and the subsequent folding pathways. ß-hairpins can represent the folding nucleus with hydrophobic collapse being an integral part of folding [35]. However, comparison of the sequences of the two zebrafish peptides indicates very little difference in the hydrophobicity of the amino acids involved in the ß-hairpin loop and therefore it is difficult to speculate the reason for the folding differences.

The structural fold present in these truncated granulin peptides has recently been referred to as the "mini-granulin" fold and is present in several peptides [36], including conotoxins ϕ-MiXXVIIA [37] and H-Vc7.2. The fold contains two conserved disulfide bonds, which correspond to Cys I-Cys III and Cys II-Cys V in ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$. Both ϕ-MiXXVIIA and H-Vc7.2 contain additional disulfide bonds, but they are distinct from the additional Cys IV-Cys VI bond present in ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$. However, there are examples of this motif containing only two disulfide bonds, such as the truncated form of carp granulin-1, which forms a well-defined ß-sheet-containing structure [16, 17], indicating that the mini-granulin fold is not always dependent on additional disulfide bonds.

The ß-sheet structure present in the carp granulin-1 truncated peptide is in contrast to the results obtained in the current study on ZF-N$_{24\_2s}$, which has no regular

secondary structure. Previous studies have shown that, when only Cys I-Cys III and Cys II-Cys V are present in a truncated form of human granulin A, the peptide does not form a well-defined structure [17]. Similarly, a truncated form of *Ov*-GRN-1 with only two disulfide bonds does not form ß-sheet structure, whereas the three-disulfide bond form does [2, 3]. The results to date suggest that the folding of the carp granulin-1 peptide could be an anomaly, but further study is required to determine residues required for effective folding.

In addition to providing insight into granulin folding, the current study also analysed the biological activity of the truncated N-terminal peptides. ZF-N$_{24\_3s}$ displayed activity in the cell proliferation assay, whereas ZF-para$_{\_3s}$ was not active. However, ZF-N$_{24\_3s}$ was not as active as the parasite-derived granulin peptide *Ov*-GRN$_{12-35\_3s}$ [2]. The cells chosen for study in the current study reflect the potential of granulin peptides as wound healing agents, but it is possible the zebrafish derived peptides would show greater potency on fish-derived cell lines. There are now several examples where it has been shown that well-defined 3D structures are not necessary for bioactivity [38], including human granulin B, which has been characterised as an intrinsically disordered protein that modestly induces NF-κB activation in SY-SH5Y human neuroblastoma cells [39]. The current study highlights the importance of the primary structure, given that all peptides contain a ß-hairpin, but the activity varies significantly.

3.6 Conclusions

Overall, this study has shown that truncated granulins can incorporate a disulfide bond not-present in the full-length module, and this bond plays a pivotal role in stabilizing a well-defined structure in the N-terminal region of truncated granulin peptides. Despite a common structural motif, the bioactivity of these engineered peptides varies, and further mutational studies will be required to determine important regions/residues for bioactivity. Ultimately, the identification of the molecular target is likely to provide the greatest advance in understanding the structure/function relationships of granulin peptides.

3.7 Funding

3.8 Acknowledgments

3.9 References

1.   Palfree, R.G.E., et al., *The evolution of the secreted regulatory protein progranulin.* PLoS One, 2015. 10(8): p. e0133749.

2.   Bansal, P.S., et al., *Development of a potent wound healing agent based on the liver fluke granulin structural fold.* J. Med. Chem., 2017. 60(10): p. 4258-4266.

3.   Dastpeyman, M., et al., *Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing.* J. Med. Chem., 2018. 61(19): p. 8746-8753.

4.   Smout, M.J., et al., *Carcinogenic parasite secretes growth factor that accelerates wound healing and potentially promotes neoplasia.* PLoS Pathog., 2015. 11(10): p. e1005209.

5.   Botelho, M.C., et al., *Wound healing and cancer progression in Opisthorchis viverrini associated cholangiocarcinoma.* Parasitol. Res., 2016. 115(7): p. 2913-4.

6.   Ding, H., et al., *Progranulin derived engineered protein Atsttrin suppresses TNF-α-mediated inflammation in intervertebral disc degenerative disease.* Oncotarget, 2017. 8(65).

7.   Qiao, G., et al., *Granulin A synergizes with cisplatin to inhibit the growth of human hepatocellular carcinoma.* Int. J. Mol. Sci., 2018. 19(10): p. 3060.

8.   Chitramuthu, B.P., et al., *Progranulin: a new avenue towards the understanding and treatment of neurodegenerative disease.* Brain, 2017. 140(12): p. 3081-3104.

9.   Jian, J., et al., *Progranulin recruits HSP70 to β-glucocerebrosidase and is therapeutic against Gaucher disease.* EBioMedicine, 2016. 13(C): p. 212-224.

10.   Abella, V., et al., *Progranulin as a biomaker and potential therapeutic agent.* Drug Discov. Today, 2017. 22(10): p. 1557-1564.

11.   Pogonowska, M., et al., *The role of progranulin (PGRN) in the modulation of anti-inflammatory response in asthma.* Cent. Eur. J. Immunol., 2019. 44(1): p. 91-101.

12.   Dastpeyman, M., et al., *Folding of granulin domains.* Pept. Sci., 2018. 110(3).

13.   Ong, C.H.P. and Bateman, A., *Progranulin (granulin-epithelin precursor, PC-cell derived growth factor, acrogranin) in proliferation and tumorigenesis.* Histol. Histopathol., 2003. 18(4): p. 1275-1288.

14.   Hrabal, R., et al., *The hairpin stack fold, a novel protein architecture for a new family of protein growth factors.* Nat. Struct. Biol., 1996. 3(9): p. 747-751.

15.   Wang, P., et al., *Structure dissection of zebrafish progranulins identifies a well-folded granulin/epithelin module protein with pro-cell survival activities.* Protein Sci., 2018. 27(8): p. 1476-1490.

16.   Tolkatchev, D., et al., *Structure dissection of human progranulin identifies well-folded granulin/epithelin modules with unique functional activities.* Protein Sci., 2008. 17(4): p. 711-724.

17.   Tolkatchev, D., et al., *Design and solution structure of a well-folded stack of two β-hairpins based on the amino-terminal fragment of human granulin A.* Biochemistry, 2000. 39(11): p. 2878-2886.

18.   Vranken, W.F., et al., *A 30-residue fragment of the carp granulin-1 protein folds into a stack of two β-hairpins similar to that found in the native protein.* J. Pept. Res., 1999. 53(5): p. 590-597.

19.   Vranken, W.F., et al., *Solution structures of a 30-residue amino-terminal domain of the carp granulin-1 protein and its amino-terminally truncated 3-30 subfragment: Implications for the conformational stability of the stack of two β-hairpins.* Proteins, 2002. 47(1): p. 14-24.

20.   Tolkatchev, D., et al., *A peptide derived from the c-terminal part of a plant cysteine protease folds into a stack of two β-hairpins, a scaffold present in the emerging family of granulin-like growth factors.* J. Pept. Res., 2001. 57(3): p. 227-233.

21.   Niezgoda, J.A., et al., *Randomized clinical trial comparing OASIS Wound Matrix to Regranex Gel for diabetic ulcers.* Adv. Skin Wound Care, 2005. 18(5 Pt 1): p. 258-266.

22.   Chan, R.K., et al., *Effect of recombinant platelet-derived growth factor (Regranex®) on wound closure in genetically diabetic mice.* J. Burn Care Res., 2006. 27(2): p. 202-205.

23.   Georgoulia, P.S. and Glykos, N.M., *Molecular simulation of peptides coming of age: Accurate prediction of folding, dynamics and structures.* Arch. Biochem. Biophys., 2019. 664: p. 76-88.

24.   Wang, P.F., *Structure genomics of zebrafish granulins.* 2004, McGill University, Montreal.

25.   Wüthrich, K., *NMR studies of structure and function of biological macromolecules (Nobel Lecture).* J. Biomol. NMR, 2003. 27(1): p. 13-39.

26. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: Development of a software pipeline.* Proteins, 2005. 59(4): p. 687-696.

27. Cierpicki, T. and Otlewski, J., *Amide proton temperature coefficients as hydrogen bond indicators in proteins.* J. Biomol. NMR, 2001. 21(3): p. 249-261.

28. Wishart, D.S., et al., *^1H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J. Biomol. NMR, 1995. 5(1): p. 67-81.

29. Koradi, R., et al., *MOLMOL: A program for display and analysis of macromolecular structures.* J. Mol. Graph., 1996. 14(1): p. 51-55.

30. Güntert, P., *Automated NMR structure calculation with CYANA.* Methods Mol. Biol., 2004. 278: p. 353-378.

31. Shen, Y. and Bax, A., *Protein structural information derived from NMR chemical shift with the neural network program TALOS-N.* Methods Mol. Biol., 2015. 1260: p. 17-32.

32. Bhandari, V. and Bateman, A., *Structure and chromosomal location of the human granulin gene.* Biochem. Biophys. Res. Commun., 1992. 188(1): p. 57-63.

33. Yang, J., et al., *Accurate disulfide-bonding network predictions improve ab initio structure prediction of cysteine-rich proteins.* Bioinformatics, 2015. 31(23): p. 3773-3781.

34. Paul George, A.A., et al., *Insights into the folding of disulfide-rich μ-conotoxins.* ACS Omega, 2018. 3(10): p. 12330-12340.

35. Lewandowska, A., et al., *β-hairpin-forming peptides; models of early stages of protein folding.* Biophys. Chem., 2010. 151(1): p. 1-9.

36. Nielsen, L.D., et al., *The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework.* J. Biol. Chem., 2019. 294(22): p. 8745-8759.

37. Jin, A.H., et al., *Conotoxin Φ-MiXXVIIA from the superfamily G2 employs a novel cysteine framework that mimics granulin and displays anti-apoptotic activity.* Angew. Chem., 2017. 129(47): p. 15169-15172.

38. Uversky, V.N., *A decade and a half of protein intrinsic disorder: Biology still waits for physics.* Protein Sci., 2013. 22(6): p. 693-724.

39. Ghag, G., et al., *Fully reduced granulin-B is intrinsically disordered and displays concentration-dependent dynamics.* Protein Eng. Des. Sel., 2016. 29(5): p. 177-186.

Chapter 4. The Granulin Disulfide Framework: Distinctions Between the N- and C-terminal Regions

4.1 Abstract

Granulins are a family of growth-factor proteins with a critical role in infection and inflammation. They have promising potential therapeutic application in a wide range of medical conditions including neurodegenerative disorders, rheumatoid arthritis, and wound repair. The granulin framework is characterised by 12 cysteine residues, but N-terminally truncated analogues, comprising two or three disulfide bonds, can fold independently and display bioactivity. The aim of this study was to determine if the C-terminal half of granulins can fold independently and display bioactivity, analogous to the N-terminal half. A truncated analogue of zebrafish granulin AaE, comprising two disulfide bonds, was synthesised using solid-phase peptide synthesis. The structure of the peptide was analysed using NMR spectroscopy and the impact on cell proliferation determined using biosensor technology. We show that the C-terminal region of granulin AaE has a disordered structure with a centrally located, well-defined β-turn, and has no inhibitory effect on the growth of human skin fibroblast cells. The well-defined β-turn in truncated granulin AaE comprises the sequence PAGY, which has previously been shown to have antioxidant properties. Interestingly, the sequence is highly conserved in granulins suggesting a role in either bioactivity or structure. The disordered structure in ZF-C$_{20\_2s}$ indicates that the folding of the N- and C-terminal halves of granulin are fundamentally different from each other given that the N-terminal region has been shown to fold into a well-defined structure in isolation. This distinction might reflect the mirror-image symmetry in the scaffold as the highly-conserved cysteine framework is reversed in the C-terminal half of the molecule compared to the N-terminal, i.e., C-C-CC-CC (N-terminal) and CC-CC-C-C (C-terminal).

4.2 Introduction

Granulins are a highly conserved group of growth factors that are involved in several biological processes such as cell growth [1, 2], angiogenesis [3], cell cycle progression [4, 5], wound repair [5], and one example has recently been shown to display anti-microbial activity [6]. They have been found in a wide range of organisms including humans [7], plants [8], and parasites [9], and it has been suggested they have an ancient evolutionary origin [10]. Granulins are characterised by a unique 12-cysteine framework (Figure 4-1A), rich in a double cysteine motif [11, 12] and in some cases display four β-hairpins linked by intramolecular disulfide bonds (Figure 4-1B) [13, 14]. Despite the conserved cysteine framework, significant sequential, structural, and bioactivity variations are present in the granulin family [10, 15]. Granulins are released from progranulin, a cysteine-rich glycoprotein [16], by proteolysis as small granulin polypeptides [10], and each granulin domain has an approximate molecular weight of 6 kDa [17].

There is evidence of glycosylation in human progranulin, including within granulin domains [18], but there is limited information of post-translational modifications across the granulin family. Human progranulin comprises a signal peptide and tandem repeats of seven-and-a-half granulin motifs arranged in the order P-G-F-B-A-C-D-E, in which A-G are full repeats and P is paragranulin [1, 19, 20], an N-terminal half granulin motif [10, 15]. Paragranulins are present in many species with a multiunit progranulin. Figure 4-1C shows the structure of a paragranulin from the model organism zebrafish [21].
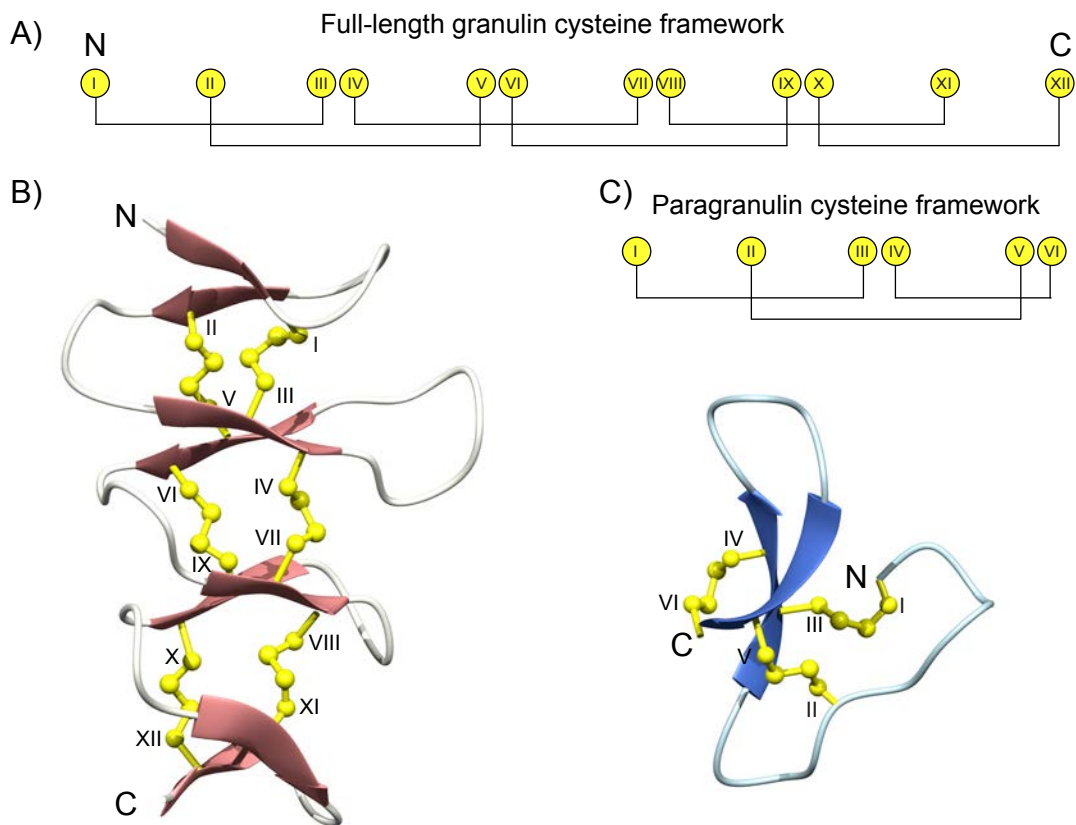
Figure 4-1 Cysteine framework and three-dimensional structures of granulin peptides. A) Full-length granulin cysteine framework. Cysteine residues are numbered using Roman numerals (I-XII) and the disulfide bonds shown as lines. B) The solution structure of zebrafish granulin AaE (PDB ID: 6CKU) [22, 23]. C) Paragranulin cysteine framework and the solution structure of zebrafish paragranulin (PDB ID: 7JIY) from zebrafish progranulin1 (Q8QGN9) [21, 23]. In panels B and C, disulfide bonds are shown in yellow (ball and stick format).

The structural disorder present in granulin peptides varies significantly [10, 15]. Carp granulin-1, extracted from the spleen and head kidney of *Cyprinus carpio*, [13] and recombinant zebrafish granulin AaE have well-defined structures over the majority of the molecules [17, 22, 23]. Human granulins A, C, and F [24], and zebrafish granulin AbB [23] have disordered C-terminal regions, but well-defined N-terminal regions based on NMR studies. By contrast, granulins B, D and E do not have well-folded structures at all [24]. However, the majority of these studies were done on recombinant versions of the peptides, and it has not been confirmed that the recombinant versions have the native fold.

Analysis of truncated granulins, with either engineered or naturally occurring sequences, has provided insight into the structure/function relationships of granulins. We have previously shown that a truncated granulin peptide from the N-terminal of the human liver fluke, *Opisthorchis viverrini*, can fold independently into a "mini-granulin" structure and has potent wound healing properties *in vivo* [25, 26]. This peptide incorporates a non-native disulfide bond, between Cys IV-Cys VI, with regard to the full-length granulin structure. We have also recently discovered that the mini-granulin fold occurs in a truncated form of a zebrafish granulin, and a naturally occurring zebrafish paragranulin (half-granulin motif) as shown in Figure 4-1C [21]. These peptides contain the first six cysteine residues of the granulin motif. Interestingly, the *in vitro* oxidative folding of zebrafish paragranulin, ZF-para$_{3s}$, produces a higher yield of this isomer compared to the truncated zebrafish peptide suggesting an evolutionary impact on the folding efficiency [21]. Even though the well-defined structure observed for parasite and zebrafish granulins includes three disulfide bonds, the N-terminal region of carp granulin-1 can fold independently, incorporating only two disulfide bonds [14, 27]. While unusual, this fold is not unique, and is present in a truncated plant cysteine protease (oryzain ß) with a granulin/epithelin like motif [8]. Additionally, this mini-granulin fold has recently been identified in conotoxins ϕ-MiXXVIIA [28] and H-Vc7.2 [29], and also in U33-theraphotoxin-Cg1c, a disulfide-rich peptide found in the venom of a Chinese theraphosid spider and belonging to the spider toxin superfamily SF22 [30].

In contrast to the N-terminal region, there is limited information on the folding and structure/function relationships related to the C-terminal region of granulins. To address this issue, we have analysed the structures of truncated form the C-terminal of zebrafish AaE. We show that peptide corresponding to the C-terminal region of zebrafish AaE cannot fold independently into well-defined structures, and the two-disulfide bond containing peptide has no inhibitory effect on cell growth.

4.3 Experimental section

4.3.1 Peptide synthesis and purification

C-terminal truncated zebrafish peptide was synthesised on a 0.1 mmole scale using Fmoc solid-phase synthesis chemistry. The peptide was assembled on 2-chlorotrityl chloride resin (Auspep, Australia), activated, and cleaved using the techniques previously reported [25, 26]. The crude peptide was precipitated and washed twice with ice-cold diethyl ether, dissolved in 50% acetonitrile:50% H_2O:0.1% TFA (v/v), and then freeze-dried at least twice to remove TFA.

The purification of the crude peptide was performed by reversed-phase high performance liquid chromatography (RP-HPLC) on a preparative C_{18} column (Phenomenex Jupiter 250 mm × 21.2 mm, 10 μm, 300 Å) using gradients of solvent B (90% acetonitrile/10% H_2O/0.045% TFA (v/v/v) and solvent A (H_2O/0.05% TFA (v/v)). The molecular mass of each collected peptide was determined using a 5800 MALDI TOF/TOF mass spectrometer (SCIEX, Framingham, MA) using α-cyano-4-hydroxycinnamic acid (CHCA; Sigma-Aldrich, Australia) matrix.

The truncated peptide was oxidised in 0.1 M ammonium bicarbonate buffer (pH 8.2), 5 mM reduced glutathione at a concentration of 0.2 mg/mL at room temperature for 48 hours. The pH was checked after dissolution to confirm the solution was still at pH 8. After the completion of the oxidation, the solution was acidified with TFA and loaded on a C_{18} preparative HPLC column (Phenomenex Jupiter 250 mm × 21.2 mm, 10 μm, 300 Å) with a flow rate of 6 mL/min. Peptide was eluted using gradients of solvent as above. Fractions were collected and the peptide masses were analysed using a SCIEX 5800 MALDI TOF/TOF spectrometer (SCIEX, Framingham, MA). Analytical RP-HPLC was carried out using a Phenomenex Jupiter 4 μm Proteo column (150 × 2.00 mm, 90 Å) and a gradient of 0-60% solvent B (Solvent A: 99.95% H_2O:0.05% TFA; Solvent B: 90% acetonitrile:10% H_2O:0.045% TFA) over 60 minutes with a flow rate of 0.4 mL/min.

4.3.2 NMR spectroscopy and structure determination

All NMR spectra were recorded on a 600 MHz AVANCE III NMR spectrometer (Bruker, Karlsruhe, Germany). The unlabeled, freeze-dried peptide (0.2 mM) was dissolved in 90% H_2O/10% D_2O (v/v). The pH of the NMR sample solution

was ~4.5. Two-dimensional NMR spectra for sequence-specific assignments were acquired. Two-dimensional TOCSY and NOESY spectra with mixing periods of 80 ms and 200 ms were recorded respectively. All spectra were processed using Bruker TopSpin (Bruker, Billerica, MA, USA) and assigned using CCPNMR Analysis [31, 32]. The αH secondary shift was determined by subtracting random coil ^{1}H NMR chemical shifts from the experimental αH chemical shifts [33]. The three-dimensional structure of ZF-C$_{20\_2s}$ was calculated using the CYANA program, based on automated assignment of the NOEs [34]. Dihedral-angle restraints predicted by TALOS-n were used in the structure calculations [35]. Structures were visualized using MOLMOL [36].

4.3.3   Human skin normal fibroblast cells

The human skin normal fibroblast cell line, 1BR.3.GN, was obtained from the European Collection of Authenticated Cell Cultures (ECACC). 1BR.3.GN cells were cultured and maintained in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) (Life Technologies, Australia) containing 1 × antibiotic/antimycotic including penicillin, streptomycin, and amphotericin B, and 1 × GlutaMAX, supplemented with 10% fetal bovine serum (FBS) (Gibco, Scotland) at 37 °C and 5% CO_2. Phosphate Buffered Saline (PBS; sodium chloride, potassium chloride, disodium phosphate, monopotassium phosphate: ThermoFisher (#14190250)) was used for negative vehicle control. Cell proliferation assays were performed with DMEM/F12 low nutrient media supplemented with 0.5% FBS.

4.3.4   Proliferation xCELLigence assay

Cells were seeded at 5000 cells per well in 150 μL of complete media in E-plates (ACEA Biosciences) and grown overnight while monitored with an xCELLigence SP system (ACEA Biosciences, San Diego, CA, USA), which monitors cellular events in real time by measuring electrical impedance across gold microelectrodes integrated into the base of tissue culture plates. Cells were washed three times with low nutrient media prior to addition of 150 μL of low nutrient media and incubated for a minimum of 6 h before further treatment. The peptide was quantified based on absorbance measurements at 214 nm and a predicted extinction coefficient based on the amino acid composition [37, 38]. The pH was checked after

dissolution to confirm the solution was still at pH 7. Treatments were prepared at 8.5 × concentration and added to each well (20 μL) for a final total volume of 170 μL. The xCELLigence system recorded cell indices at intervals of 1 h for 5−6 days following treatment. Readings for the cell index were normalized before treatment, and cell proliferation ratios characterised the relative numbers of cells compared to control cells. Cell proliferation rates were compared between treatment and control wells, and a two-way ANOVA test was used for multiple comparisons with Holm-Sidak's correction, using GraphPad Prism 9.0.

4.4 Results

4.4.1 Design and synthesis of C-terminal peptide

A peptide corresponding to the C-terminal region of zebrafish granulin AaE (residues 35 to 54) was synthesised using Fmoc solid-phase peptide synthesis and 2-chlorotrityl chloride resin. The peptide is referred to as ZF-C$_{20\_2s}$ (the "2s" refer to the presence of two disulfide bonds). In ZF-C$_{20\_2s}$, Cys IX was mutated to an alanine residue, as shown in Figure 4-2. The synthesised peptide was purified using RP-HPLC, and the mass analysed using MALDI-TOF mass spectrometry. The peptide was oxidised in a single-step reaction without selective protection of the cysteine residues. The folding yield for ZF-C$_{20\_2s}$ was 30%.

The analytical RP-HPLC trace of the oxidation reaction contains three peaks at different retention times, corresponding to three isomers (I$_1$, I$_2$ and I$_3$) with the same masses of 2009.7340 Da (Figure 4-3A). The major peak (I$_2$) in the RP-HPLC trace, produced under the *in vitro* oxidative folding conditions and purified to greater than 92% purity, is shown in Figure 4-3B. The SCIEX TOF/TOF™ 5800 MALDI mass spectrum of oxidised and reduced ZF-C$_{20\_2s}$ using CHCA matrix are shown in Figure 4-4. There was no evidence of dimer formation between 4000 and 6000 for oxidised peptide. The overlay of the oxidation reaction with the reduced peptide is shown in appendix 7.
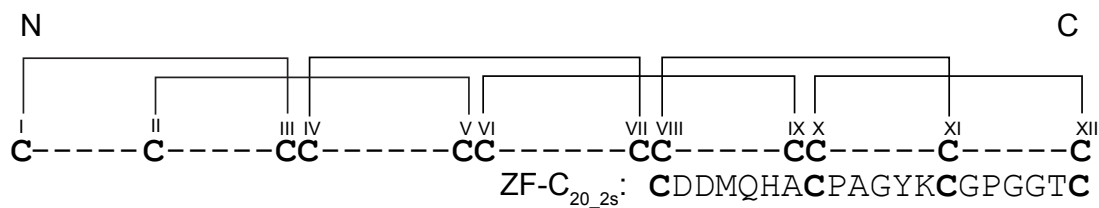
Figure 4-2 Schematic representation of the cysteine framework present in zebrafish AaE and the sequences of the truncated analogue. The cysteine residues are numbered using Roman numerals (I-XII) and connected based on the three-dimensional structure of the full-length peptide. The sequence of ZF-C$_{20\_2s}$ is shown below the full-length framework.
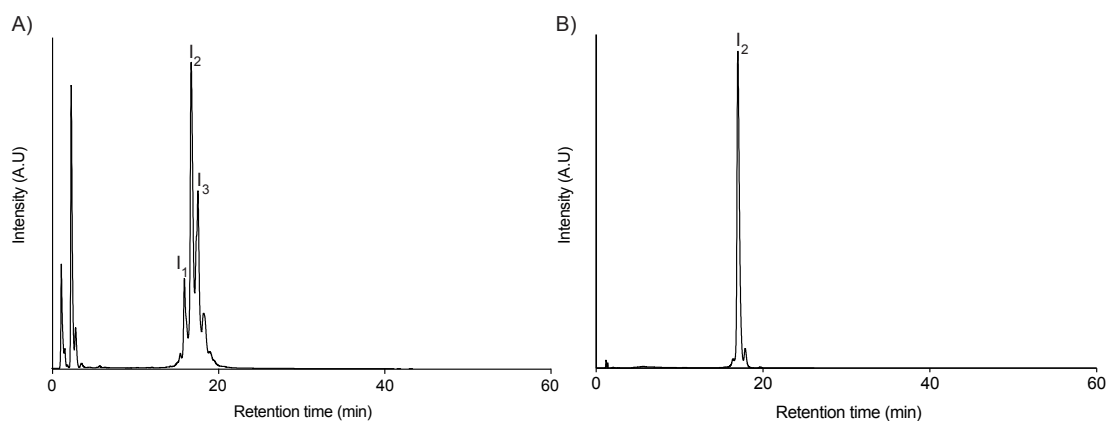
Figure 4-3 HPLC analysis of oxidised ZF-C$_{20\_2s}$. A) Three peaks at different retention times corresponding to three isomers. B) Purified peak corresponding to isomer 2.
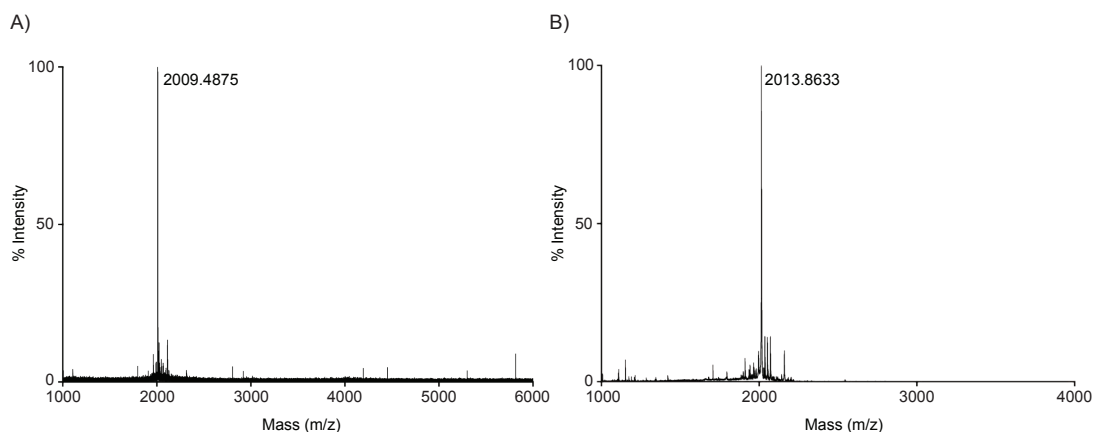
Figure 4-4 MALDI-TOF mass spectrometry spectrum of ZF-C$_{20\_2s}$. The spectra in panels A, B correspond to the $[M+H]^+$ of the oxidised and reduced peptides, respectively.

4.4.2   Structural Analysis and NMR Spectroscopy

The one-dimensional proton NMR spectra of the ZF-C$_{20\_2s}$ isomers had limited dispersion in the amide region (Figure 4-5), and the peaks were much sharper in the I$_2$ spectrum compared to I$_1$ and I$_3$. However, most of the backbone and side-chain protons were assigned for all three isomers. Analysis of the secondary shifts indicates that the ZF-C$_{20\_2s}$ isomers do not have regular secondary structure (Figure 4-6), based on the majority of the shifts being close to random coil (within $\pm$ 0.1 ppm). As a comparison, the secondary shifts of residues 35 to 54 in the full-length zebrafish granulin AaE [21], is shown. The large values of secondary shifts of zebrafish granulin AaE (blue) are consistent with the ß strands present in these regions. By contrast, ß-strands were not present in the ZF-C$_{20\_2s}$ isomers and consequently much lower secondary shifts were observed.
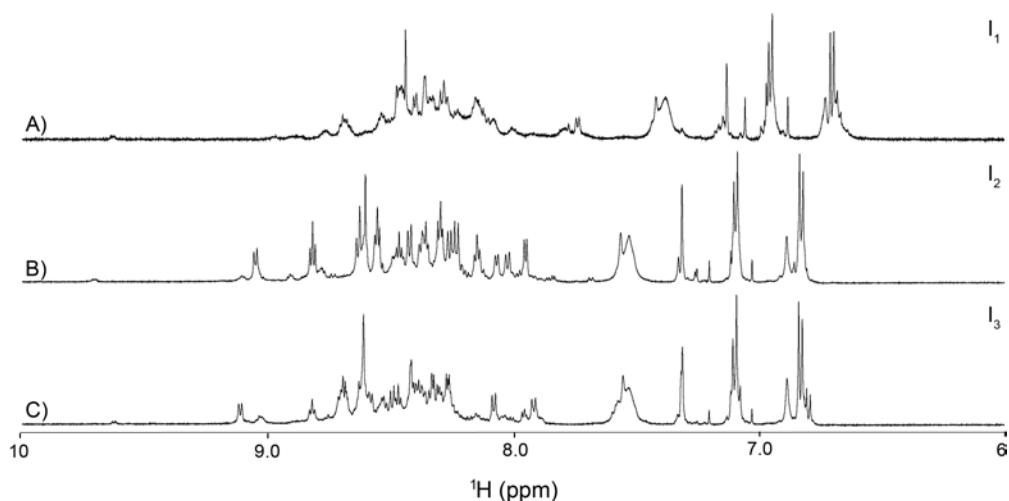
Figure 4-5 One-dimensional proton NMR spectra for ZF-C$_{20\_2s}$. A) I$_1$, B) I$_2$, and C) I$_3$. Spectra are shown from 6.0 to 10 ppm; this region contains primarily the amide and the aromatic ring protons.
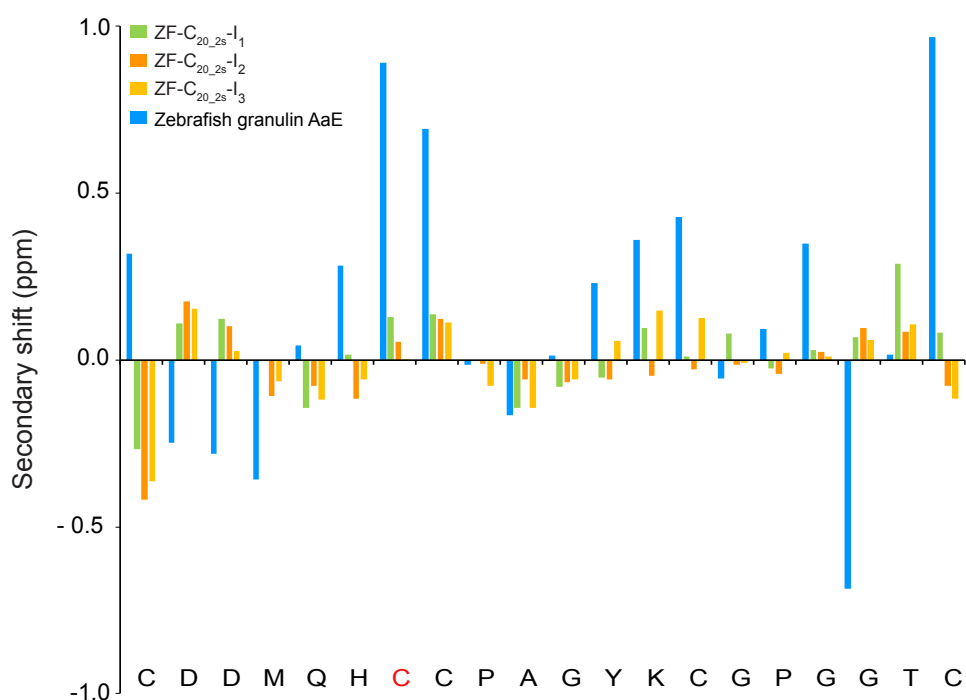
Figure 4-6 αH Secondary-shifts for ZF-C$_{20\_2s}$ isomers. The αH secondary shift was calculated by subtracting the random coil ^{1}H NMR chemical shifts previously reported by Wishart et al. [33] from the experimental αH chemical shifts. The sequence of the C-terminal of AaE homologue to ZF-C$_{20\_2s}$ is given at the bottom of the diagram (BMRB: 30422). In ZF-C$_{20\_2s}$, the Cys in red was replaced with alanine.

Three-dimensional structures of the major isomer, ZF-C$_{20\_2s}$-I$_2$, were calculated with the three possible disulfide bond connectivities included in the calculations; the cyana [34] target functions for these three connectivities are shown in Table 4-1. Set 3 including the connectivity of Cys VIII-Cys X, Cys XI-Cys XII had the lowest target function which does not correspond to the disulfide bonds present in the full-length structures of well characterised granulins [23]. It appears likely that the isomer I$_1$ of ZF-C$_{20\_2s}$ is corresponding to the disulfide bonds present in the full-length granulins as it is quite common for the earliest eluting peak on RP-HPLC to correspond to the natively folded peptide in small disulfide-rich peptides; however, further study would be required to definitively determine the connectivity such as selective reduction and alkylation of the peptide [39, 40] or regioselective formation of the predicted disulfide bonds using a combination of two or three types of thiol protecting groups, such as acetamidomethyl (Acm) [41, 42]. Overall, the structures are poorly defined regardless of the disulfide connectivity used, likely reflecting structural flexibility, but a short stretch (residues 9-12) corresponding to a β-turn is relatively well-defined. Superposition of the backbone atoms of all three sets and a ribbon representation of the three-dimensional structure of Set 3 are shown in Figures 4-7A and -7B respectively. The refinement statistics for the structural ensembles are given in Table 4-2.
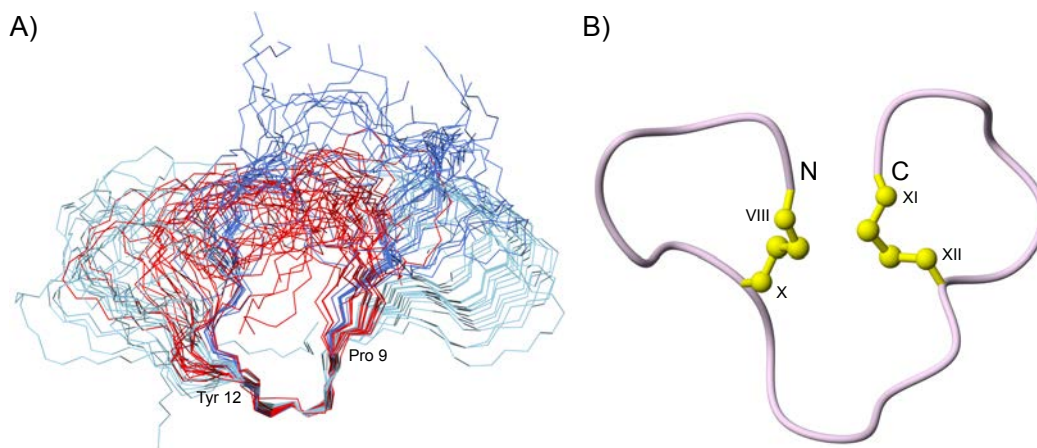
Figure 4-7 Backbone overlay and three-dimensional structure of ZF-C$_{20\_2s}$-I$_2$ determined based on NMR data in aqueous solution. A) Backbone overlay of set 1 (red), set 2 (dark blue), and set 3 (light blue). Structures were superimposed over residues 9-12 with correspond to a ß-turn. B) The backbone ribbon representation of the structure of ZF-C$_{20\_2s}$-I$_2$, set 3 (Cys VIII-Cys X, Cys XI-Cys XII). The disulfide bonds are shown in yellow. The cysteine residues are numbered using Roman numerals corresponding to the numbers in full-length protein as shown in Figure 4-1.

Table 4- 1 Cyana target functions for ZF-C$_{20\_2s}$-I$_2$.

Connectivity	Target functions
Set 1: Cys VIII-Cys XI, Cys X-Cys XII	0.11 ± 0.0212
Set 2: Cys VIII-Cys XII, Cys X-Cys XI	0.0883 ± 0.0197
Set 3: Cys VIII-Cys X, Cys XI-Cys XII	0.0496 ± 0.0159

Table 4- 2 Structural statistics for ZF-C$_{20\_2s}$-I$_2$.

Experimental restraints	
Interproton distance restraints	
All	101
Intraresidue, |i-j|=0	23
Sequential, |i-j|=1	62
Medium range, 1 <|i-j| < 5	16
Long range, |i-j| >=5	0
Disulfide-bond restraints (3 restraints per bond)	6
Dihedral-angle restraints	23

R.m.s. deviations from mean coordinate structure (Å)	
Backbone atoms	2.70 ± 0.90
All heavy atoms	3.44 ± 1.02
Backbone atoms (9-12)	0.39 ± 0.15
All heavy atoms (9-12)	1.30 ± 0.61

Ramachandran Statistics	
% In most favored region	62.5
% Residues in additionally allowed regions	37.5

4.4.3 Cell Proliferation Monitoring in Real Time Using xCELLigence

The effect of the three isomers of ZF-C$_{20\_2s}$ on the growth of 1BR.3.GN fibroblasts was assessed using an xCELLigence system. 1BR.3.GN is a human fibroblast skin cell line derived from transformed normal fibroblasts. Cells were cultured with two concentrations of the isomers, 200 nM and 1 μM (Figure 4-8). The 200 nM datasets are shown in appendix 8. Interestingly, despite being only partially folded, ZF-C$_{20\_2s}$-I$_2$ and -I$_3$ appear to have a slight inhibitory effect on fibroblast cell growth at 200 nM and 1 μM on day 3 after treatment, compared to the negative control peptide (a 21-residue tropomyosin derived peptide with the sequence of RAETGESKIVELEEELRVVGN), respectively. A tropomyosin fragment, previously shown to have no activity in a cell proliferation assay was used as a negative control. Importantly, this peptide was produced in the laboratory using the same conditions to the ZF-C$_{20\_2s}$. However, the variability on the data prevents this from being statistically significant.
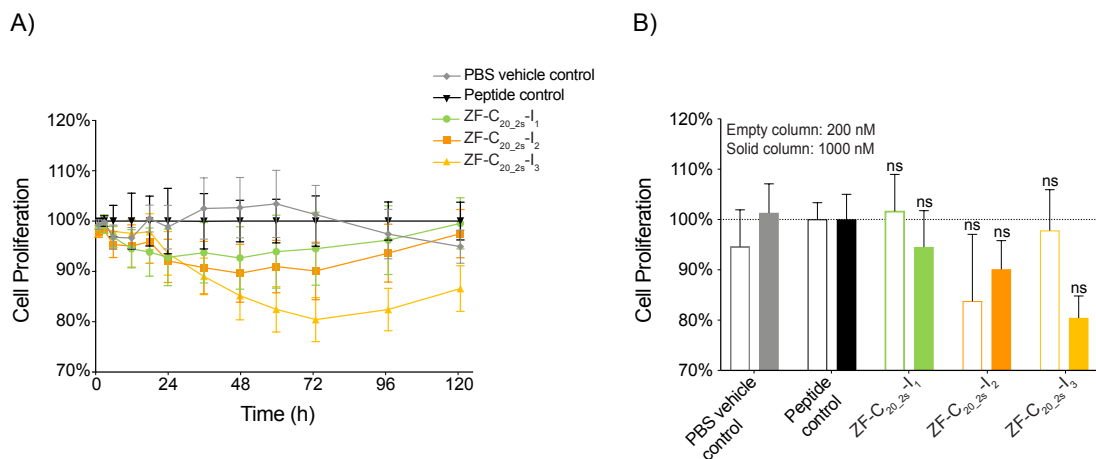
Figure 4-8 Real-time cell proliferation analysis for three isomers of ZF-C$_{20\_2s}$ using xCELLigence. A) Fibroblast cells were treated with three isomers of ZF-C$_{20\_2s}$ at concentrations of 200 nM and 1 μM. For clarity only 1 μM is shown. Cell index was measured 4 days after treatment. B) Day 3 relative proliferation. Empty columns and solid columns represent concentrations of 200 nM and 1 μM, respectively. Both panels: The proliferation rates relative to peptide control are plotted as mean ± SEM bars. Data were analysed by two-way ANOVA against peptide control (Not significant = ns).

4.5 Discussion

Full-length granulin modules are present in genes from premetazoan organisms to mammals [10]. Despite this conservation of the full-length module, several studies have shown that the N-terminal region of granulins (equivalent to paragranulin) can fold independently, and in some cases display bioactivity, in the absence of the C-terminal region. Here, we show that the C-terminal region of zebrafish AaE does not fold independently. ZF-C$_{20\_2s}$-I$_2$ and -I$_3$ displayed slight inhibitory effects on cell growth, but these effects were not statistically significant, as shown in Figure 4-7. Our study intended to analyze the properties of the C-terminal region of granulins, but the lack of folding of the C-terminal region of zebrafish granulin AaE poses interesting questions regarding the role of this conserved granulin region.

Neither of the ZF-C$_{20\_2s}$ isomers displayed significant dispersion in the amide region of the one-dimensional NMR spectra, but the spectra of three isomers were of sufficient quality to allow assignment of the majority of the resonances. The overall structure of ZF-C$_{20\_2s}$-I$_2$ was not well defined, with the exception of a well-defined β-turn involving residues Pro-Ala-Gly-Tyr (PAGY). A tetrapeptide comprising this PAGY sequence has previously been shown to display antioxidant properties [43, 44], and a recent study has suggested that granulins can mitigate oxidative stress by reducing Cu-induced reactive oxygen species cytotoxicity in cells [45]. Whether ZF-C$_{20\_2s}$, or indeed granulin AaE, displays antioxidant activity, and if the PAGY motif is involved in such activity, requires further study.

In contrast to the limited structure observed for ZF-C$_{20\_2s}$-I$_2$, the N-terminal truncated peptides from zebrafish granulin AaE and *Ov*-GRN-1 [21, 25, 26] have previously been shown to contain well-defined β-hairpin structures [21]. These N-terminally truncated forms contain the first six cysteine residues of the granulin framework and incorporate a non-native disulfide bond. They are analogous to paragranulins found in several species. In the current study ZF-C$_{20\_2s}$ were designed to contain, or at least have the possibility of forming, native disulfide bonds. The ideal comparison would be a C-terminal paragranulin, but there is no evidence of C-terminal paragranulins comprising the six C-terminal cysteine residues in any species.

The current study also analysed the biological function of the truncated C-terminal peptides. Unexpectedly, the poorly folded peptide ZF-C$_{20\_2s}$-I$_2$, and -I$_3$ displayed a slight, but not significant inhibitory effect in the cell proliferation assay. Given the sequence similarity between ZF-C$_{20\_2s}$ isomers, it appears likely that the disulfide bonds might play a role in this difference in bioactivity, and it would be of interest for future studies to analyze the bioactivity of reduced forms of these peptides to determine the influence of the disulfide bonds. There is precedent for reduced forms of granulin peptides to have bioactivity, with reduced granulin-B shown to activate NF-κB in human neuroblastoma cells [46].

Full-length granulin AaE promotes the survival of neuronal cells with no indication of an inhibitory effect [23]. However, there are other granulin peptides that have been shown to have inhibitory effects on cancer cell lines. For example, the interaction of human granulin A and enolase 1 inhibits the growth of hepatocellular carcinoma [47]. A similar inhibitory effect on the growth of the same cancer was demonstrated by the synergistic effect of human granulin A and cisplatin [48]. In addition, epithelin 1 and 2, purified from rat kidneys, inhibit the growth of A431 cells, which are derived from a human epidermal carcinoma of the vulva [49]. The bioactivity observed in the current study for two isomers of ZF-C$_{20\_2s}$ might be related to the poorly defined structure facilitating binding to different targets, which are not relevant for the full-length protein [50-52].

4.6 Conclusions

In summary, the C-terminal region of granulin AaE does not appear to fold independently into a well-defined structure, and isomers of ZF-C$_{20\_2s}$ did not present any cell proliferation. Our findings appear consistent with the presence of structural disorder in the C-terminal region of several full-length granulin modules, and the lack of individual C-terminal regions present in genomes that are analogous to paragranulin. It is tempting to speculate that the C-terminal region is consequently less important but given the complexity of the structure/function relationships of this family and the conservation of the full-length module, it appears likely that the C-terminal region, as part of the full-length module, has been conserved for a reason; the elucidation of which will await further study.

4.7 Funding

4.8 Acknowledgements

4.9 References

1. Bateman, A. and Bennett, H.P.J., *The granulin gene family: From cancer to dementia.* BioEssays, 2009. 31(11): p. 1245-1254.

2. Wang, C., et al., *A novel granulin homologue isolated from the jellyfish Cyanea capillata promotes proliferation and migration of human umbilical vein endothelial cells through the ERK1/2-signaling pathway.* Int. J. Biol. Macromol., 2019. 135: p. 212-225.

3. Haugen, B., et al., *Granulin secreted by the food-forne liver fluke opisthorchis viverrini promotes angiogenesis in human endothelial cells.* Front. Med., 2018. 5: p. 30.

4. Hanington, P.C., et al., *Molecular and functional characterization of granulin-like molecules of insects.* Insect Biochem. Mol. Biol., 2008. 38(5): p. 596-603.

5. He, Z. and Bateman, A., *Progranulin (granulin-epithelin precursor, PC-cell-derived growth factor, acrogranin) mediates tissue repair and tumorigenesis.* J. Mol. Med., 2003. 81(10): p. 600-612.

6. Wu, S.-H., et al., *Granulin peptide GRN-41 of Mozambique tilapia is a novel antimicrobial peptide against Vibrio species.* Biochem. Biophys. Res. Commun., 2019. 515(4): p. 706-711.

7. Bhandari, V. and Bateman, A., *Structure and chromosomal location of the human granulin gene.* Biochem. Biophys. Res. Commun., 1992. 188(1): p. 57-63.

8. Tolkatchev, D., et al., *A peptide derived from the c-terminal part of a plant cysteine protease folds into a stack of two β-hairpins, a scaffold present in the emerging family of granulin-like growth factors.* J. Pept. Res., 2001. 57(3): p. 227-233.

9. Smout, M.J., et al., *A granulin-like growth factor secreted by the carcinogenic liver fluke, Opisthorchis viverrini, promotes proliferation of host cells.* PLoS Pathog., 2009. 5(10): p. e1000611.

10. Palfree, R.G., et al., *The evolution of the secreted regulatory protein progranulin.* PLoS One, 2015. 10(8): p. e0133749.

11. Bateman, A. and Bennett, H.P., *Granulins: the structure and function of an emerging family of growth factors.* J. Endocrinol., 1998. 158(2): p. 145-151.

12. Bhandari, V., et al., *Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains.* Proc. Natl. Acad. Sci., 1992. 89(5): p. 1715-1719.

13. Hrabal, R., et al., *The hairpin stack fold, a novel protein architecture for a new family of protein growth factors.* Nat. Struct. Biol., 1996. 3(9): p. 747-751.

14. Tolkatchev, D., et al., *Design and solution structure of a well-folded stack of two β-hairpins based on the amino-terminal fragment of human granulin A.* Biochemistry, 2000. 39(11): p. 2878-2886.

15. Dastpeyman, M., et al., *Folding of granulin domains.* Pept. Sci., 2018. 110(3).

16. Tian, Q., et al., *Three TNFR-binding domains of PGRN act independently in inhibition of TNF-alpha binding and activity.* Front. Biosci., 2014. 19(7): p. 1176-1185.

17. Cadieux, B., et al., *The zebrafish progranulin gene family and antisense transcripts.* BMC Genom., 2005. 6(1): p. 156-156.

18. Songsrirote, K., et al., *Development and application of mass spectrometric methods for the analysis of progranulin N-glycosylation.* J. Proteom., 2010. 73(8): p. 1479-90.

19. Jian, J., et al., *Insights into the role of progranulin in immunity, infection, and inflammation.* J. Leukoc. Biol., 2013. 93(2): p. 199-208.

20. Wang, B.C., et al., *New discovery rarely runs smooth: An update on progranulin/TNFR interactions.* Protein Cell, 2015. 6(11): p. 792-803.

21. Takjoo, R., et al., *Folding of truncated granulin peptides.* Biomolecules, 2020. 10(8): p. 1152.

22. Wang, P.F., *Structure genomics of zebrafish granulins.* 2004, McGill University, Montreal.

23. Wang, P., et al., *Structure dissection of zebrafish progranulins identifies a well-folded granulin/epithelin module protein with pro-cell survival activities.* Protein Sci., 2018. 27(8): p. 1476-1490.

24. Tolkatchev, D., et al., *Structure dissection of human progranulin identifies well-folded granulin/epithelin modules with unique functional activities.* Protein Sci., 2008. 17(4): p. 711-724.

25. Bansal, P.S., et al., *Development of a potent wound healing agent based on the liver fluke granulin structural fold.* J. Med. Chem., 2017. 60(10): p. 4258-4266.

26. Dastpeyman, M., et al., *Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing.* J. Med. Chem., 2018. 61(19): p. 8746-8753.

27. Vranken, W.F., et al., *A 30-residue fragment of the carp granulin-1 protein folds into a stack of two β-hairpins similar to that found in the native protein.* J. Pept. Res., 1999. 53(5): p. 590-597.

28. Jin, A.H., et al., *Conotoxin Φ-MiXXVIIA from the superfamily G2 employs a novel cysteine framework that mimics granulin and displays anti-apoptotic activity.* Angew. Chem., 2017. 129(47): p. 15169-15172.

29. Nielsen, L.D., et al., *The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework.* J. Biol. Chem., 2019. 294(22): p. 8745-8759.

30. Pineda, S.S., et al., *Structural venomics reveals evolution of a complex venom by duplication and diversification of an ancient peptide-encoding gene.* Proc. Natl. Acad. Sci., 2020. 117(21): p. 11399-11408.

31. Wüthrich, K., *NMR studies of structure and function of biological macromolecules (Nobel Lecture).* J. Biomol. NMR, 2003. 27(1): p. 13-39.

32. Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: Development of a software pipeline.* Proteins, 2005. 59(4): p. 687-696.

33. Wishart, D.S., et al., *^1H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects.* J. Biomol. NMR, 1995. 5(1): p. 67-81.

34. Güntert, P., *Automated NMR structure calculation with CYANA.* Methods Mol. Biol., 2004. 278: p. 353-378.

35. Shen, Y. and Bax, A., *Protein structural information derived from NMR chemical shift with the neural network program TALOS-N.* Methods Mol. Biol., 2015. 1260: p. 17-32.

36. Koradi, R., et al., *MOLMOL: A program for display and analysis of macromolecular structures.* J. Mol. Graph., 1996. 14(1): p. 51-55.

37. Gill, S.C. and von Hippel, P.H., *Calculation of protein extinction coefficients from amino acid sequence data.* Anal. Biochem., 1989. 182(2): p. 319-26.

38. Hilario, E.C., et al., *An improved method of predicting extinction coefficients for the determination of protein concentration.* PDA J. Pharm. Sci. Technol., 2017. 71(2): p. 127-135.

39. Daly, N.L., et al., *Disulfide folding pathways of cystine knot proteins. Tying the knot within the circular backbone of the cyclotides.* J. Biol. Chem., 2003. 278(8): p. 6314-22.

40.   Hale, J.E., et al., *A simplified procedure for the reduction and alkylation of cysteine residues in proteins prior to proteolytic digestion and mass spectral analysis.* Anal. Biochem., 2004. 333(1): p. 174-81.

41.   Okumura, M., et al., *A chemical method for investigating disulfide-coupled peptide and protein folding.* FEBS J., 2012. 279(13): p. 2283-2295.

42.   Hidaka, Y. and Shimamoto, S., *Folding of peptides and proteins: Role of disulfide bonds, recent developments.* Biomol. Concepts, 2013. 4(6): p. 597-604.

43.   Maharani R., et al., *Synthesis of tetrapeptides and screening of their antioxidant properties.* Curr. Bioact. Compd., 2019. 15: p. 680-685.

44.   Nikoo, M., et al., *Antioxidant and cryoprotective effects of Amur sturgeon skin gelatin hydrolysate in unwashed fish mince.* Food Chem., 2015. 181: p. 295-303.

45.   Bhopatkar, A.A. and Rangachari, V., *Are granulins copper sequestering proteins?* Proteins, 2021. 89(4): p. 450-461.

46.   Ghag, G., et al., *Fully reduced granulin-B is intrinsically disordered and displays concentration-dependent dynamics.* Protein Eng. Des. Sel., 2016. 29(5): p. 177-186.

47.   Chen, X.L., et al., *Interaction between granulin A and enolase 1 attenuates the migration and invasion of human hepatoma cells.* Oncotarget, 2017. 8(18): p. 30305-30316.

48.   Qiao, G., et al., *Granulin A synergizes with cisplatin to inhibit the growth of human hepatocellular carcinoma.* Int. J. Mol. Sci., 2018. 19(10): p. 3060.

49.   Shoyab, M., et al., *Epithelins 1 and 2: Isolation and characterization of two cysteine-rich growth-modulating proteins.* Proc. Natl. Acad. Sci., 1990. 87(20): p. 7912-7916.

50.   Kriwacki, R.W., et al., *Structural studies of p21Waf1/Cip1/Sdi1 in the free and Cdk2-bound state: conformational disorder mediates binding diversity.* Proc. Natl. Acad. Sci., 1996. 93(21): p. 11504-11509.

51.   Uversky, V.N., *Natively unfolded proteins: A point where biology waits for physics.* Protein Sci., 2002. 11(4): p. 739-756.

52.   Uversky, V.N., *A decade and a half of protein intrinsic disorder: Biology still waits for physics.* Protein Sci., 2013. 22(6): p. 693-724.

Chapter 5. Structural Analysis of an *Asterias rubens* Peptide Indicates the Presence of a Disulfide Directed ß-hairpin

5.1 Abstract

Sea stars are an abundant group of marine invertebrates that display remarkably robust regenerative capabilities throughout all life stages. Numerous proteins and peptides have recently been identified in a proteome study on the coelomic fluid (biofluid) of the common sea star *Asterias rubens*, which appear to be involved with the wound-healing response. Several of these injury-responsive peptides have no known structures and functions, including the 28-residue peptide KASH2. Here we show this peptide does not enhance proliferation of human fibroblasts, a property commonly found in other wound-healing peptides. However, interestingly, this peptide adopts a disulfide-directed β-hairpin fold (DDH). The DDH motif appears to be evolutionarily related to the inhibitor cystine knot motif, which is one of the most widespread disulfide-rich peptide folds. The DDH motif was originally thought to be restricted to arachnids, but our study suggests it is more widespread, and, based on phylogenetic analyses it appears that the presence of this motif in sea stars is the result of convergent evolution. Further study is required to determine the role of KASH2 in the sea star injury response.

5.2 Introduction

Sea stars are invertebrates with exceptional powers of regeneration and the capability to regrow lost limbs [1]. Several molecules have been shown to be involved in this regeneration including the serine protease calpain [2], neurotransmitters (monoamines), neuropeptides (substance P, SALMFamides 1 and 2), and growth-factor-like molecules [3], but a recent study suggested there could be many more [4].

Proteomic analysis of cell-free coelomic fluid (biofluid) of the common sea star *Asterias rubens* highlighted several proteins that were upregulated in response to injury, and therefore are potentially involved in defence, cell migration, and wound healing [4]. The majority of the upregulated proteins were classified as pattern-recognition receptors or peptidase inhibitors. However, there were several uncharacterised proteins including the disulfide-rich proteins KARTESH (KASH1 and KASH2) and CHUPA (CHU1, CHU2, and CHU3) (Figure 5-1). These proteins are expressed containing a signal peptide, an acidic region, and the mature peptide regions containing four cysteine residues. Convertase cleavage sites between the acidic region and the cysteine-rich regions result in the predicted mature peptides for the KARTESH peptides containing 28-30 residues, and the CHUPA peptides containing 35-48 residues.
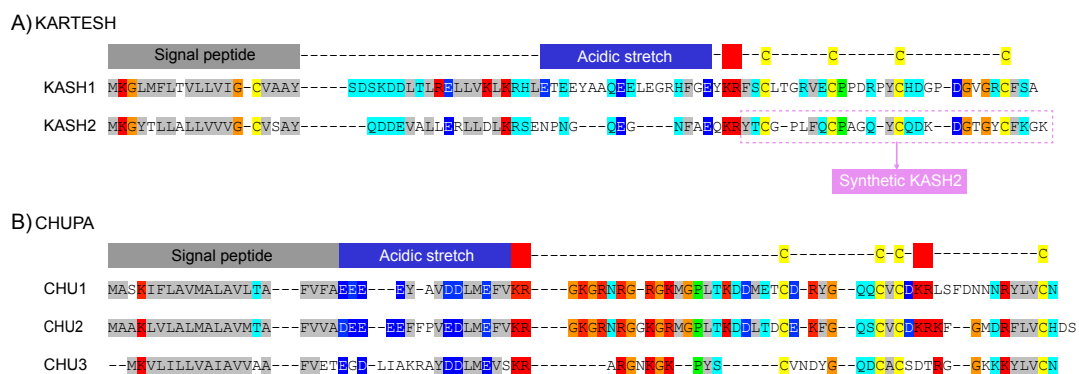
Figure 5-1 Multiple sequence alignments of two peptides in *Asterias rubens* transcriptome. A) KARTESH. B) CHUPA. Positions of predicted signal peptide, acidic stretch and dibasic convertase cleavage sites are shown in grey, blue, and red boxes, respectively. Cysteine residues and synthetic KASH2 sequence are shown in yellow and purple boxes, respectively.

Several homologues of the KARTESH and CHUPA proteins are present in other echinoderms, but no other phyla. The cysteine residues are conserved throughout the predicted mature peptides but there is variation in the residue type and the number of residues in the inter-cysteine loops. This phenomenon of families of peptides with inter-cysteine loop sequence variations combined with conservation of the cysteine residues is prevalent throughout nature including in the peptides in the venoms of cone snails and spiders, with the former containing numerous four-cysteine containing peptides [5-12].

The upregulation of the uncharacterised sea star proteins in response to injury warrants further investigation. In particular, determination of the three-dimensional structures might provide insight into their function and evolution. Here we have determined the three-dimensional structure of KASH2 using NMR spectroscopy. The rationale for the selection of KASH2 for this study was based on two factors; the cysteine-rich nature of the peptide which is likely to confer stability in combination with the high upregulation after injury. Intriguingly, the three-dimensional solution structure contains an ancestral protein fold termed the disulfide directed β-hairpin (DDH), previously thought to be restricted to arachnids. We also explored whether the potential wound-healing properties are associated with an influence on fibroblast proliferation, similar to what we have observed for granulin-derived peptides. We have previously shown that the human parasitic liver fluke *Opisthorchis viverrini*, secretes a granulin referred to as *Ov*-GRN-1 [13]. Recombinant *Ov*-GRN-1 and truncated peptides derived from the full-length protein accelerate wound repair in mice [14, 15]. For the truncated peptides we have shown that the *in vivo* wound healing correlates with an increase in cell proliferation of human normal skin fibroblast cells [15, 16]. However, KASH2 did not show any effect on cell proliferation, suggesting it might not be directly related to wound healing in the sea star, or that phyla specific processes are involved.

5.3 Experimental section

5.3.1 Peptide synthesis and purification

The peptide was synthesised by manual solid-phase peptide synthesis on a 0.1 mmole scale using Fmoc-SPPS chemistry and was assembled on 2-chlorotrityl chloride resin (Auspep). Amino acids were activated using O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate (HCTU; [Peptides International, KY, USA]) in peptide grade dimethylformamide (DMF; [Auspep]). Resin was cleaved using a mixture of 95% trifluoroacetic acid (TFA; Auspep, Australia)/2.5% triisopropyl silane (TIPS; [Aldrich, USA])/2.5% H_2O at room temperature for 2 h [15, 17], and the TFA was removed by evaporation. Cleaved peptide was precipitated with ice-cold diethyl ether. Diethyl ether was removed by filtration, and the precipitated peptide was dissolved in 50% acetonitrile/50% H_2O/0.1% TFA (v/v/v) and subsequently lyophilized at least twice to remove TFA. The resulting crude peptide was purified with RP-HPLC on a C_{18} preparative column (Phenomenex Jupiter 10 μm, C_{18}, 300 Å, 250 mm × 21.2 mm) using an appropriate gradient of solvents A and B. The purity of the peptides was assessed using analytical RP-HPLC on a C_{18} analytical column (Agilent Eclipse Plus C_{18}, 3.5 μm, 4.6 × 100 mm) with a gradient of 0-50% Solvent B in 50 min, 50-90% Solvent B in 5 min, 90-100% Solvent B in 5 min at 1 mL/min.

5.3.2 Disulfide formation

KASH2 was oxidised by dissolving lyophilized purified peptide in 100 mM ammonium bicarbonate (pH 8.2) for 24 h at room temperature. The pH was checked after dissolution to confirm the solution was still at pH 8. Oxidised peptide was acidified, filtered, and purified using RP-HPLC on a C_{18} preparative column (Phenomenex Jupiter 10 μm C_{18} 300 Å, 250 x 21.2 mm). The purity of the peptide was assessed using analytical RP-HPLC on the same C_{18} analytical column and gradient as above. The peptide mass was analysed using a 5800 MALDI TOF/TOF spectrometer (SCIEX, Framingham, MA).

Although the native connectivity of KASH2 is not known to determine the disulfide connectivity present in the major isomer in the one-step oxidative folding, selective protection of the cysteine residues was carried out. Side-chain protection with

acetamidomethyl (Acm) groups was used for Cys II and IV, and trityl (Trt) protecting groups were used for Cys I and Cys III in the synthesis. Following TFA cleavage and deprotection in the presence of scavenger (TIPS), the disulfide bond between Cys I and Cys III was formed by air oxidation in pH 8.2 buffer for 24 h, and the peptide was purified using the procedure described above. The second disulfide bond formed following removal of the Acm protecting group from Cys II and Cys IV. Cys-Acm peptide was dissolved in 50% aqueous acetic acid ($\sim$ 0.5 mg/mL). Then 1 M HCl (0.1 mL per mg of peptide) was added followed immediately by I_2 in 100% aqueous acetic acid until solution was yellow. The flask was flushed with nitrogen and sealed. The reaction was completed in 2 h. The oxidation reaction was quenched by adding 1 M ascorbic acid solution dropwise until solution was colourless. Finally, the quenched reaction was diluted with solvent A until the percentage of acetic acid was $\sim$10%. Oxidised peptide was purified using a C_{12} semi-preparative column (Phenomenex Jupiter Proteo C_{12}, 4 μm, 90 Å, 250 × 10.0 mm). The purity of the peptides was assessed using analytical RP-HPLC on a C_{18} analytical column (Agilent Eclipse Plus C_{18}, 3.5 μm, 4.6 × 100 mm) using a gradient of 0-50% Solvent B in 50 min, 50-90% Solvent B in 5 min, 90-0% Solvent B in 5 min at 1 mL/min.

5.3.3   NMR spectroscopy and structure determination

NMR spectra were recorded at 290 K on Bruker Avance III 600 MHz spectrometer (Bruker, Karlsruhe, Germany) equipped with a TCI cryoprobe. The sample was prepared from lyophilized peptide at concentration of approximately 0.2 mM in 90% H_2O/10% D_2O (D_2O; Cambridge Isotope Laboratories, Woburn, MA, USA). The pH of the NMR sample solution was $\sim$4.5. Two-dimensional spectra including ^{1}H−^{1}H TOCSY, ^{1}H−^{1}H NOESY, ^{1}H−1H DQF-COSY, ^{1}H−^{15}N HSQC, and ^{1}H−^{13}C HSQC were used for assignment. TOCSY and NOESY spectra were collected using mixing times of 80 and 200 ms, respectively. All spectra were analysed using TopSpin (Bruker, Billerica, MA, USA) and assigned using CCPNMR based on the approach described by Wüthrich [18, 19]. The αH secondary shifts were determined by subtracting random coil ^{1}H NMR chemical shifts from the experimental αH chemical shifts [20]. The 2D NOESY spectra were assigned and an ensemble of structures calculated using the program CYANA [21]. A total of 100 initial structures were calculated using the CYANA program. Torsion-angle restraints predicted using

TALOS-n were used in the structure calculations [22]. The disulfide bond connectivities were included in the calculations for KASH2. Structures were visualized using MOLMOL [23].

5.3.4 Sequence retrieval and alignment

A disulfide-directed β-hairpin (DDH) nucleotide dataset was assembled using sequences from the NCBI-NR database (https://www.ncbi.nlm.nih.gov/). KASH2, a peptide from the common sea star (*Asterius rubens*) and associated with wound-healing and adopting the DDH fold, was used as query sequences for BLAST searches (https://blast.ncbi.nlm.nih.gov/BlastAlign.cgi). Sequences obtained were manually curated before alignment. Translated sequences were aligned in MEGA7 using the MUSCLE algorithm [24, 25]. Universally conserved cysteines conforming to the DDH motif were used as guides to refine the alignment. Regions with gaps in the alignment of >50% were trimmed prior to phylogenetic analyses.

5.3.5 Phylogenetic analysis

The evolutionary relationship between DDH sequences was determined by subjecting the nucleotide dataset to Bayesian and Maximum likelihood analyses. The MPI version of MrBayes 3.2.6 [26, 27] was implemented to perform Bayesian inferences. Analyses were run for a minimum of two hundred million generations using twelve Markov chains across four runs, sampling every 100th tree. Twenty-five per cent of the sampled trees were discarded as burn-in. The log-likelihood score for each of the saved trees was plotted against the number of generations to assess if the log-likelihood scores of the analyses had reached asymptote. Bayesian posterior probability (BPP) was used to evaluate the branch node support. ML analyses were performed using IQ-TREE v1.6.12 [28, 29] with an Edge-proportional partition model coupled with ModelFinder for tree reconstruction using the best-fit partitioning scheme. Bootstrap was used to evaluate branch node support and analyses were run for 100 bootstrap replicates. All phylogenetic trees were midpoint rooted and were visualized using Figtree v1.4.4 (http://tree.bio.ed.ac.uk/software/figtree/).

5.3.6   Human skin normal fibroblast cells

The human skin normal fibroblast cell line 1BR.3.GN was obtained from a European Collection of Authenticated Cell Cultures (ECACC, Porton Down, UK). The 1BR.3.GN cells were grown and maintained in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) (Life Technologies, Melbourne, Australia) containing 1 × antibiotic/antimycotic and 1 × GlutaMAX, supplemented with 10% foetal bovine serum (FBS) (Gibco, Glasgow, Scotland) at 37 °C and 5% CO_2. Cell proliferation assays were performed with DMEM/F12 low nutrient media supplemented with 0.5% FBS.

5.3.7   The real time xCELLigence cell proliferation assay

Cells were seeded at 5000 cells per well in 150 µL of complete media in E-plates (ACEA Biosciences ACEA Biosciences, San Diego, CA, United States) and grown overnight while monitoring with an xCELLigence SP system (ACEA Biosciences), which monitors cellular events in real time by measuring electrical impedance across gold microelectrodes integrated into the base of tissue culture plates. Cells were washed three times with low-nutrient media prior to addition of 150 µL of low-nutrient media and incubated for a minimum of 6 h before further treatment. The peptide was quantified based on absorbance measurements at 214 nm and a predicted extinction coefficient based on the amino acid composition [30, 31]. The pH was checked after dissolution to confirm the solution was still at pH 7.

Treatments were prepared at 8.5 × concentration and added to each of 6 replicate wells in 20 µL, for a final 1 × concentration in a total volume of 170 µL. The xCELLigence system recorded cell indices at intervals of 1 h for 5−6 days following treatment. Readings for the cell index were normalized before treatment, and cell proliferation ratios characterised the relative numbers of cells compared to control cells. Day 4 cell proliferation rates were compared between treatment and control wells, and a one-way ANOVA test was used for multiple comparisons with Holm-Sidak's correction, using GraphPad Prism 9.0.

5.4 Results

5.4.1 Peptide synthesis

The sequence corresponding to KASH2 (Figure 5-1) was synthesised using Fmoc solid-phase peptide synthesis (Fmoc-SPPS). The crude peptide was purified using RP-HPLC, mass analysis was carried out using MALDI mass spectrometry, and a single step oxidation reaction with no selective protection of cysteine residues carried out in ammonium bicarbonate. The folding yield for KASH2 was 25%. The analytical RP-HPLC trace of the oxidation reaction contains three main peaks at different retention times, corresponding to three distinct isomers with the same masses of 3074.5527 Da as shown in Figure 5-2A. The isomers were purified using RP-HPLC.
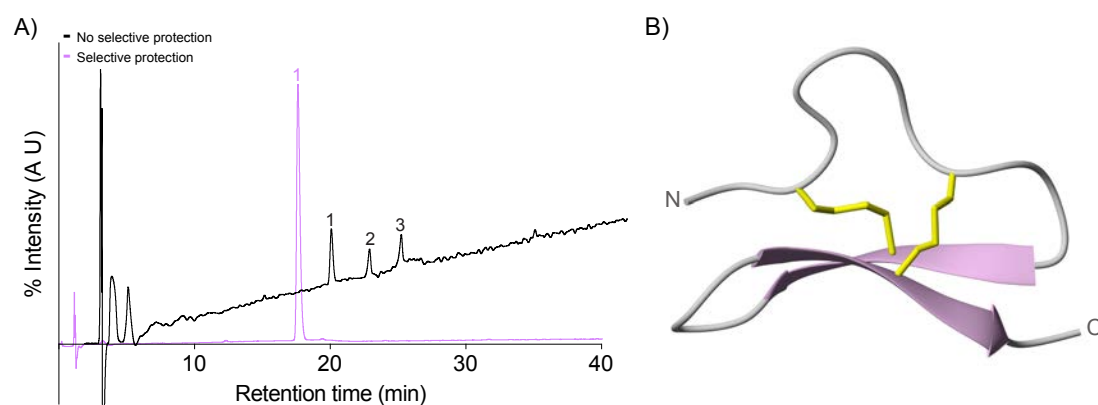
Figure 5-2 Characterisation of KASH2. A) HPLC analysis of oxidised KASH2. B) Three-dimensional representation of KASH2 consistent with the DDH motif. The HPLC traces are offset for clarity.

5.4.2 NMR spectroscopy and structure determination

The structures of the KASH2 isomers were analysed using NMR spectroscopy. The first peak (isomer 1) in the HPLC trace had significant dispersion in the amide region indicative of a well-structured peptide (Appendix 9), in contrast to the other isomers which had broad and overlapped peaks in the NMR spectra. The three-dimensional structure of isomer 1 was determined using torsion-angle dynamics in the CYANA program. The major element of secondary structure was a β-hairpin braced by two disulfide bonds. The β-sheet structure was consistent with the secondary shifts (Appendix 10). The preliminary structures and the temperature coefficient data

(Appendix 11) were used to infer the hydrogen bonds which were then included in the structure calculations (Appendix 12).

Structures were initially calculated with no disulfide-bond restraints and, based on the inter-cysteine distances measured in MOLMOL, the most likely disulfide bond was between Cys II to Cys IV. Consequently, the most likely connectivity involves Cys I-Cys III, Cys II-Cys IV. Structures calculated with this connectivity are shown in Figure 5-2B. The refinement statistics for the final structures are provided in Appendix 13. Analysis of the three-dimensional structure of KASH2 indicates that it contains the DDH motif based on the disulfide connectivity with respect to the β-hairpin; the DDH motif was initially defined by Wang et al. [5]. The structures and chemical shifts of KASH2 have been deposited into the Protein Data Bank and the Biological Magnetic Resonance Data Bank (KASH2—PDB ID: 7U9L, BMRB ID: 30996).

5.4.3 KASH2 synthesis using selective protection of the cysteine residues

To confirm the disulfide connectivity present in isomer 1 of KASH2, selective protection of the cysteine residues was used to direct the folding to form the predicted disulfide connectivity (Cys I-Cys III and Cys II-Cys IV). The fully oxidised peptide was purified using RP-HPLC and shown to have the same retention time on RP-HPLC and TOCSY NMR spectra as isomer 1.

5.4.4 Phylogenetic analyses of DDH peptide identified from the common sea star *Asterias rubens*

The DDH is a part of the disulfide-rich peptide class of proteins, which includes other members such as the inhibitory cystine knot (ICK) [12, 32, 33] and the disulfide-stabilized antiparallel β-hairpin stack (DABS) [12]. The various disulfide linkages in these peptides play a crucial role in determining their native conformation and biological function. They are commonly found in arachnid toxins [32, 34] and have also been detected in other organisms including plants and fungi [35]. KASH2 is the first peptide to contain the DDH fold reported from outside the arachnid lineage, and warrants further investigation to understand the evolutionary origin and diversification of disulfide-rich peptides.

In our in-depth searches against various sea star genomes and other sequences from public repositories, we could not retrieve homologues for KASH2. Very limited

positive hits in certain sea star genomes turned out to be characterised by very poor coverage and therefore we could not use these in our analyses. The only exceptions were the sequences that we retrieved from the *A. rubens* genome. When compared and aligned with the known DDH sequences from arachnids, we found that the sequences from *A. rubens* are relatively short in length, and share very little sequence similarity with their arachnid DDH counterparts. The only exception to this is the four corresponding cysteine residues. The three-dimensional structure of KASH2 is found it to be consistent with the DDH motif, as it adopts the typical disulfide bridging [C1-C3] [C2-C4] pattern, similar to U1-LITX-Lw1a identified from the scorpion *Liocheles waigiensis* (Figure 5-3).
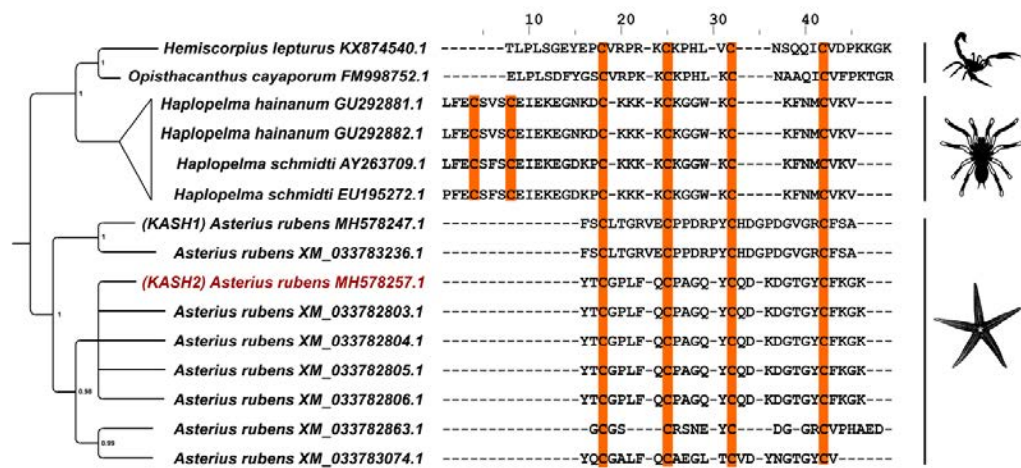
Figure 5-3 The Bayesian phylogeny and sequence alignment of DDH peptides from *A. rubens* and Arachnida. This figure depicts the sequence alignment of the mature DDH peptides from *A. rubens* and Arachnida. Here, the KASH2 peptide from *A. rubens* is shown in red, while the cysteine residues involved in the formation of the DDH fold are highlighted in orange.

Further, to understand their evolutionary history, we performed phylogenetic reconstructions of all known peptides from arachnids and sequences retrieved from *A. rubens* containing the DDH motif. We observed that the DDH-containing sequences identified from *A. rubens* form a distinct clade from the known arachnid DDH sequences. Based on our phylogenetic analyses and the sequence comparisons of DDH containing peptides, we conclude that the presence of DDH motifs found in the European or common sea star is likely a result of convergent evolution (Appendix 14).

5.4.5 Cell proliferation monitoring in real time using xCELLigence

The effect of KASH2 on cell proliferation of 1BR.3.GN fibroblasts was assessed using an xCELLigence system. KASH2 displayed no significant cell proliferation on fibroblast cell growth compared to the negative control peptide (a 21-residue tropomyosin derived peptide) (Figure 5-4). A tropomyosin fragment, previously shown to have no activity in a cell proliferation assay was used as a negative control. Importantly, this peptide was produced in the laboratory using the same conditions to the KASH2.
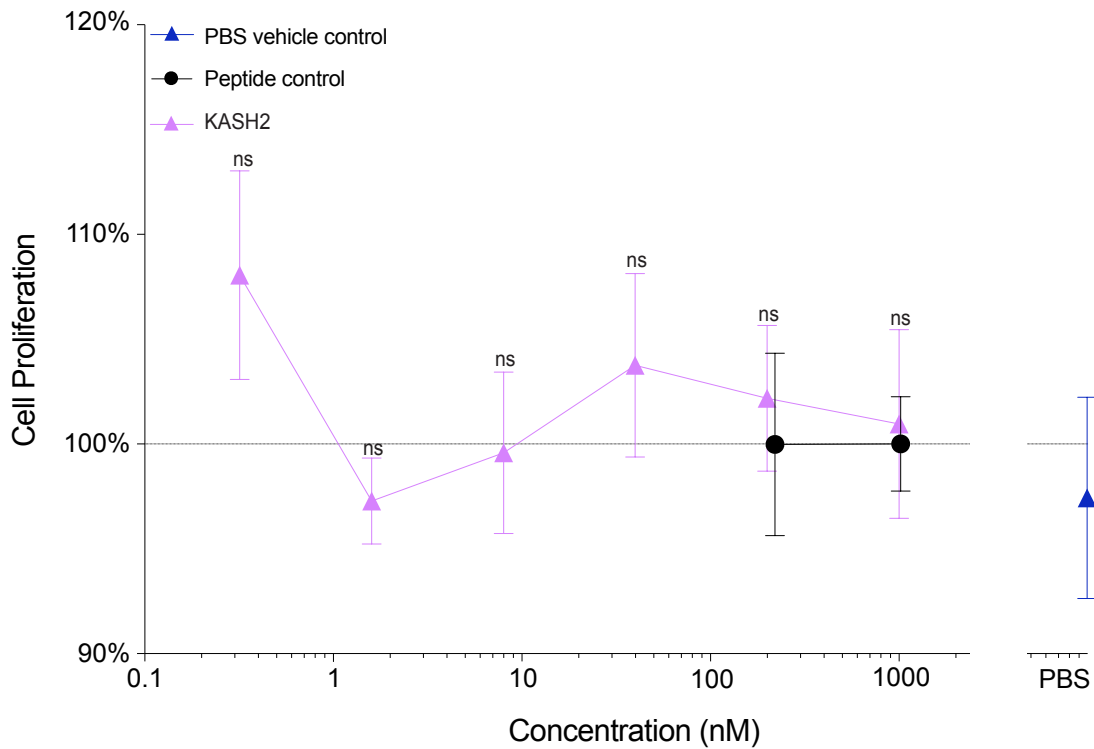
Figure 5-4 KASH2 cell assay. Cell proliferation assay of KASH2 using xCELLigence presenting their proliferative effect on human normal fibroblast cells relative to peptides control. Treatments were measured with six replicates from three independent experiments. Cell index was measured day 4 after treatment. The proliferation rates relative to peptide control are plotted as mean ± SEM bars. Data were analysed by one-way ANOVA against peptide control (Not significant = ns).

5.5 Discussion

Several proteins from *Asterias rubens* have recently been shown to be upregulated in response to injury, including proteins with no known function or structural information [4]. Determination of the three-dimensional structure of one of the mature peptides predicted from the identified precursor proteins has provided insight into the molecular evolution of a disulfide-rich structural motif.

The synthetic version of KASH2 produced three isomers during the oxidative folding reaction, which most likely correspond to the three possible disulfide connectivities for a four-cysteine containing peptide. For small disulfide-rich peptides, it is quite common for the earliest eluting peak on RP-HPLC to correspond to the natively folded peptide. This appears to be the case for KASH2, as the earliest eluting peak was the major form and displays a well-defined structure. Preliminary structural analyses of this isomer suggested that the disulfide connectivity was Cys I-Cys III and Cys II-Cys IV, and this was subsequently confirmed by selective protection of the cysteine residues.

Analysis of the three-dimensional structure of KASH2 with the VAST [36] structural similarity program indicated that KASH2 does not have significant structural similarity to any characterised peptides. However, toxins are well known to contain disulfide-rich motifs and we noticed that KASH2 has the same disulfide connectivity and similar fold to the scorpion toxin U1-LITX-Lw1a [9] as shown in Figure 5-5A. U1-LITX-Lw1a has more residues than KASH2 and contains an additional ß-hairpin, but the RMSD between residues 17-23, 28-30 in U1-LITX-Lw1a, and 9-15, 23-25 in KASH2 is 0.532, highlighted in the overlay in Figure 5-5B and -5C. U1-LITX-Lw1a was the first native peptide shown to adopt a single-domain DDH fold [9]. The DDH fold has also been reported in the Australian funnel-web spider toxin, J-ACTX-Hv1c [5, 6].

A

| U1-LITX-Lw1a | DFPLSKEYES**C**VRPRK**C**KPPLK**C**NKAQI---**C**VDPNKGW |
| Kartesh-2 | YT**C**GPLFQ**C**PAGQY**C**QDKDGTGY**C**FKGK |

B

C
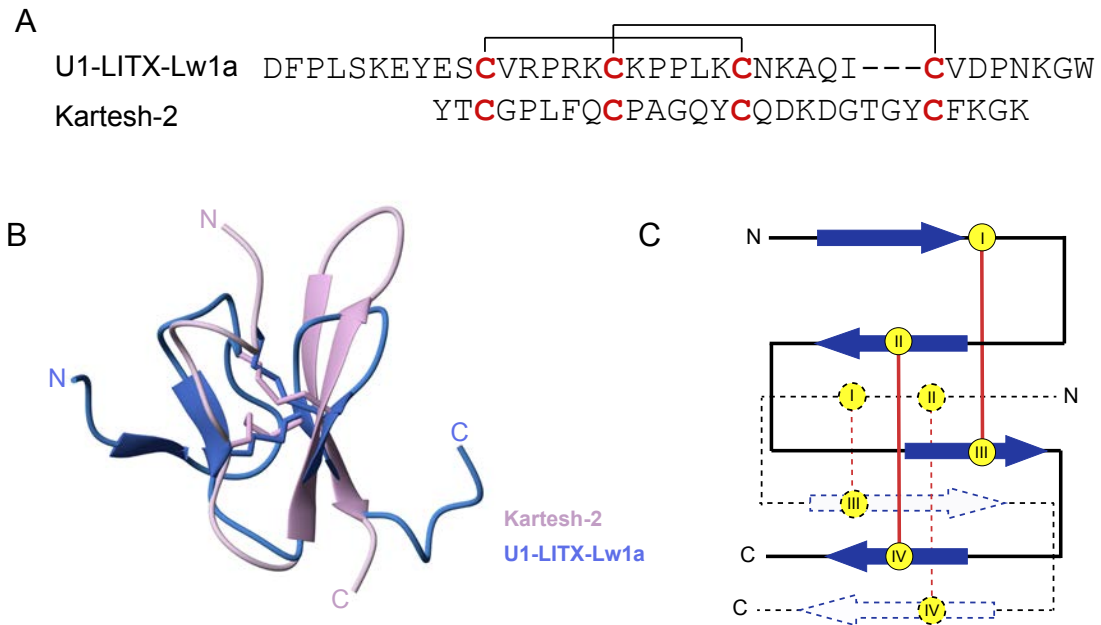
Figure 5-5 Conserved structural fold between sea star *Asterias rubens* and scorpion *Liocheles waigiensis* peptides. A) Alignment of the sequences demonstrating the disulfide connectivity. B) Overlay of the three-dimensional structures, superimposed over the cysteine residues. U1-LITX-Lw1a PDB ID: 2KJY. C) Graphical representation of the DDH motif, with the disulfide and β-sheet shown in red and blue, respectively. The KASH2 DDH motif is shown using dashed line.

The structural evolution study of KASH2 revealed that arachnids and Asteroidea have been phylogenetically separated for nearly 560 million years. While we cannot confidently rule out the common origin of disulfide-rich peptides in these organisms, which was, perhaps, followed by an extensive diversification, independent origin seems to be parsimonious. It is also likely that the two lineages adopted the DDH scaffold from the same/related cysteine-rich scaffold subsequent to their divergence from the most recent common ancestor, but further analyses, such as gene synteny and the identification of similar disulfide-rich peptides along the phylogenetic tree leading to Arachnida and Asteroidea, can help in answering this question more precisely. However, with the existing evidence, the independent origin of DDH in sea star and arachnids seems to be the most parsimonious explanation.

There has been speculation regarding the evolution of the ICK motif, with one hypothesis being that the DDH motif is an ancestral fold of the ICK motif [5]. However, a subsequent study suggested that although the structural data of U1-LITX-

Lw1a support an ancestral link between the ICK and DDH motifs, molecular evolution data and the apparent restriction of single-domain DDH peptides to arachnids suggests that DDH is a derived ICK and that this derivation has occurred on multiple occasions [32]. Consequently, our finding of a single-domain DDH peptide in sea stars might have implications for the evolution of the ICK fold.

As sea stars have been shown to have remarkable abilities to regenerate tissue [37, 38], it was of interest to investigate the biological activity of KASH2 in terms of cell proliferation. The lack of activity of KASH2 on the proliferation of fibroblast has several implications, but might simply be a reflection of the difference between human and sea star tissue regeneration processes, or that KASH2 is not directly related to wound healing in the sea star. Nevertheless, it is a preliminary screen and further analysis, and investigations are required to understand the potential functions of KASH2.

5.6 Conclusions

Overall, we have identified the first DDH structural motif characterised from the sea star *Asterias rubens*. We have significantly expanded the sequence diversity known to adopt this fold. Although KASH2 was upregulated in response to injury, we show here that it did not have any effect on cell proliferation of human fibroblasts. There are numerous other roles it might play in injury, which remain to be determined. Disulfide-rich scaffolds, such as the DDH fold, have previously been shown to be useful as scaffolds in drug design and agricultural applications [10, 39-41], therefore our finding might have implications for a range of studies.

5.7 Funding

5.8 Acknowledgements

5.9 References

1. Gupta, A., et al., *Wound healing in guinea pigs after topical application of starfish Pentaceraster regulus extract.* J. Wound Care, 2008. 17(10): p. 441-4.

2. Franco, C.F., et al., *Radial nerve cord protein phosphorylation dynamics during starfish arm tip wound healing events.* Electophoresis, 2012. 33(24): p. 3764-3778.

3. Thorndyke, M.C. and Carnevali, C., *Regeneration neurohormones and growth factors in echinoderms.* Can. J. Zool., 2001. 79(7): p. 1171-1208.

4. Shabelnikov, S.V., et al., *Injury affects coelomic fluid proteome of the common starfish, Asterias rubens.* J. Exp. Biol., 2019. 222(6): p. jeb198556.

5. Wang, X.-h., et al., *Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge.* Nat. Struct. Mol. Biol., 2000. 7(6): p. 505.

6. Shu, Q., et al., *The structure of spider toxin huwentoxin-II with unique disulfide linkage: Evidence for structural evolution.* Protein Sci., 2002. 11(2): p. 245-252.

7. Cremonez, C.M., et al., *Structural and functional elucidation of peptide Ts11 shows evidence of a novel subfamily of scorpion venom toxins.* Toxins, 2016. 8(10): p. 288.

8. Sunagar, K., et al., *Evolution stings: The origin and diversification of scorpion toxin peptide scaffolds.* Toxins, 2013. 5(12): p. 2456-2487.

9. Smith, J.J., et al., *Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif.* Proc. Natl. Acad. Sci., 2011. 108(26): p. 10478-83.

10. Smith, J.J., et al., *The insecticidal potential of venom peptides.* Cell. Mol. Life. Sci., 2013. 70(19): p. 3665-3693.

11. Chen, J., et al., *Cloning and genomic characterization of a natural insecticidal peptide LaIT1 with unique DDH structural fold.* J. Biochem. Mol. Toxicol., 2015. 29(5): p. 207-12.

12. Pineda, S.S., et al., *Structural venomics reveals evolution of a complex venom by duplication and diversification of an ancient peptide-encoding gene.* Proc. Natl. Acad. Sci., 2020. 117(21): p. 11399-11408.

13. Smout, M.J., et al., *A granulin-like growth factor secreted by the carcinogenic liver fluke, Opisthorchis viverrini, promotes proliferation of host cells.* PLoS Pathog., 2009. 5(10): p. e1000611.

14.   Smout, M.J., et al., *Carcinogenic parasite secretes growth factor that accelerates wound healing and potentially promotes neoplasia.* PLoS Pathog., 2015. 11(10): p. e1005209.

15.   Bansal, P.S., et al., *Development of a potent wound healing agent based on the liver fluke granulin structural fold.* J. Med. Chem., 2017. 60(10): p. 4258-4266.

16.   Dastpeyman, M., et al., *Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing.* J. Med. Chem., 2018. 61(19): p. 8746-8753.

17.   Takjoo, R., et al., *Folding of truncated granulin peptides.* Biomolecules, 2020. 10(8): p. 1152.

18.   Wüthrich, K., *NMR studies of structure and function of biological macromolecules (Nobel Lecture).* J. Biomol. NMR, 2003. 27(1): p. 13-39.

19.   Vranken, W.F., et al., *The CCPN data model for NMR spectroscopy: Development of a software pipeline.* Proteins, 2005. 59(4): p. 687-696.

20.   Wishart, D.S., et al., *^{1}H, ^{13}C and ^{15}N chemical shift referencing in biomolecular NMR.* J. Biomol. NMR, 1995. 6(2): p. 135-140.

21.   Güntert, P., *Automated NMR structure calculation with CYANA.* Methods Mol. Biol., 2004. 278: p. 353-378.

22.   Shen, Y. and Bax, A., *Protein structural information derived from NMR chemical shift with the neural network program TALOS-N.* Methods Mol. Biol., 2015. 1260: p. 17-32.

23.   Koradi, R., et al., *MOLMOL: A program for display and analysis of macromolecular structures.* J. Mol. Graph., 1996. 14(1): p. 51-55.

24.   Edgar, R.C., *MUSCLE: multiple sequence alignment with high accuracy and high throughput.* Nucleic Acids Res., 2004. 32(5): p. 1792-7.

25.   Kumar, S., et al., *MEGA7: Molecular evolutionary genetics analysis version 7.0 for bigger datasets.* Mol. Biol. Evol., 2016. 33(7): p. 1870-4.

26.   Altekar, G., et al., *Parallel Metropolis coupled Markov chain Monte Carlo for Bayesian phylogenetic inference.* Bioinformatics, 2004. 20(3): p. 407-15.

27.   Ronquist, F., et al., *MrBayes 3.2: Efficient Bayesian phylogenetic inference and model choice across a large model space.* Syst. Biol., 2012. 61(3): p. 539-42.

28. Nguyen, L.T., et al., *IQ-TREE: A fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies.* Mol. Biol. Evol., 2015. 32(1): p. 268-74.

29. Kalyaanamoorthy, S., et al., *ModelFinder: Fast model selection for accurate phylogenetic estimates.* Nat. Methods, 2017. 14(6): p. 587-589.

30. Gill, S.C. and von Hippel, P.H., *Calculation of protein extinction coefficients from amino acid sequence data.* Anal. Biochem., 1989. 182(2): p. 319-26.

31. Hilario, E.C., et al., *An improved method of predicting extinction coefficients for the determination of protein concentration.* PDA J. Pharm. Sci. Technol., 2017. 71(2): p. 127-135.

32. Undheim, E.A.B., et al., *Toxin structures as evolutionary tools: Using conserved 3D folds to study the evolution of rapidly evolving peptides.* BioEssays, 2016. 38(6): p. 539-548.

33. Nielsen, L.D., et al., *The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework.* J. Biol. Chem., 2019. 294(22): p. 8745-8759.

34. Zhu, L.M., et al., *Origin of neurotoxins from defensins.* Acta Physiol. Sin., 2015. 67(3): p. 239-47.

35. Zhu, S., et al., *Evolutionary origin of inhibitor cystine knot peptides.* FASEB J., 2003. 17(12): p. 1765-1767.

36. Gibrat, J.-F., et al., *Surprising similarities in structure comparison.* Curr. Opin. Struct. Biol., 1996. 6(3): p. 377-385.

37. Dai, Y., et al., *Tissue extract fractions from starfish undergoing regeneration promote wound healing and lower jaw blastema regeneration of zebrafish.* Sci. Rep., 2016. 6: p. 38693.

38. Ben Khadra, Y., et al., *Wound repair during arm regeneration in the red starfish Echinaster sepositus.* Wound Repair Regen., 2015. 23(4): p. 611-22.

39. King, G.F., *Venoms as a platform for human drugs: translating toxins into therapeutics.* Expert Opin. Biol. Ther., 2011. 11(11): p. 1469-1484.

40. King, G.F. and Hardy, M.C., *Spider-venom peptides: Structure, pharmacology, and potential for control of insect pests.* Annu. Rev. Entomol., 2013. 58(1): p. 475-496.

41. Windley, M.J., et al., *Spider-venom peptides as bioinsecticides.* Toxins, 2012. 4(3): p. 191-227.

Chapter 6. Conclusions and Future Directions

6.1 Conclusions

Chronic wounds are a major health problem worldwide, and new treatments are needed that are effective and relatively low-cost. Peptides have the potential to fill this need, but more information on structure/function relationships is required to direct drug development studies. The focus of this thesis was on characterising naturally occurring peptides with potential as wound healing agents.

The hypothesis in Chapter 2 of this thesis was that one of the inter-cysteine loops in the parasite-derived granulin *Ov*-GRN-1 is responsible for the wound healing properties. We confirmed this hypothesis by showing a peptide corresponding to loop 2 of *Ov*-GRN-1 promoted significant cell proliferation activity. A similar downsizing approach has previously been applied to the scorpion venom peptide chlorotoxin, where fragments of the peptide were synthesised and analysed for cell migration effects [1]. A peptide corresponding to loop 4 of chlorotoxin inhibited cell migration, in contrast to the other three fragments synthesised. In a separate study an analogue of loop 4 of chlorotoxin has been shown to transport nanoparticles across the blood-brain barrier [2]. Although these studies were carried out in parallel the development of loop 4 of chlorotoxin as a drug lead, highlights the potential of identifying bioactive regions of disulfide-rich peptides [2].

Both the truncated loop 2 peptide of *Ov*-GRN-1 and the loop 4 chlorotoxin truncated peptide were unstructured in solution indicating that a well-defined structure is not critical for bioactivity. This phenomenon has been previously observed for other peptides and their targets including angiotenin II, one of the most potent systemic vasopressors known [3] which lacks tertiary structure [4]. It is possible that structure might be induced, upon binding to a biological target.

The overall objective of Chapters 3, 4, and 5 was to explore the folding of truncated and full-length peptides/proteins into bioactive peptides. There are two main strategies in oxidative folding, regioselective directed and non-selective (single-step) formation of the disulfide bonds [5]. Regioselective oxidation requires knowledge, or at least prediction, of the desired disulfide connectivity. However, for non-selective formation the conditions chosen for folding, such as temperature, pH and hydrophobicity can have a significant influence of the isomers produced. For example, the cyclic peptide kalata B1 folded more efficiently in the presence of hydrophobic

solvents as the native structure has surface exposed hydrophobic residues [6]. To address the impact of the *in vitro* conditions on folding, several conditions were used in the folding of the disulfide-rich peptides in this thesis.

The objective of Chapters 3 and 4 was to explore the structures and functions of the N- and C-terminal regions of granulin peptides from the zebrafish model organism. We aimed to explore how effectively the engineered peptides folded into well-defined conformations and therefore chose an undirected disulfide bond formation strategy. The results from Chapter 3 showed that a naturally occurring paragranulin and a truncated zebrafish granulin adopt the mini-granulin fold initially identified in the *Ov*-GRN-1 granulin truncated peptide [7, 8]. However, there was distinct differences in folding efficiency and bioactivity between the two zebrafish peptides. Despite the differences in folding and bioactivity, this study indicates that the N-terminal region of granulins folding into an independent domain is likely to be a common feature of granulins. By contrast, Chapter 4 showed that the C-terminal region of a zebrafish granulin does not fold independently into a well-defined structure, and it also had no significant inhibitory effect on cell growth.

The structural difference between the N- and C-terminal regions suggests fundamental distinctions in the folding of the two halves of the granulin framework. In some ways this is consistent with the structures of some of the full-length granulin peptides, which display well defined N-terminal regions but disordered C-terminal regions [9, 10]. Moreover, this appears consistent with the lack of a C-terminal paragranulin, in contrast to the naturally occurring paragranulins (N-terminal half) present in a range of species.

Although a C-terminal paragranulin equivalent has not been found, there are examples of granulins that do not contain the full 12-cysteine residue framework, but rather only contain ten cysteine residues. The sequences of zebrafish granulin AaG and human granulin G [9, 10], both of which only contain ten cysteine residues are shown in Appendix 15, along with human granulin A and zebrafish granulin AaE, which contains 12 cysteine residues. The 10-cysteine framework lacks the equivalent of Cys IV and Cys VII in the 12-cysteine framework [9, 10]. The structural and functional consequences of granulins that only contain ten cysteine residues are poorly understood, but *in vitro* studies have shown that human granulin G does not fold into

a well-defined conformation and displays weak growth inhibitory activity in a breast cancer cell line [9, 11]. Similarly, zebrafish granulin AaG lacks structure and had low *in vitro* production yields [10]. It appears possible that the disordered structures of zebrafish granulin AaG and human granulin G are associated with the removal of a disulfide bond, but this has yet to be fully explored.

Although granulins are an intriguing family of peptides with significant potential in the development of wound healing agents, nature offers a broad range of other peptides with potential in this area. Organisms that can regenerate are a relatively untapped resource in the field of wound healing and was the subject of Chapter 5. The specific aim of Chapter 5 was to provide information regarding the structure and function of KASH2, a disulfide rich peptide, with injury-related effects identified in the common sea star, *Asterias rubens*. Numerous proteins/peptides from the coelomic fluid of this sea star have been shown to be potentially involved in wound healing; however, there was limited structural or functional information available. Interestingly, KASH2 had no effect on the cell growth of human normal skin fibroblast cells; but structural analysis demonstrated the presence of an ancestral protein fold termed the disulfide directed β-hairpin (DDH). We show the DDH fold is more widespread than originally thought [12-14], but is likely the result of convergent evolution. Further study is required to illuminate the role of KASH2 in the sea star.

The approaches used in this thesis, in terms of using synthetic methods to produce and dissect naturally occurring peptides, have the advantage of producing chemically clean products, in contrast to cell-based approaches. However, in cell-based systems, chaperones can assist peptide and protein folding [15] [16], whereas formation of disulfide bonds can be complicated *in vitro*. Regardless of the methods used to produce the peptides, for naturally occurring peptides such as KASH2 where native peptide is not available, the native disulfide connectivity is not known, and the analyses are restricted to the isomers produced *in vitro*. Despite this limitation, this thesis provides insight into the structures and bioactivities of the granulin framework and the DDH motif. Both of these structural frameworks are widespread throughout nature and therefore the results are likely to have implications for a range of fields. It has highlighted the potential of utilizing the downsizing approach in identifying

bioactive regions of disulfide-rich peptides, which might be useful in the design of novel wound healing agents.

6.2 Future directions

The research presented in the previous chapters opens up many promising possibilities for future research. This section looks at some of these future directions. Perhaps the most important direction is related to bioactivity, target identification and mechanism of action as the mechanism of action of the granulin and sea star peptides has not been determined. In the case of GRN-L2, it would be of interest to determine if the cell proliferation effects are observed with different cells lines and if they are correlated with *in vivo* activity in a wound healing model. Similarly, it would be of interest to examine the cell proliferation effects of the zebrafish derived peptides on fish-derived cell lines as the lack of activity observed in the current study might be related to the choice of human cell lines. Moreover, extracts derived from several sea stars have been shown to have antimicrobial activity [17], making it of interest to determine if KASH2 or the other uncharacterised sea star peptides have this activity. In order to further explore the potential bioactivities associated with this peptide, collecting of coelomic fluid (containing coelomocytes) from live animals and measuring cell growth *in vitro* immediately after collection and addition of KASH2 might be considered.

Following on from bioactivity studies/screens, target identification can be approached by direct biochemical methods, genetic interactions or computational inference. In many cases combinations of approaches may be required to fully characterise on-target and off-target effects and to understand mechanisms of action [18]. Computational prediction can be used to describe the interactions of peptides with a biological target such as high-content screening, which is transforming drug discovery by allowing simultaneous measurement of multiple features of cellular phenotype that are relevant to therapeutic and toxic activities of compounds [19]. The machine-based strategies, such as random forest algorithm and Support Vector Machine algorithm, could identify the target, and have higher efficiency and lower cost while maintaining higher accuracy [20].

Protein array technology can also be used to identify possible interactions of engineered peptides with a biological target. These approaches have a wide range of applications including the identification of protein–protein interactions, protein-phospholipid interactions, small molecule targets, and substrates of proteins kinases

[21]. Pull down experiments are also an invaluable *in vitro* technique to detect physical interactions between two or more proteins and confirm a predicted protein–protein interaction or identify novel interacting partners [22, 23].

Additional engineering studies related to GRN-L2 might also prove useful in elucidating the role of this peptide sequence in bioactivity and in drug design studies. To recognise the residues important for bioactivity, characterisation, and manufacturing of over-lapping fragments from the N-terminal region of *Ov*-GRN-1 might be useful. Furthermore, production of granulin hybrid sequences might also be useful. The rationale for this suggestion is that the inter-cysteine loop corresponding to GRN-L2 is not conserved in granulin proteins and if it is playing a role in the potent wound healing properties of *Ov*-GRN-1 protein, it would be of interest to take a granulin protein that does not promote cell growth and replace the inter-cysteine loops and measure the wound healing properties of the hybrid protein. There is precedent for the production of hybrid granulin proteins. A modified hybrid granulin peptide, Atsttrin, composed of three fragments of progranulin, was even more effective than full-length PGRN in diminishing the pro-inflammatory signaling cascade stimulated by TNF-α [24].

Another possibility for future work is to examine the effects of grafting GRN-L2 into a stable scaffold. Peptide grafting is reproduction of the structural and functional properties of active epitopes on an appropriate scaffold or a template with high conformational and proteolytic stability, providing a potential solution for overcoming poor stability and bioavailability of the peptide epitopes [25]. For example, grafting the tripeptide MC-12 into the SFTI-1 scaffold improves its therapeutic efficacy in a murine model of chemically induced acute colitis while also improving its *in vitro* stability [26].

It would be of interest to individually make each of the three disulfide isomers of ZF-C$_{20\text{-}2s}$ via directed disulfide bond formation to conclusively determine the disulfide connectivity present in each of the three peaks. These isomers could be subsequently used for analysis in biological assays. Moreover, structural and biological analysis of the S-reduced and carbamidomethylated (capped) forms of the linear peptide could provide further insight into the structure/function relationships.

KASH2 is one of the many peptides and proteins identified in the sea star *A. rubens* peptides, that have potential as a wound-healing agents. Further structural and functional studies will provide fundamental information in this field. Determination of the structure of KASH1 and CHUPA peptides (CHU1, CHU2, and CHU3) will provide a better understanding of structure activity relationships in this particular sea star [27].

Overall, the future directions given in this section might further advance the potential of both granulin and sea star derived peptides as drug leads.

6.3 References

1. Dastpeyman, M., et al., *A C-Terminal fragment of chlorotoxin retains bioactivity and inhibits cell migration.* Front. Pharmacol., 2019. 10: p. 250.

2. Díaz-Perlas, C., et al., *From venoms to BBB-shuttles. MiniCTX3: a molecular vector derived from scorpion venom.* Chem. Comm., 2018. 54(90): p. 12738-12741.

3. Benigni, A., et al., *Angiotensin II revisited: New roles in inflammation, immunology and aging.* EMBO Mol. Med., 2010. 2(7): p. 247-257.

4. Spyroulias, G.A., et al., *Comparison of the solution structures of angiotensin I & II. Implication for structure-function relationship.* Eur. J. Biochem., 2003. 270(10): p. 2163-73.

5. Moroder, L., et al., *Oxidative folding of cystine-rich peptides vs regioselective cysteine pairing strategies.* Biopolymers, 1996. 40(2): p. 207-34.

6. Daly, N.L., et al., *Chemical synthesis and folding pathways of large cyclic polypeptides: Studies of the cystine knot polypeptide kalata B1.* Biochemistry, 1999. 38(32): p. 10606-14.

7. Bansal, P.S., et al., *Development of a potent wound healing agent based on the liver fluke granulin structural fold.* J. Med. Chem., 2017. 60(10): p. 4258-4266.

8. Dastpeyman, M., et al., *Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing.* J. Med. Chem., 2018. 61(19): p. 8746-8753.

9. Tolkatchev, D., et al., *Structure dissection of human progranulin identifies well-folded granulin/epithelin modules with unique functional activities.* Protein Sci., 2008. 17(4): p. 711-724.

10. Wang, P., et al., *Structure dissection of zebrafish progranulins identifies a well-folded granulin/epithelin module protein with pro-cell survival activities.* Protein Sci., 2018. 27(8): p. 1476-1490.

11. Dastpeyman, M., et al., *Folding of granulin domains.* Pept. Sci., 2018. 110(3).

12. Wang, X.-h., et al., *Discovery and characterization of a family of insecticidal neurotoxins with a rare vicinal disulfide bridge.* Nat. Struct. Mol. Biol., 2000. 7(6): p. 505.

13. Smith, J.J., et al., *Unique scorpion toxin with a putative ancestral fold provides insight into evolution of the inhibitor cystine knot motif.* Proc. Natl. Acad. Sci., 2011. 108(26): p. 10478-83.

14. Undheim, E.A.B., et al., *Toxin structures as evolutionary tools: Using conserved 3D folds to study the evolution of rapidly evolving peptides.* BioEssays, 2016. 38(6): p. 539-548.

15. Beissinger, M. and Buchner, J., *How chaperones fold proteins.* Biol. Chem., 1998. 379(3): p. 245-59.

16. Saibil, H., *Chaperone machines for protein folding, unfolding and disaggregation.* Nat. Rev. Mol. Cell Biol., 2013. 14(10): p. 630-642.

17. Haug, T., et al., *Antibacterial activity in Strongylocentrotus droebachiensis (Echinoidea), Cucumaria frondosa (Holothuroidea), and Asterias rubens (Asteroidea).* J. Invertebr. Pathol., 2002. 81(2): p. 94-102.

18. Schenone, M., et al., *Target identification and mechanism of action in chemical biology and drug discovery.* Nat. Chem. Biol., 2013. 9(4): p. 232-40.

19. Young, D.W., et al., *Integrating high-content screening and ligand-target prediction to identify mechanism of action.* Nat. Chem. Biol., 2008. 4(1): p. 59-68.

20. Shangguan, Z. *A review of target identification strategies for drug discovery: from database to machine-based methods*. in J. Phys. Conf. Ser. 2021. Bristol: IOP Publishing.

21. Hall, D.A., et al., *Protein microarray technology.* Mech. Ageing Dev., 2007. 128(1): p. 161-7.

22. Brymora, A., et al., *Protein-protein interactions identified by pull-down experiments and mass spectrometry.* Curr. Protoc. Cell Biol., 2004. Chapter 17: p. Unit 17.5.

23. Louche, A., et al., *Protein-Protein Interactions: Pull-Down Assays.* Methods Mol. Biol., 2017. 1615: p. 247-255.

24. Kleinberger, G., et al., *Mechanisms of granulin deficiency: Lessons from cellular and animal models.* Mol. Neurobiol., 2013. 47(1): p. 337-360.

25. Chittoor, B., et al., *The single disulfide-directed beta-hairpin fold. Dynamics, stability, and engineering.* Biochemistry, 2017. 56(19): p. 2455-2466.

26. Cobos Caceres, C., et al., *An engineered cyclic peptide alleviates symptoms of inflammation in a murine model of inflammatory bowel disease.* J. Biol. Chem., 2017. 292(24): p. 10288-10294.

27. Shabelnikov, S.V., et al., *Injury affects coelomic fluid proteome of the common starfish, Asterias rubens.* J. Exp. Biol., 2019. 222(6): p. jeb198556.

Appendix

Appendix 1. SCIEX TOF/TOF™ 5800 MALDI mass spectra of ZF-N$_{24\_3s}$, ZF-para$_{\_3s}$ and ZF-N$_{24\_2s}$.
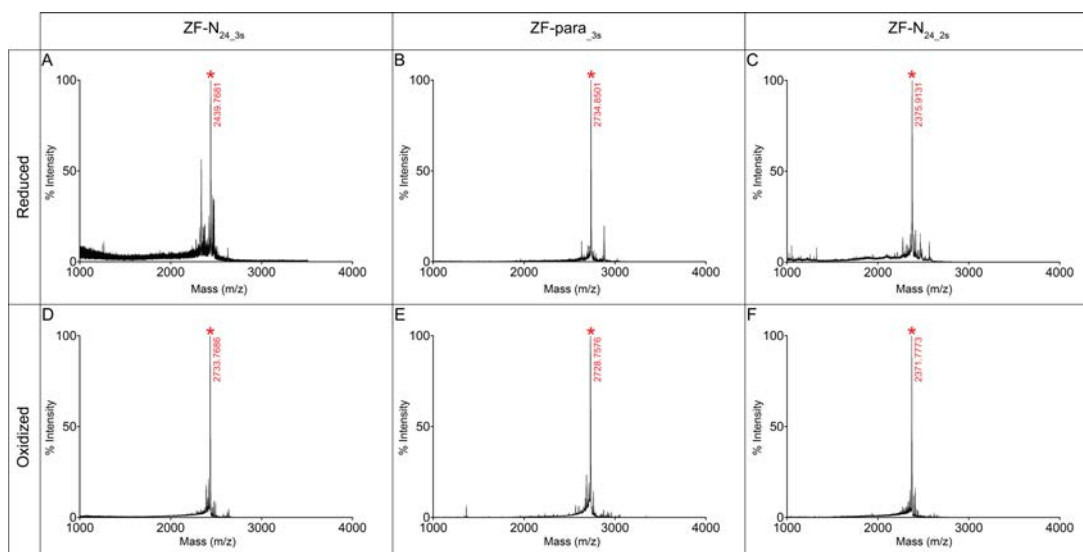

The spectra in panels A, B and C correspond to the reduced peptides and the spectra in panels D, E and F correspond to the oxidised peptides. The masses of reduced and oxidised peptides are highlighted by asterisks (*).

Appendix 2. Hydrogen bond restraints.

ZF-N$_{24\_3s}$	ZF-para$_{\_3s}$
11 GLU HN 8 HIS O	14 CYS HN 22 GLY O
17 SER HN 20 THR O	16 THR HN 20 GLN O
	24 CYS HN 12 THR O
	19 GLY HN 16 THR O

Appendix 3. HPLC analysis of the oxidised A) ZF-N$_{24\_3s}$, B) ZF-para$_{\_3s}$, and C) ZF-N$_{24\_2s}$.
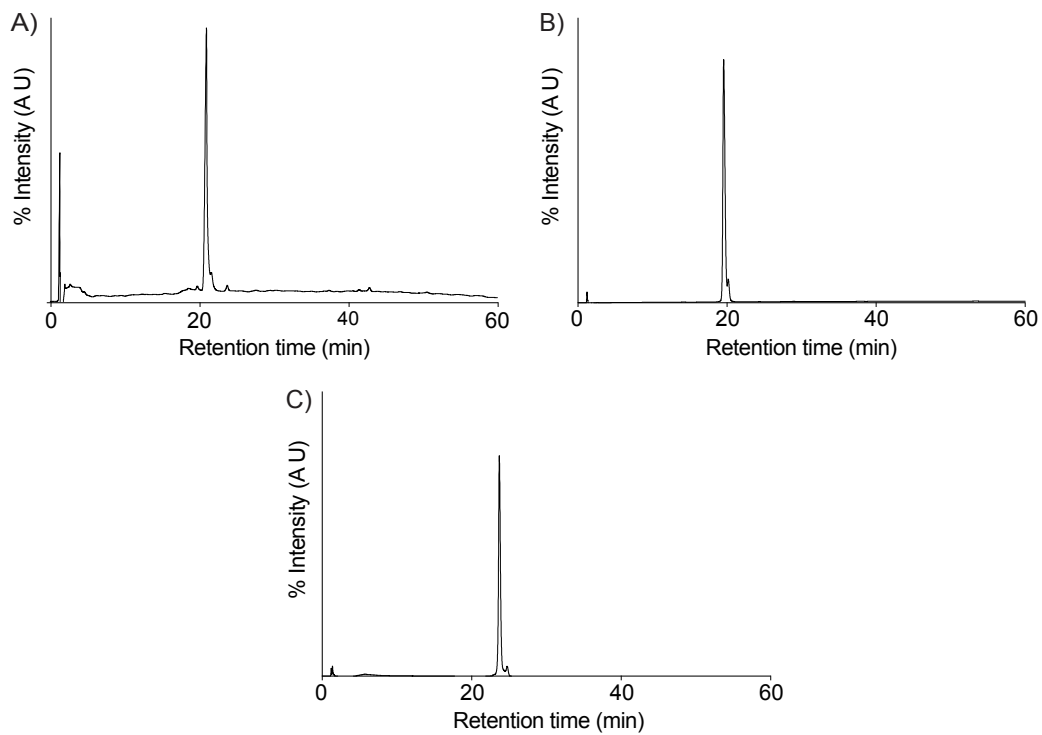
Appendix 4. Cyana target functions for all 15 possible disulfide bond connectivities for each peptide.

	Connectivity	ZF-N$_{24\_3s}$	ZF-para$_{\_3s}$
1	1-13, 7-23, 14-24	0.0254 ± 0.00303	0.10 ± 0.000367
2	1-13, 7-14, 23-24	1.42 ± 0.20	1.74 ± 0.0151
3	1-13, 7-24, 14-23	0.10 ± 0.0397	0.43 ± 0.15
4	1-7, 14-24, 13-23	0.0683 ± 0.0205	0.28 ± 0.0179
5	1-7, 13-24, 14-23	0.0741 ± 0.0228	0.51 ± 0.0359
6	1-7, 13-14, 23-24	4.82 ± 0.00338	10.44 ± 0.0497
7	1-14, 7-13, 23-24	0.99 ± 0.00248	1.11 ± 0.00313
8	1-14, 7-23, 13-24	0.0542 ± 0.0177	3.55 ± 0.0785
9	1-14, 7-24, 13-23	0.11 ± 0.0587	0.17 ± 0.0497
10	1-23, 7-13, 14-24	0.0310 ± 0.00613	0.28 ± 0.0592
11	1-23, 7-14, 13-24	0.31 ± 0.10	1.19 ± 0.0178
12	1-23, 7-24, 13-14	3.94 ± 0.0457	9.36 ± 0.0445
13	1-24, 7-13, 14-23	0.0584 ± 0.0257	0.65 ± 0.0273
14	1-24, 7-14, 13-23	0.68 ± 0.23	1.26 ± 0.0377
15	1-24, 7-23, 13-14	3.88 ± 0.0353	9.12 ± 0.0118

The numbers in red indicate the connectivity with the lowest target function for each peptide.

Appendix 5. Temperature coefficients for ZF-N$_{24\_3s}$.

Residue	$-\Delta\delta_{NH}/\Delta T$ (ppb/K)
11-Glu	-1.584
17-Ser	-1.124

Appendix 6. Temperature coefficients for ZF-para$_{\_3s}$.

Residue	$-\Delta\delta_{NH}/\Delta T$ (ppb/K)
24-Cys	-3.204
14-Cys	-4.052
16-Thr	-4.048
19-Gly	-1.844
18-Thr	-1.962
20-Gln	-3.006

Appendix 7. Overlay of chromatograms of purified reduced and oxidation reactions of ZF-C$_{20\_2s}$.
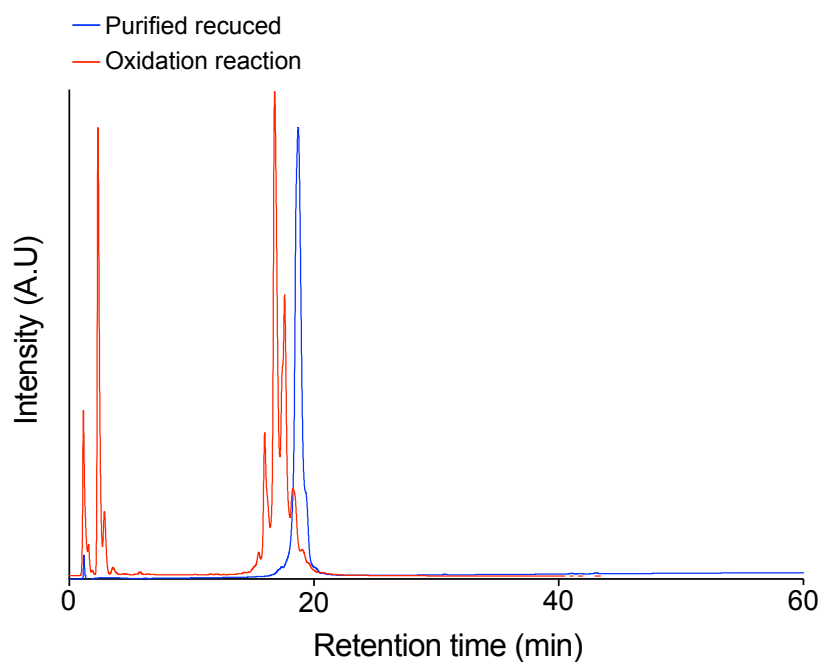
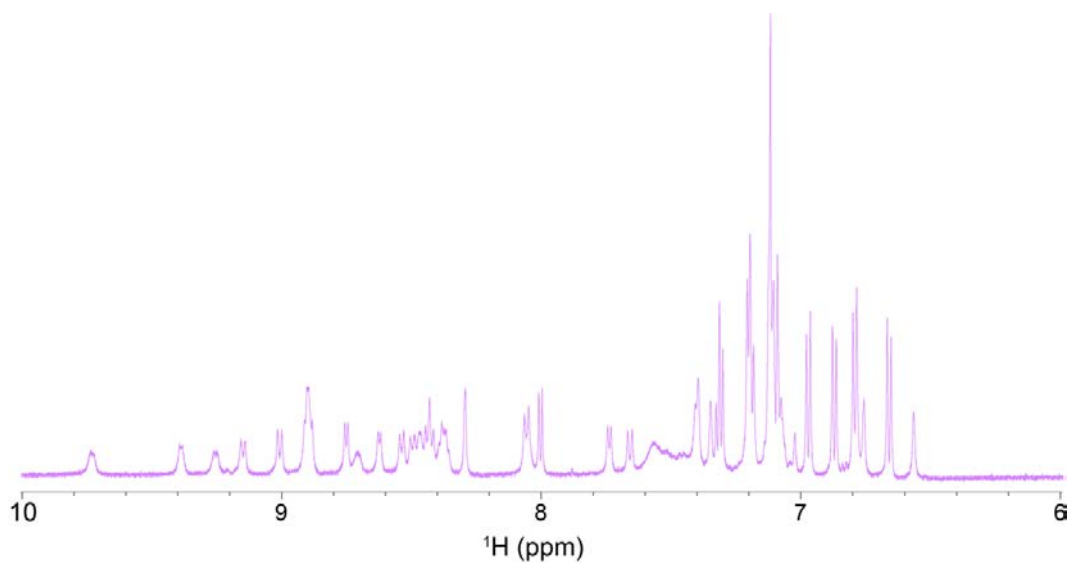The purified reduced and oxidation reactions are shown in blue and red, respectively.

Appendix 8. Real-time cell proliferation analysis for three isomers of ZF-C$_{20\_2s}$ using xCELLigence.
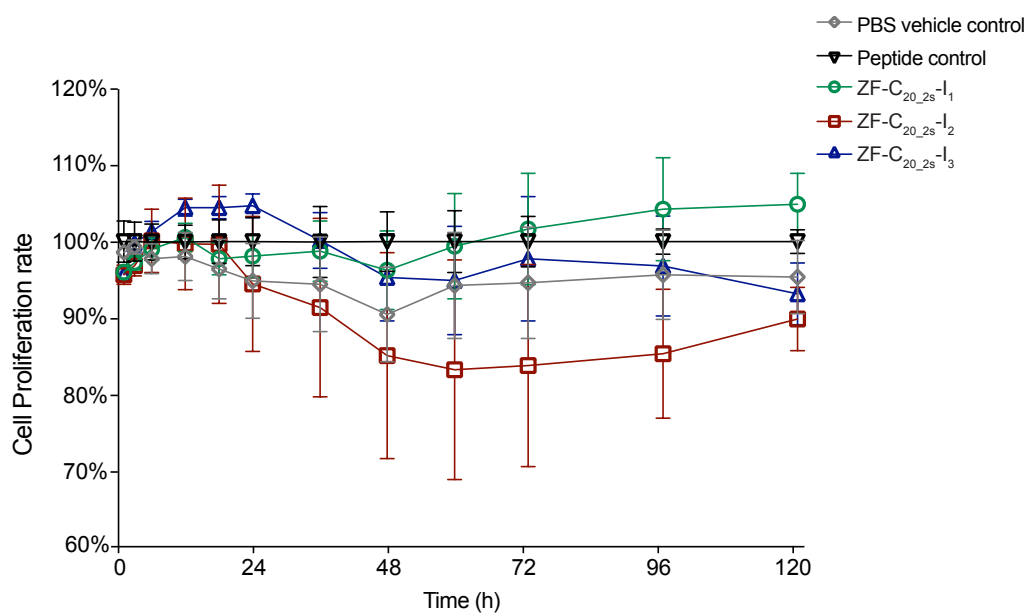
Fibroblast cells were treated with three isomers of ZF-C$_{20\_2s}$ at concentrations of 200 nM.

Appendix 9. KASH2 ^{1}H NMR (600 MHz, 90% H₂O/10% D₂O v/v) spectrum.

Spectrum is shown from 6 to 10 ppm; this region contains primarily the amide protons.

Appendix 10. αH Secondary shifts for KASH2.

The sequence of the peptides is given at the bottom of the diagram.

Appendix 11. Temperature coefficients for KASH2.

Residue	$-\Delta\delta_{NH}/\Delta T$ (ppb/K)
23-Tyr	-0.74
25-Phe	-1.02

Appendix 12. Hydrogen bond restraints for KASH2.

25 PHE HN 14 TYR O
23 TYR HN 16 GLN O

Appendix 13. Structural statistics for KASH2.

Experimental restraints	
All	149
Intraresidue, |i-j|=0	59
Sequential, |i-j|=1	61
Medium range, 1 <|i-j| < 5	6
Long range, |i-j| >=5	23
Disulfide-bond restraints (3 restraints per bond)	6
Dihedral-angle restraints	30
Hydrogen bond restraints (2 restraints per bond)	4

R.m.s. deviations from mean coordinate structure (Å)	
Backbone atoms	2.12 ± 0.6
All heavy atoms	3.12 ± 0.70
Backbone atoms (13-26)	0.50 ± 0.18
All heavy atoms (13-26)	1.36 ± 0.29
Ramachandran Statistics	
% In most favoured region	81.8
% Residues in additionally allowed regions	18.2

Appendix 14. A Bayesian phylogeny of DDH motifs.
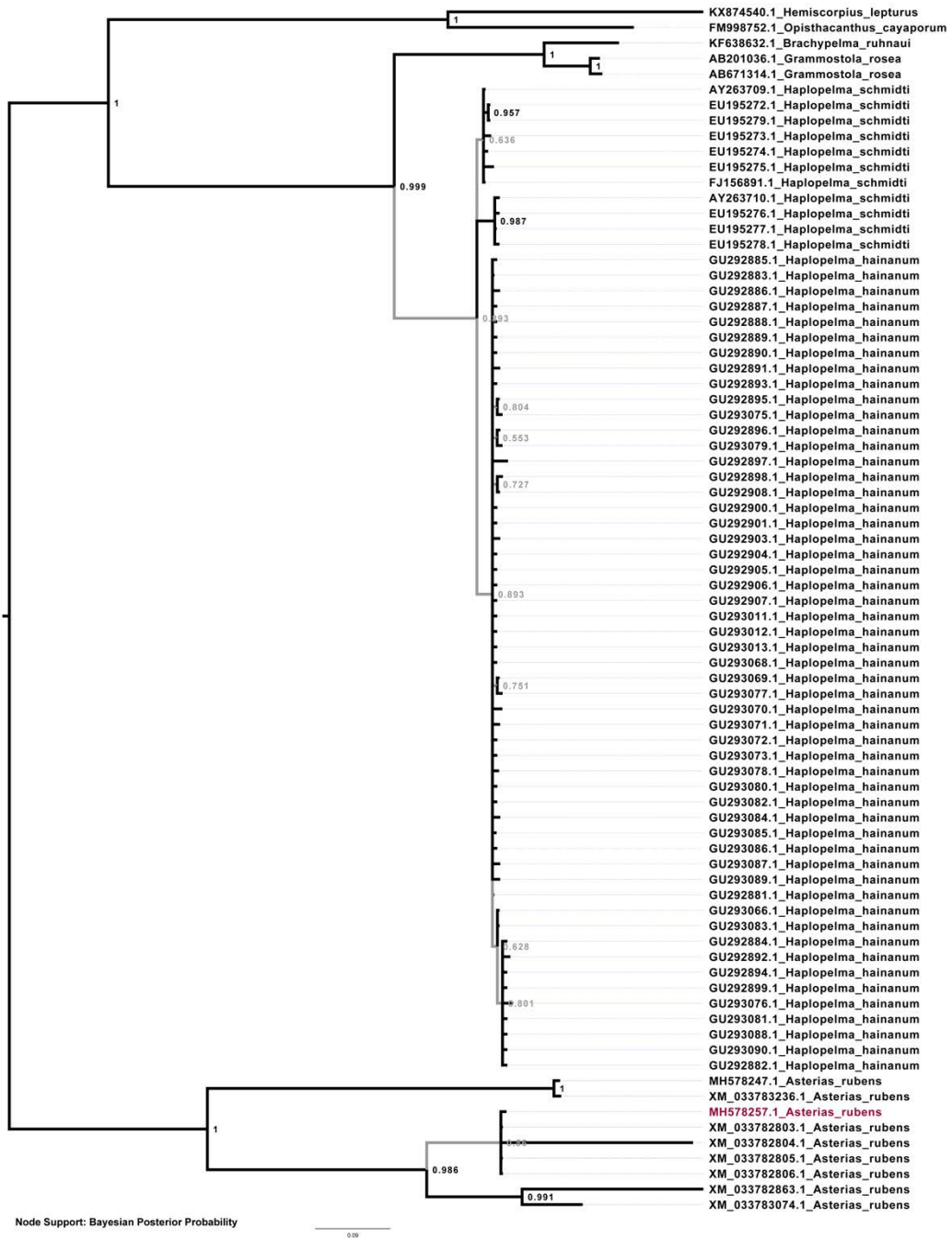
This figure highlights the Bayesian phylogeny of DDH scaffold. Here, the node supports are shown as Bayesian Posterior Probabilities (BPP), where nodes with BPP < 0.9 and those with BPP ≥ 0.90 are shown in light grey and thick black lines, respectively. DDH containing KASH2 from *A. rubens* is shown in red.

Appendix 15. Sequences of selected granulin peptides.

	Peptide	Sequence
12-cysteine	Human grn A	DVK**C**DMEVS**C**PDGYT**CC**RLQSGAWG**CC**PFTQAV**CC**EDHIH**CC**PAGFT**C**DTQKGT-**C**EQ
	Zebrafish grn AaE	DVQ**C**GGGFS**C**HDGET**CC**PTSQTTWG**CC**PSPKAV**CC**DDMQH**CC**PAGYK**C**GPGGT--**C**IS
10-cysteine	Human grn G	GGP**C**QVDAH**C**SAGHS**C**IFTVSGTSS**CC**PFPEAVA**C**GDGHH**CC**PRGFH**C**SADGRS-**C**FQ
	Zebrafish grn AaG	DIS**C**PDGSS**C**PAEFS**C**LLMSTSYG-**CC**PVAQGLA**C**SDGKH**CC**PNDHE**C**SSDSSL-**C**VK

Hyphen (–) indicates a sequence gap. Cysteine residues are shown in bold and red. Human granulin A and zebrafish granulin AaE comprise 12-cysteine residues, whereas human granulin G and zebrafish AaG contain 10-cysteine residues.

Appendix 16. Chapter 3 publication.

This chapter, titled "Folding of truncated granulin peptides", is published and presented on the next page.

Folding of Truncated Granulin Peptides

Rozita Takjoo, David Wilson, Paramjit S. Bansal, Alex Loukas, Michael J. Smout * and **Norelle L. Daly ***

Centre for Molecular Therapeutics, Australian Institute of Tropical Health and Medicine, James Cook University, Cairns, QLD 4870, Australia; rozita.takjoochelaras@my.jcu.edu.au (R.T.); david.wilson4@jcu.edu.au (D.W.); psoondh@hotmail.com (P.S.B.); alex.loukas@jcu.edu.au (A.L.)

*** Correspondence: michael.smout@jcu.edu.au (M.J.S.); norelle.daly@jcu.edu.au (N.L.D)

Abstract: Granulins are a family of unique protein growth factors which are found in a range of species and have several bioactivities that include cell proliferation and wound healing. They typically contain six disulfide bonds, but the sequences, structures and bioactivities vary significantly. We have previously shown that an N-terminally truncated version of a granulin from the human liver fluke, *Opisthorchis viverrini*, can fold independently into a "mini-granulin" structure and has potent wound healing properties in vivo. The incorporation of a non-native third disulfide bond, with respect to the full-length granulin module, was critical for the formation of regular secondary structure in the liver fluke derived peptide. By contrast, this third disulfide bond is not required for a carp granulin-1 truncated peptide to fold independently. This distinction led us to explore granulins from the zebrafish model organism. Here we show that the mini-granulin fold occurs in a naturally occurring paragranulin (half-domain) from zebrafish, and is also present in a truncated form of a full-length zebrafish granulin, suggesting this structure might be a common property in either naturally occurring or engineered N-terminally truncated granulins and the carp granulin-1 folding is an anomaly. The in vitro folding yield is significantly higher in the naturally occurring paragranulin, but only the truncated zebrafish granulin peptide promoted the proliferation of fibroblasts consistent with a growth factor function, and therefore the function of the paragranulin remains unknown. These findings provide insight into the folding and evolution of granulin domains and might be useful in the elucidation of the structural features important for bioactivity to aid the design of more potent and stable analogues for the development of novel wound healing agents.
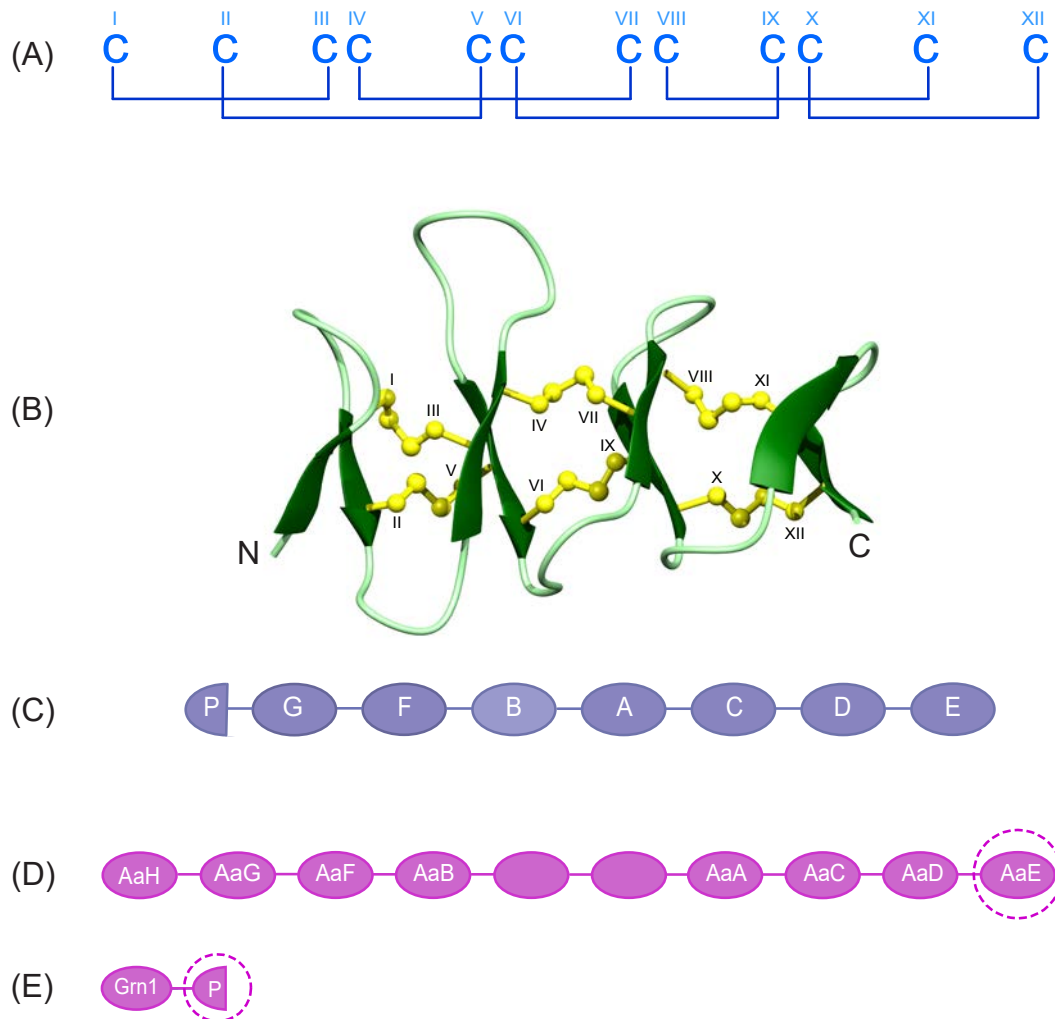
Keywords: granulins; peptide; oxidative folding; NMR spectroscopy; cell proliferation

1. Introduction

Granulin proteins are prevalent throughout nature, being found in a wide variety of organisms [1]. They have potential as a novel class of therapeutic agents for treating a range of diseases and conditions, including chronic wounds [2–11]. Most granulins contain twelve cysteine residues, that form six disulfide bonds (Figure 1A,B), and are often expressed as progranulins containing multiple granulin modules. For example, the human and other mammalian granulin precursors (PGRN) contain a signal peptide and tandem repeats of seven-and-a-half granulin modules, as shown in Figure 1C for the human precursor [1,12]. Granulin G contains only ten cysteine residues and the half granulin module, paragranulin, comprises only the first six cysteine residues [13].

Significant sequence, structural and bioactivity variations exist across the granulin family [1,12]. The sequence conservation is primarily limited to the cysteine framework [12], with significant variation in the inter-cysteine loops. Structural diversity has been observed through the analysis of fish and human derived granulins. In contrast to the carp granulin-1 and zebrafish granulin AaE,

which display relatively well-defined structures containing a four β-hairpin stack [14,15], several of the human granulin peptides display disordered regions. Human granulins A, C and F, for example, have well-defined N-terminal regions, but disordered C-terminal regions when produced recombinantly as single units, but it is unknown if this reflects the native granulin structures in biological systems. The bioactivity of the human granulins also varies, with granulin A shown to inhibit the proliferation of a breast cancer cell line, whereas human granulin F stimulates cell proliferation [16].

Figure 1. Granulin cysteine framework and precursor organization. (**A**) Schematic representation of the highly conserved cysteine framework and disulfide bond pairing present in the granulin family. The cysteine residues are numbered using Roman numerals (I-XII). (**B**) The solution structure of full-length zebrafish granulin AaE (PDB ID: 6cku). Disulfide bonds are shown in yellow and cysteine residues are labelled using Roman numerals. (**C**) The order of the 7.5 granulin motifs in human progranulins, including A to G and a half-granulin motif named paragranulin (P) that includes the N-terminal six cysteines of the full length granulin. (**D**) Schematic representation of zebrafish progranulin A. Granulin modules with no label were missing from the partial clone [15]. (**E**) Granulin modules from zebrafish progranulin 1 including one full-length, and a half-granulin module named paragranulin (P) that comprises six cysteines of the full length granulin [15]. Target peptides of interest in the current study are shown with dashed circles.

Truncated forms of granulins can fold into well-defined structures with only two or three disulfide bonds and, in some cases, have potent bioactivities. Peptides corresponding to the N-terminal region of carp granulin-1, an analogue of the N-terminal region of human granulin A (hGRNA) and a plant

cysteine protease (oryzain β) analogue, can all fold independently into β-sheet structures with only two disulfide bonds [17–20]. However, we have recently shown that analogous peptides from the N-terminus of a granulin derived from the human liver fluke *Opisthorchis viverrini* (*Ov*-GRN-1) do not form a regular secondary structure [2,3]. By contrast, an *Ov*-GRN-1 derived peptide, (*Ov*-GRN$_{12-35\_3s}$), that contains three disulfide bonds, including a disulfide bond (CysIV-CysVI) not present in the full length granulins, can fold independently with a β-sheet structure. Interestingly, both the two and three-disulfide bond containing *Ov*-GRN-1 derived peptides display potent in vivo wound healing properties, similar to the full-length *Ov*-GRN-1 [2,3]. Indeed, an analogue of the *Ov*-GRN-1 peptide with three disulfide bonds has significant potential as a wound healing agent and is as potent as Regranex [2,3], a recombinant human platelet-derived growth factor, but can be made with synthetic methods rather than recombinant technologies. Regranex is currently the only approved biologic on the market for the treatment of chronic wounds [21,22].

The determination of whether granulins with sequences that differ significantly from *Ov*-GRN-1 can also fold independently while incorporating the CysIV-CysVI disulfide bond, is a critical step in advancing the understanding of the folding of this structural motif. Such an advance might have significant implications in the field of disulfide-rich peptide/protein structure and folding pathway predictions using molecular simulations [23]. On a more applied note, further understanding of the structure/function relationships of granulin-derived peptides might be useful in the design of wound healing agents for treating chronic wounds, such as diabetic foot ulcers.

To address the question regarding the folding of the N-terminal regions of granulins, we have analysed the structures of two truncated forms of zebrafish AaE, as well as a paragranulin from zebrafish progranulin 1. The peptides of interest in this study are highlighted with dashed circles in Figure 1D,E. In addition to a well-defined three-dimensional structure, zebrafish granulin AaE [15,24] promotes the survival of neuronal cells [15]. However, the structure or bioactivity of the half-module granulin from zebrafish has not been determined. Indeed, there have been limited studies on the naturally occurring paragranulins (half-modules). We show that incorporating a CysIV-CysVI disulfide bond into the N-terminal region of granulin might be a common feature across the family, but that the proliferative bioactivity of human cell lines is highly dependent on the primary structure.

2. Materials and Methods

2.1. Peptide Synthesis

The synthesis of truncated granulin peptides was undertaken by manual solid-phase peptide synthesis using standard 9-fluorenylmethyloxycarbonyl (Fmoc) chemistry. All peptides were assembled on 2-chlorotrityl chloride resin (Auspep, Tullamarine, Australia). The Fmoc amino acids (Auspep, Tullamarine, Australia) were activated using O-(1H-6-Chlorobenzotriazole-1-yl)-1,1,3,3-tetramethyluronium hexafluorophosphate, HCTU, (Iris Biotech GMBH, Marktredwitz, Germany) and coupled on the resin with DIPEA/DMF. Peptides were cleaved by a cleavage cocktail, including 95% trifluoracetic acid (TFA)/2.5% triisopropylsilane (TIPS)/2.5% H$_2$O (*v/v/v*). TFA was removed with a stream of nitrogen gas. Peptides were then precipitated in ice-cold diethyl ether and dissolved in 50% acetonitrile: 50% H$_2$O:0.1% TFA (*v/v*) and subsequently lyophilised. The peptides were synthesised without isotope labelling and the subsequent NMR spectra were recorded with natural abundance.

2.2. Purification

The crude peptides were purified using reversed-phase high performance liquid chromatography (RP-HPLC) on a C$_{18}$ preparative column (Phenomenex Jupiter 250 × 21.2 mm, 10 μm, 300 Å) with gradients of solvent B (90% acetonitrile/10% H$_2$O/0.045% TFA (*v/v/v*)) and solvent A (H$_2$O/0.05% TFA (*v/v*)). The masses of collected fractions were determined using a 5800 MALDI TOF/TOF mass spectrometer (SCIEX, Foster City, CA, USA).

2.3. Disulfide Bond Formation

The disulfide bonds of the truncated peptides were formed using non-regioselective oxidation in a solution of peptide in 0.1 M ammonium bicarbonate buffer (pH 8–8.2) and 5 mM reduced glutathione at room temperature. The oxidation reaction was left for 48 h prior to acidification with TFA (30–50 μL TFA to a 10 mL oxidation reaction) and loaded on a C_{18} preparative HPLC column with a flow rate of 6 mL/min. Fractions were collected and the peptide mass was analysed using a SCIEX 5800 MALDI TOF/TOF spectrometer [2,3] (Supplementary Figure S1).

2.4. NMR Spectroscopy

NMR samples were prepared from unlabelled purified peptide (0.2 mm) in 90% H_2O/10% D_2O. All NMR spectra were recorded on a 600 MHz AVANCE III NMR spectrometer (Bruker, Karlsruhe, Germany), equipped with a 5 mm TCI cryoprobe. Two-dimensional $^1H–^1H$ TOCSY, $^1H–^1H$ NOESY, $^1H–^1H$ DQF-COSY and $^1H–^{13}C$ HSQC spectra were acquired at 290 K. Spectra were recorded using an interscan delay of 1 s. NOESY spectra were acquired with mixing times of 200 ms, and TOCSY spectra were acquired with isotropic mixing periods of 80 ms. All spectra were processed using Bruker TopSpin (Version 3.5pl7) and assigned using CCPNMR analysis 2.1, based on the approach described in Wüthrich et al. [25,26]. More than 90% of the protons were assigned and 70% of the Cα and Cβ atoms were unambiguously assigned. Amide temperature coefficients were calculated based on TOCSY spectra directly referenced to DSS (4,4-dimethyl-4-silapentane-1-sulfonic acid) and recorded at temperatures ranging from 290 K to 305 K using the method outlined in Cierpicki et al. [27]. Hydrogen bond restraints (Supplementary Table S1) were included in the calculations based on hydrogen bonds identified in preliminary structures and amide protons with temperature coefficients more positive than −4.6 ppb/K.

The αH secondary shifts were determined by subtracting random coil ^1H NMR chemical shifts from the experimental αH chemical shifts [28]. Root mean square deviation (RMSD) values were calculated relative to a mean structure using MOLMOL [29]. The structures and chemical shifts have been deposited into the Protein Data Bank and the Biological Magnetic Resonance Data Bank (ZF-N$_{24\_3}$—PDB ID: 7JIA, BMRB ID: 30780; ZF-para$_{\_3s}$—PDB ID: 7JIY, BMRB ID: 30781).

2.5. Structure Calculations

The three-dimensional structures of peptides were calculated using the CYANA program, based on automated assignment of the NOEs [30]. Torsion-angle restraints predicted by TALOS-N were used in the structure calculations [31]. Structures were visualised using MOLMOL [29].

2.6. Mammalian Cell Culture

The 1BR.3.GN human skin normal fibroblast cell line was obtained from a European Collection of Authenticated Cell Cultures (ECACC, Porton Down, UK). The 1BR.3.GN cells were grown and maintained in Dulbecco's Modified Eagle Medium/Nutrient Mixture F-12 (DMEM/F12) (Life Technologies, Melbourne, Australia) containing 1 × antibiotic/antimycotic and 1 × GlutaMAX, supplemented with 10% foetal bovine serum (FBS) (Gibco, Glasgow, Scotland) at 37 °C and 5% CO_2. Cell proliferation assays were performed with DMEM/F12 media, supplemented with 10% FBS.

2.7. Cell Proliferation Monitoring in Real Time Using xCELLigence

Cells were seeded at 5000 cells/well in 170 μL of complete media in E-plates (ACEA Biosciences, San Diego, CA, United States) and grown overnight and monitored with an xCELLigence SP system (ACEA Biosciences), which monitors cellular events in real time by measuring electrical impedance across gold microelectrodes integrated into the base of tissue culture plates. Cells were washed three times with PBS prior to addition of 150 μL of low nutrient media and incubated for a minimum of 6 h before further treatment. Treatments were prepared at 8.5 × concentration and added to each

well in a total volume of 20 µL. The xCELLigence system recorded cell indexes at intervals of 1 h for 5−6 days following treatment. Readings for the cell index were normalized prior to treatment, and cell proliferation ratios represent the relative numbers of cells compared to control cells at day 4. Comparisons of induction of cell proliferation in response to treatments were accomplished using the Two-Way ANOVA test with Dunnett's multiple comparison correction, using GraphPad Prism 8.0.

3. Results

3.1. Design and Synthesis of Zebrafish Granulin Peptides

To analyse the folding of zebrafish granulins, three peptides were designed: ZF-N$_{24\_2s}$, ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$. The first peptide contains two disulfide bonds, and the latter two peptides contain three disulfide bonds. The sequences and sources of the synthetic peptides are shown in Table 1. An *Ov*-GRN-1 derived peptide is given in the table to highlight the differences in the inter-cysteine loop sequences.

Table 1. Granulin-derived peptide sequences.

Peptide	Source	Sequence
ZF-N$_{24\_3s}$	Zebrafish granulin AaE	CGGGF–SCHDGETCCPTSQTTWGCC
ZF-N$_{24\_2s}$	Zebrafish granulin AaE	CGGGF–SCHDGETCAPTSQTTWGCA
ZF-para$_{\_3s}$	Zebrafish paragranulin	CEGNFY–CPAEKFCCKTRTGQWGCC
Ov-GRN$_{12-35\_3s}$[#]	*Ov*-GRN-1	CPDPVYTCRPGQTCCRGLHG–YGCC

Hyphen (–) indicates a sequence gap. Cysteine residues are shown in bold and red. [#] The sequence of the previously studied *Ov*-GRN$_{12-35\_3s}$ [2], is cited accordingly and presented in this table for comparison purposes. The sequences of the peptides synthesised in the current study are highlighted in blue. The cysteine residues involved in the non-native (CysIV-CysVI) disulfide bond in zebrafish granulin AaE, and by homology in *Ov*-GRN-1 (the 3D structure of *Ov*-GRN-1 has not been determined), have been underlined. The sequences are derived from the precursor proteins of zebrafish granulin A (Q90ZD0), zebrafish progranulin-1 (Q8QGN9) and *Ov*-GRN-1 (B8XSI4).

Peptides were synthesised using Fmoc chemistry on 2-chlorotrityl resin without selective protection of the cysteine residues, and the disulfide bonds formed by air oxidation in ammonium bicarbonate (0.1 M) and reduced glutathione (5 mM) at room temperature. The analytical RP-HPLC trace of all three oxidation reactions contain several peaks, including a sharp peak that eluted earlier than the other peaks and corresponds to a fully oxidised isomer, based on mass spectrometry. The early eluting peaks were purified (>95% purity based on analytical RP-HPLC) and the structures analysed using NMR spectroscopy. The yields of the early eluting peaks varied amongst the peptides. A comparison of the folding of the peptides is shown in Figure 2, highlighting the higher yield obtained for the naturally occurring paragranulin compared to the truncated version of zebrafish granulin AaE (ZF-N$_{24\_3s}$).

### 3.2. Structural Analysis with NMR Spectroscopy

The one-dimensional proton NMR spectra of ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ display large chemical shift dispersion in the amide region, and analysis of the secondary shifts indicates that the peptides contain β-sheet structure based on the consecutive positive shifts (Figure 3). The three-dimensional structures of both peptides were calculated using CYANA based on distance and dihedral angle restraints derived from one- and two-dimensional homonuclear and heteronuclear NMR experiments. Calculation of the structures without disulfide bond restraints indicated that the most likely connectivity was CysI-CysIII, CysII-CysV and CysIV-CysVI, consistent with the previous study on *Ov*-GRN-1 peptides [2,3]. Calculation of the structures with all 15 possible disulfide bond connectivities also indicated that this connectivity had the lowest target function (Supplementary Table S2). Subsequent structures were calculated using restraints for this disulfide connectivity. Hydrogen bond restraints were included based on analysis of the preliminary structures and temperature coefficients (Supplementary Tables S1, S3 and S4). The refinement statistics are given in Table 2.
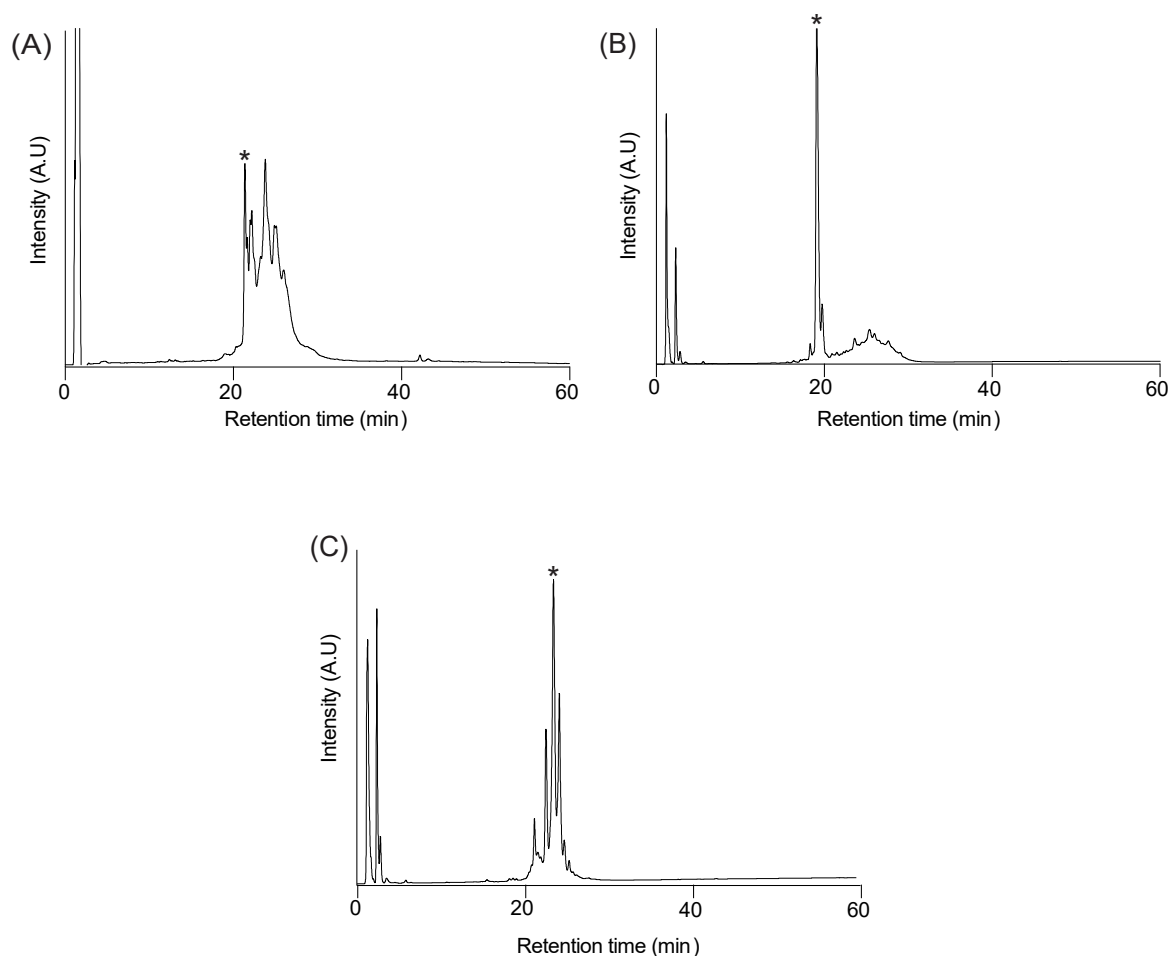
Figure 2. RP-HPLC analysis of the oxidation reaction of granulin N-terminal truncated analogues (**A**) ZF-N$_{24\_3s}$, (**B**) ZF-para$_{\_3s}$ and (**C**) ZF-N$_{24\_2s}$. Analytical RP-HPLC was carried out using a Phenomenex Jupiter 4 μm Proteo column (C$_{12}$, 150 × 2.00 mm, 10 μm, 90 Å) using a gradient of 0–60% solvent B (Solvent A: 99.95% H$_2$O: 0.05% TFA; Solvent B: 90% acetonitrile: 10% H$_2$O: 0.045% TFA) over 60 min with a flow rate of 0.4 mL/min. The absorbance was monitored at 214 nm. The early eluting sharp peaks (highlighted by asterisks (*)) were purified for subsequent analyses.

In contrast to the three-disulfide bond containing peptides, ZF-N$_{24\_2s}$ did not display significant dispersion in the amide region of the one-dimensional proton spectrum. Despite this lack of dispersion in the amide region, the backbone protons and most of the side chain protons could be assigned based on two-dimensional spectra. Analysis of the secondary shifts (Figure 3) indicates that the peptide does not contain regular secondary structure, as the majority of the shifts are within 0.1 ppm of the random coil values.

The major element of secondary structure in ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ is a β-hairpin comprising residues 13–23. An NOE between the α-protons of Cys 13 and Cys 23 is clearly present in both spectra, and represents a key restraint in defining the alignment of the β-strands. The structures appear to be stabilised by hydrogen bonds consistent with amide protons, having temperature coefficients more positive than −4.6 ppb/K [27]. The disulfide bonds have the connectivity CysI-CysIII, CysII-CysV and CysIV-CysVI; the two former disulfide bonds are equivalent to those present in the full-length structures of the well characterised granulins [14,18]. In the full-length structures CysIV is bonded to CysVII and CysVI is bonded to CysIX (refer to Figure 1A for disulfide bond connectivity) [15,16,32].
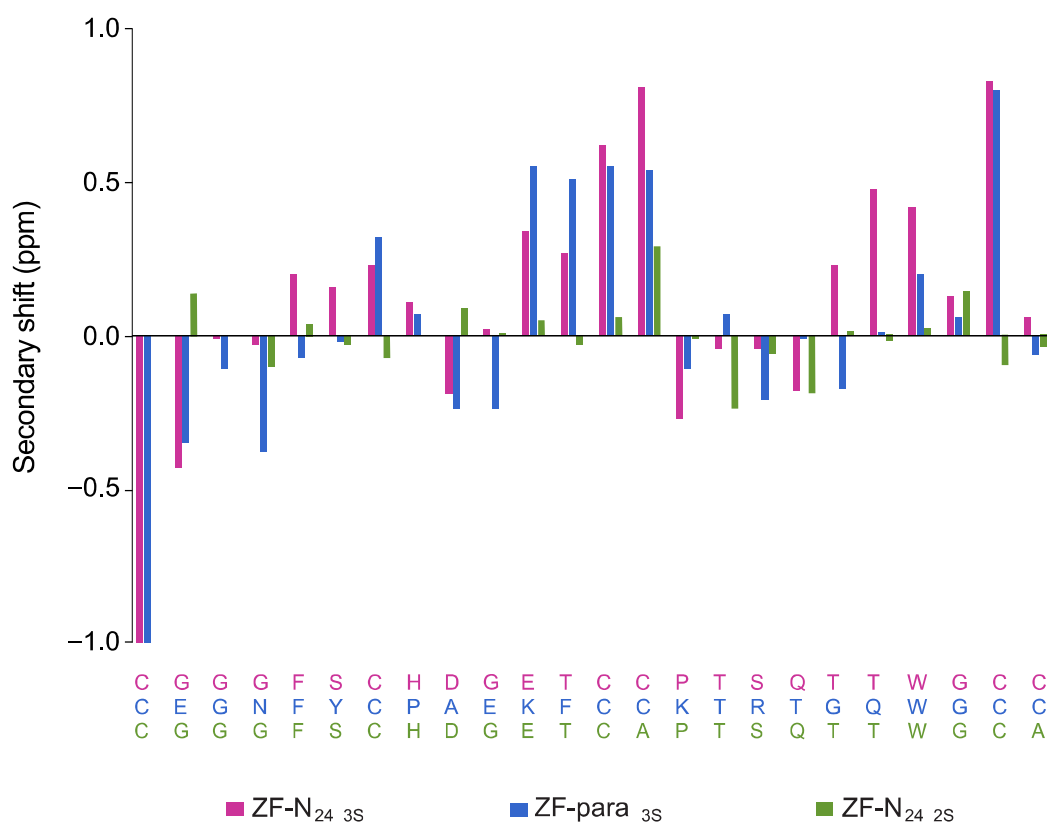
Figure 3. αH secondary-shift comparison for truncated granulin analogues. The αH secondary shifts were calculated by subtracting the random coil ^{1}H NMR chemical shifts previously reported by Wishart et al. [28] from the experimental αH chemical shifts. The sequences of ZF-N$_{24\_3s}$ (violet), ZF-para$_{\_3s}$ (blue) and ZF-N$_{24\_2s}$ (green) are given at the bottom of the diagram and are colour coded with respect to the graph.

Table 2. Structural statistics.

| Experimental Restraints	ZF-N$_{24\_3s}$	ZF-para$_{\_3s}$
Interproton distance restraints		
Intra-residue, |i − j| = 0	41	72
Sequential, |i − j| = 1	65	88
Medium range, 1 < |i − j| < 5	12	36
Long range, |i − j| > = 5	31	64
Disulfide-bond restraints (3 restraints per bond)	9	9
Dihedral-angle restraints	27	32
Hydrogen bond restraints (2 restraints per bond)	4	8
Root Mean Square Deviations from Mean Coordinate Structure (Å)		
Backbone atoms	1.29 ± 0.38	0.63 ± 0.20
All heavy atoms	1.81 ± 0.41	1.31 ± 0.35
Ramachandran Statistics		
% in most favoured region	80.6	96.4
% Residues in additionally allowed regions	19.4	3.6

Comparison of the structures of ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ with the structure of a truncated form of *Ov*-GRN-1 (*Ov*-GRN$_{12-35\_3s}$) [2,3], also containing the first six cysteine residues, indicates the peptides have similar overall folds despite the sequence differences (37.5% sequence identity to *Ov*-GRN$_{12-35\_3s}$; the sequence identity is primarily related to the conserved cysteine residues) (Figure 4). A superposition

of ZF-N$_{24\_3s}$ with the structure of zebrafish granulin AaE is given in Figure 5 and highlights the similarity between the truncated and full-length versions, despite the differences in disulfide bonds.
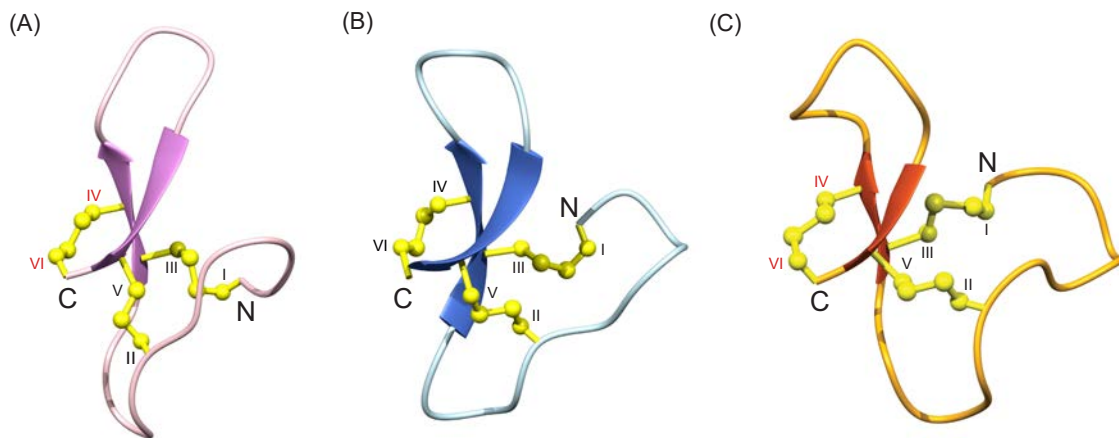

Figure 4. Three-dimensional structures of (**A**) ZF-N$_{24\_3s}$, (**B**) ZF-para$_{\_3s}$ and (**C**) *Ov*-GRN$_{12-35\_3s}$ [2] N-terminal truncated peptides. The β-hairpins are shown as arrows, and ZF-N$_{24\_3s}$ is shown in violet, ZF-para$_{\_3s}$ in blue and *Ov*-GRN$_{12-35\_3s}$ (PDB: 5UJG) in orange. Disulfide bonds are shown in yellow and cysteine residues are labelled using Roman numerals. The cysteine residues involved in forming the disulfide bond (IV-VI) not present in the full-length granulin structures of zebrafish granulin AaE and by homology *Ov*-GRN-1, are labelled with red letters (**A**,**C**). The figure was prepared using MOLMOL [29].
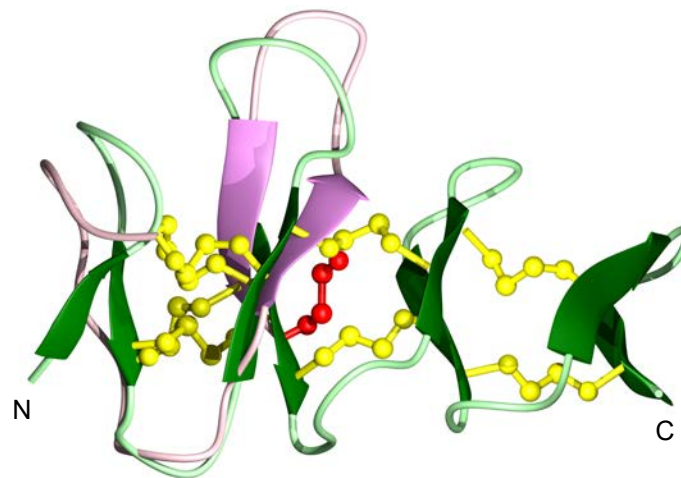
Figure 5. Superposition of the structures of the ZF-N$_{24-3s}$ and zebrafish AaE. The β-hairpins are shown as arrows—in violet for ZF-N$_{24-3s}$ (PDB ID: 7JIA), and dark green for zebrafish AaE (PDB ID: 6cku). Disulfide bonds are shown in yellow. The cysteine residues involved in forming the non-native disulfide bond (IV-VI) in ZF-N$_{24-3s}$ is shown in red. The figure was prepared using MOLMOL [29].

3.3. Cell Proliferation Assay

The effect of the truncated peptides on the growth of 1BR.3.GN fibroblasts was evaluated using an xCELLigence system. The 1BR.3.GN is a human fibroblast skin cell line derived from transformed normal fibroblasts. Cells were cultured with two concentrations of the peptides, 200 nM and 1 μM. ZF-N$_{24\_3s}$ significantly promoted the proliferation of fibroblasts compared to the negative control peptide (a 19-residue peptide from tropomyosin) (Figure 6). The proliferation rate for ZF-N$_{24\_3}$ was 113.6% ($p < 0.05$), relative to the control peptide at day 4 after treatment. By contrast, the effect of ZF-para$_{\_3s}$ or ZF-N$_{24\_2S}$ on fibroblast cell proliferation was not statistically significant at either 200 nM or 1 μM (Figure 6).
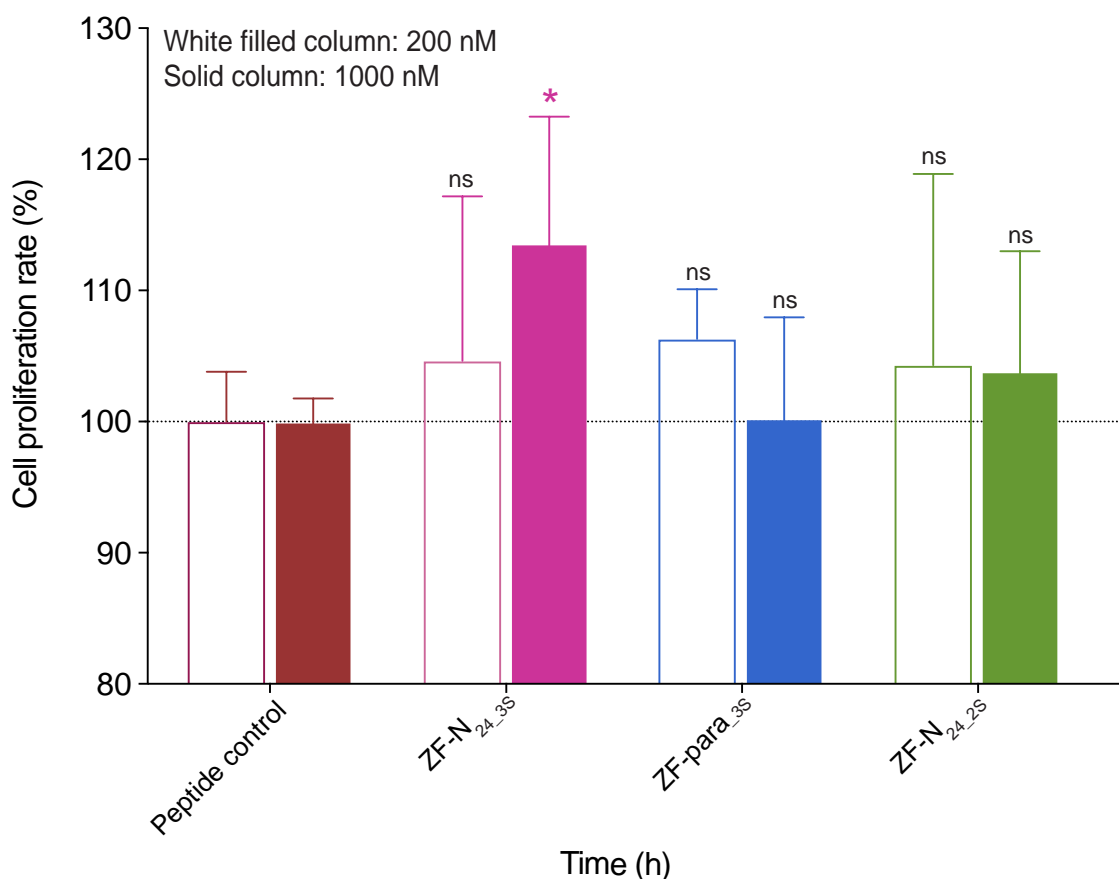
Figure 6. Cell proliferation analysis of granulin peptides using xCELLigence technology. Fibroblast cells were treated with peptides at concentrations of 200 nM and 1 μM. Data represent mean ± SD of three independent experiments. Data were analysed by one-way ANOVA against peptide control: ns—not significant; *—$p < 0.05$. Cell index was measured at day 4 after treatment. White-filled columns and solid columns represent concentrations of 200 nM and 1 μM, respectively.

4. Discussion

Folding of disulfide rich peptides still represents one of the great challenges in structure prediction and simulation [33,34]. Experimental analysis of the folding processes has to underlie these simulations, and analysis of the granulin structural framework present in the vast majority of organisms is likely to provide valuable insight for future computational studies. The potential as a wound healing agent might also be enhanced as a result of greater understanding of the structure/function relationships. In the current study we show that the N-terminal half of a range of granulins can fold independently of the C-terminal region via the incorporation of a CysIV-CysVI disulfide bond, indicating this is a common feature in this family. It appears likely that the naturally occurring half-motifs are optimised for folding compared to engineered versions derived from full-length granulin sequences, and that the CysIV-CysVI disulfide bond likely constitutes a native disulfide bond. However, the structure of a paragranulin on native material will be required to confirm this suggestion.

The secondary shifts of the granulin peptides ZF-N$_{24\_3s}$ and ZF-para$_{\_3s}$ were similar (Figure 3), indicating they have similar structures, and this was subsequently confirmed by the determination of the three-dimensional structures of the peptides (Figure 4). Both peptides contain a β-hairpin and similar arrangement of the disulfide bonds to *Ov*-GRN$_{12-35\_3s}$ [2,3], the first granulin peptide shown to adopt this conformation with three disulfide bonds. The common fold we have identified for the half-granulin domain could represent an evolutionary ancestor of the full-length scaffold, hence the propensity for a range of granulins to adopt this fold in the truncated form even though the disulfide connectivity differs in the full-length proteins.

The more efficient in vitro folding of the paragranulin-derived peptide, ZF-para$_{3s}$, relative to the engineered version (ZF-N$_{24\_3S}$), poses questions regarding the sequence differences between the peptides and the subsequent folding pathways. β-hairpins can represent the folding nucleus with hydrophobic collapse being an integral part of folding [35]. However, comparison of the sequences of the two zebrafish peptides indicates very little difference in the hydrophobicity of the amino acids involved in the β-hairpin loop and therefore it is difficult to speculate the reason for the folding differences.

The structural fold present in these truncated granulin peptides has recently been referred to as the "mini-granulin" fold and is present in several peptides [36], including conotoxins φ-MiXXVIIA [37] and H-Vc7.2. The fold contains two conserved disulfide bonds, which correspond to CysI-CysIII and CysII-CysV in ZF-N$_{24\_3s}$ and ZF-para$_{3s}$. Both φ-MiXXVIIA and H-Vc7.2 contain additional disulfide bonds, but they are distinct from the additional CysIV-CysVI bond present in ZF-N$_{24\_3s}$ and ZF-para$_{3s.}$ However, there are examples of this motif containing only two disulfide bonds, such as the truncated form of carp granulin-1, which forms a well-defined β-sheet-containing structure [16,17], indicating that the mini-granulin fold is not always dependent on additional disulfide bonds.

The β-sheet structure present in the carp granulin-1 truncated peptide is in contrast to the results obtained in the current study on ZF-N$_{24\_2s}$, which has no regular secondary structure. Previous studies have shown that, when only CysI-CysIII and CysII-CysV are present in a truncated form of human granulin A, the peptide does not form a well-defined structure [17]. Similarly, a truncated form of *Ov*-GRN-1 with only two disulfide bonds does not form β-sheet structure, whereas the three-disulfide bond form does [2,3]. The results to date suggest that the folding of the carp granulin-1 peptide could be an anomaly, but further study is required to determine residues required for effective folding.

In addition to providing insight into granulin folding, the current study also analysed the biological activity of the truncated N-terminal peptides. ZF-N$_{24\_3s}$ displayed activity in the cell proliferation assay, whereas ZF-para$_{3s}$ was not active. However, ZF-N$_{24\_3s}$ was not as active as the parasite-derived granulin peptide *Ov*-GRN$_{12-35\_3s}$ [2]. The cells chosen for study in the current study reflect the potential of granulin peptides as wound healing agents, but it is possible the zebrafish derived peptides would show greater potency on fish-derived cell lines. There are now several examples where it has been shown that well-defined 3D structures are not necessary for bioactivity [38], including human granulin B, which has been characterised as an intrinsically disordered protein that modestly induces NF-κB activation in SY-SH5Y human neuroblastoma cells [39]. The current study highlights the importance of the primary structure, given that all peptides contain a β-hairpin, but the activity varies significantly.

5. Conclusions

Overall, this study has shown that truncated granulins can incorporate a disulfide bond not-present in the full-length module, and this bond plays a pivotal role in stabilizing a well-defined structure in the N-terminal region of truncated granulin peptides. Despite a common structural motif, the bioactivity of these engineered peptides varies, and further mutational studies will be required to determine important regions/residues for bioactivity. Ultimately, the identification of the molecular target is likely to provide the greatest advance in understanding the structure/function relationships of granulin peptides.

References

1. Palfree, R.G.; Bennett, H.P.; Bateman, A. The evolution of the secreted regulatory protein progranulin. *PLoS ONE* **2015**, *10*, e0133749. [CrossRef] [PubMed]

2. Bansal, P.S.; Smout, M.J.; Wilson, D.; Caceres, C.C.; Dastpeyman, M.; Sotillo, J.; Seifert, J.; Brindley, P.J.; Loukas, A.; Daly, N.L. Development of a potent wound healing agent based on the liver fluke granulin structural fold. *J. Med. Chem.* **2017**, *60*, 4258–4266. [CrossRef] [PubMed]

3. Dastpeyman, M.; Bansal, P.S.; Wilson, D.; Sotillo, J.; Brindley, P.J.; Loukas, A.; Smout, M.J.; Daly, N.L. Structural variants of a liver fluke derived granulin peptide potently stimulate wound healing. *J. Med. Chem.* **2018**, *61*, 8746–8753. [CrossRef] [PubMed]

4. Smout, M.J.; Sotillo, J.; Laha, T.; Papatpremsiri, A.; Rinaldi, G.; Pimenta, R.N.; Chan, L.Y.; Johnson, M.S.; Turnbull, L.; Whitchurch, C.B.; et al. Carcinogenic parasite secretes growth factor that accelerates wound healing and potentially promotes neoplasia. *PLoS Pathog.* **2015**, *11*, e1005209. [CrossRef]

5. Botelho, M.C.; Alves, H.; Richter, J. Wound healing and cancer progression in Opisthorchis viverrini associated cholangiocarcinoma. *Parasitol. Res.* **2016**, *115*, 2913–2914. [CrossRef]

6. Ding, H.; Wei, J.; Zhao, Y.; Liu, Y.; Liu, L.; Cheng, L. Progranulin derived engineered protein Atsttrin suppresses TNF-α-mediated inflammation in intervertebral disc degenerative disease. *Oncotarget* **2017**, *8*. [CrossRef]

7. Qiao, G.; Xu, H.L.; Li, C.; Li, X.; Farooqi, A.A.; Zhao, Y.M.; Liu, X.H.; Liu, M.; Stagos, D.; Lin, X.K. Granulin A synergizes with cisplatin to inhibit the growth of human hepatocellular carcinoma. *Int. J. Mol. Sci.* **2018**, *19*, 3060. [CrossRef]

8. Chitramuthu, B.P.; Bennett, H.P.J.; Bateman, A. Progranulin: A new avenue towards the understanding and treatment of neurodegenerative disease. *Brain* **2017**, *140*, 3081–3104. [CrossRef]

9. Jian, J.; Tian, Q.-Y.; Hettinghouse, A.; Zhao, S.; Liu, H.; Liu, C.-J.; Wei, J.; Grunig, G.; Zhang, W.; Setchell, K.D.R.; et al. Progranulin recruits HSP70 to β-glucocerebrosidase and is therapeutic against Gaucher disease. *EBioMedicine* **2016**, *13*, 212–224. [CrossRef]

10. Abella, V.; Pino, J.; Scotece, M.; Conde, J.; Lagoa, F.; Gonzalez-Gay, M.A.; Mera, A.; Gomez, R.; Mobasheri, A.; Gualillol, O. Progranulin as a biomaker and potential therapeutic agent. *Drug Discov. Today* **2017**, *22*, 1557–1564. [CrossRef]

11. Pogonowska, M.; Poniatowski, Ł.A.; Wawrzyniak, A.; Królikowska, K.; Kalicki, B. The role of progranulin (PGRN) in the modulation of anti-inflammatory response in asthma. *Cent. Eur. J. Immunol.* **2019**, *44*, 91–101. [CrossRef] [PubMed]

12. Dastpeyman, M.; Smout, M.J.; Wilson, D.; Loukas, A.; Daly, N.L. Folding of granulin domains. *Pept. Sci.* **2018**, *110*. [CrossRef]

13. Ong, C.H.P.; Bateman, A. Progranulin (granulin-epithelin precursor, PC-cell derived growth factor, acrogranin) in proliferation and tumorigenesis. *Histol. Histopathol.* **2003**, *18*, 1275–1288. [PubMed]

14. Hrabal, R.; Chen, Z.; James, S.; Bennett, H.P.J.; Ni, F. The hairpin stack fold, a novel protein architecture for a new family of protein growth factors. *Nat. Struct. Biol.* **1996**, *3*, 747–751. [CrossRef] [PubMed]

15. Wang, P.; Chitramuthu, B.; Bateman, A.; Bennett, H.P.J.; Xu, P.; Ni, F. Structure dissection of zebrafish progranulins identifies a well-folded granulin/epithelin module protein with pro-cell survival activities: Dynamic structures of zebrafish granulin modules. *Protein Sci.* **2018**, *27*, 1476–1490. [CrossRef]

16. Tolkatchev, D.; Malik, S.; Vinogradova, A.; Wang, P.; Chen, Z.; Xu, P.; Bennett, H.P.J.; Bateman, A.; Ni, F. Structure dissection of human progranulin identifies well-folded granulin/epithelin modules with unique functional activities. *Protein Sci.* **2008**, *17*, 711–724. [CrossRef]

17. Tolkatchev, D.; Ng, A.; Vranken, W.; Ni, F. Design and solution structure of a well-folded stack of two β-hairpins based on the amino-terminal fragment of human granulin A. *Biochemistry* **2000**, *39*, 2878–2886. [CrossRef]

18. Vranken, W.F.; Chen, Z.G.; Xu, P.; James, S.; Bennett, H.P.J.; Ni, F. A 30-residue fragment of the carp granulin-1 protein folds into a stack of two β-hairpins similar to that found in the native protein. *J. Pept. Res.* **1999**, *53*, 590–597. [CrossRef]

19. Vranken, W.F.; James, S.; Bennett, H.P.J.; Ni, F. Solution structures of a 30-residue amino-terminal domain of the carp granulin-1 protein and its amino-terminally truncated 3-30 subfragment: Implications for the conformational stability of the stack of two β-hairpins. *Proteins* **2002**, *47*, 14–24. [CrossRef]

20. Tolkatchev, D.; Xu, P.; Ni, F. A peptide derived from the C-terminal part of a plant cysteine protease folds into a stack of two β-hairpins, a scaffold present in the emerging family of granulin-like growth factors. *J. Pept. Res.* **2001**, *57*, 227–233. [CrossRef]

21. Niezgoda, J.A.; Van Gils, C.C.; Frykberg, R.G.; Hodde, J.P. Randomized clinical trial comparing OASIS Wound Matrix to Regranex Gel for diabetic ulcers. *Adv. Skin Wound Care* **2005**, *18*, 258–266. [CrossRef] [PubMed]

22. Chan, R.K.; Liu, P.H.; Pietramaggiori, G.; Ibrahim, S.I.; Hechtman, H.B.; Orgill, D.P. Effect of recombinant platelet-derived growth factor (Regranex®) on wound closure in genetically diabetic mice. *J. Burn Care Res.* **2006**, *27*, 202–205. [CrossRef] [PubMed]

23. Georgoulia, P.S.; Glykos, N.M. Molecular simulation of peptides coming of age: Accurate prediction of folding, dynamics and structures. *Arch. Biochem. Biophys.* **2019**, *664*, 76–88. [CrossRef] [PubMed]

24. Wang, P.F. *Structure Genomics of Zebrafish Granulins*; McGill University: Montreal, QC, Canada, 2004.

25. Wüthrich, K. NMR studies of structure and function of biological macromolecules (Nobel Lecture). *J. Biomol. NMR* **2003**, *27*, 13–39. [CrossRef]

26. Vranken, W.F.; Boucher, W.; Stevens, T.J.; Fogh, R.H.; Pajon, A.; Llinas, M.; Ulrich, E.L.; Markley, J.L.; Ionides, J.; Laue, E.D. The CCPN data model for NMR spectroscopy: Development of a software pipeline. *Proteins* **2005**, *59*, 687–696. [CrossRef] [PubMed]

27. Cierpicki, T.; Otlewski, J. Amide proton temperature coefficients as hydrogen bond indicators in proteins. *J. Biomol. NMR* **2001**, *21*, 249–261. [CrossRef]

28. Wishart, D.S.; Bigam, C.G.; Holm, A.; Hodges, R.S.; Sykes, B.D. ^{1}H, ^{13}C and ^{15}N random coil NMR chemical shifts of the common amino acids. I. Investigations of nearest-neighbor effects. *J. Biomol. NMR* **1995**, *5*, 67–81. [CrossRef] [PubMed]

29. Koradi, R.; Billeter, M.; Wüthrich, K. MOLMOL: A program for display and analysis of macromolecular structures. *J. Mol. Graph.* **1996**, *14*, 51–55. [CrossRef]

30. Güntert, P. Automated NMR structure calculation with CYANA. *Methods Mol. Biol.* **2004**, *278*, 353–378.

31. Shen, Y.; Bax, A. Protein structural information derived from NMR chemical shift with the neural network program TALOS-N. *Methods Mol. Biol.* **2015**, *1260*, 17–32. [CrossRef] [PubMed]

32. Bhandari, V.; Roger, G.E.P.; Bateman, A. Isolation and sequence of the granulin precursor cDNA from human bone marrow reveals tandem cysteine-rich granulin domains. *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1715–1719. [CrossRef]

33. Yang, J.; He, B.-J.; Jang, R.; Zhang, Y.; Shen, H.-B. Accurate disulfide-bonding network predictions improve ab initio structure prediction of cysteine-rich proteins. *Bioinformatics* **2015**, *31*, 3773–3781. [CrossRef] [PubMed]

34. Paul George, A.A.; Heimer, P.; Maaß, A.; Hamaekers, J.; Hofmann-Apitius, M.; Biswas, A.; Imhof, D. Insights into the folding of disulfide-rich μ-conotoxins. *ACS Omega* **2018**, *3*, 12330–12340. [CrossRef] [PubMed]

35. Lewandowska, A.; Ołdziej, S.; Liwo, A.; Scheraga, H.A. β-hairpin-forming peptides; models of early stages of protein folding. *Biophys. Chem.* **2010**, *151*, 1–9. [CrossRef] [PubMed]

36. Nielsen, L.D.; Foged, M.M.; Albert, A.; Bertelsen, A.B.; Søltoft, C.L.; Robinson, S.D.; Petersen, S.V.; Purcell, A.W.; Olivera, B.M.; Norton, R.S.; et al. The three-dimensional structure of an H-superfamily conotoxin reveals a granulin fold arising from a common ICK cysteine framework. *J. Biol. Chem.* **2019**, *294*, 8745–8759. [CrossRef] [PubMed]

37. Jin, A.H.; Dekan, Z.; Smout, M.J.; Wilson, D.; Dutertre, S.; Vetter, I.; Lewis, R.J.; Loukas, A.; Daly, N.L.; Alewood, P.F. Conotoxin φ-MiXXVIIA from the superfamily G2 employs a novel cysteine framework that mimics granulin and displays anti-apoptotic activity. *Angew. Chem. Int. Ed.* **2017**, *56*, 14973–14976. [CrossRef]

38. Uversky, V.N. A decade and a half of protein intrinsic disorder: Biology still waits for physics. *Protein Sci.* **2013**, *22*, 693–724. [CrossRef]

39. Ghag, G.; Wolf, L.M.; Reed, R.G.; Van Der Munnik, N.P.; Mundoma, C.; Moss, M.A.; Rangachari, V. Fully reduced granulin-B is intrinsically disordered and displays concentration-dependent dynamics. *Protein Eng. Des. Sel.* **2016**, *29*, 177–186. [CrossRef]