UC Santa Barbara
UC Santa Barbara Electronic Theses and Dissertations

Title
Fish Scale Collagen Liquid-Liquid Phase Separation

Permalink
https://escholarship.org/uc/item/222527rc

Author
Ignatenko, Dariya

Publication Date
2021

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

Fish Scale Collagen Liquid-Liquid Phase Separation

A thesis submitted in partial satisfaction of the

requirements for the degree Master of Arts

in Molecular, Cellular, and Developmental Biology

by

Dariya A. Ignatenko

Committee in charge:

Professor J. Herbert Waite, Chair

Professor Daniel E. Morse

Professor Leslie Wilson

September 2021

The thesis of Dariya A. Ignatenko is approved.

_____

Dr. Daniel E. Morse

_____

Dr. Leslie Wilson

_____

Dr. J. Herbert Waite, Committee Chair

September 2021

ACKNOWLEDGEMENTS

I thank my advisor Dr. J. Herbert Waite, who ultimately created an environment that cultivates and inspires curiosity and appreciation for the natural world. In a field where the study of biological structures often strays from the organism that makes them, working in Herb's lab has countlessly shown me that pursuing scientific research as it relates to nature unveils countless wonders and possibilities. My time in the graduate program would not have been the same without this lesson and also the many wonderful stories and discussions Herb offers.

I express my gratitude to members of the lab, who provided priceless feedback and support in group meetings, chats at the lab bench, and lunch conversations. I especially thank Dr. William Wonderly. Billy's endless curiosity and generosity made overcoming some of the biggest obstacles as simple as stepping over a pebble. As said by others before me, the success of this thesis would not have been possible without his help.

I would be remiss not to acknowledge my amazing friends for providing me with endless joyful memories and giving me the strength to finish this thesis. Thank you, Alexian Alton, Morgan Mutch, Olecya Tyaglo, Anjali Rai, Bianca Bacaltos, and Laura Busby.

Furthermost, I thank my parents, Tatiana and Andrey, for their unwavering love and support. Your encouraging affirmation and eagerness to learn about my interests have always given me courage in my pursuits. Lastly, I want to thank my brother, Ivan, for always finding new ways to make me laugh, especially during the past tumultuous year, and for being someone I can always relate to.

ABSTRACT

Fish Scale Collagen Liquid-Liquid Phase Separation

by

Dariya A. Ignatenko

Type I collagen and gelatin have previously been shown to form complex and simple coacervates. We have consistently observed purified fish scale collagens (FSc) from several species to liquid-liquid phase separate (LLPS) and form coacervate-like droplets upon heating from 4˚C to room temperature. The observation of this phenomena led us to investigate the composition of fish scale collagen extracts. Collagen extract was purified from scales of several fish species and submitted to gel electrophoresis, amino acid analysis, liquid chromatography-mass spectrometry (LC-MS/MS), and circular dichroism (CD) for compositional analysis. We found that the purified extracts from each fish species contained proteins consistent with collagens in structure and composition, differing only slightly from mammalian collagens as well as other marine type I collagens. Typically, soluble marine type I collagens have lower thermal stability than the corresponding collagens from found in warm-blooded mammalian relatives. Optical rotation of FSc collagen molecules at increasing temperatures indicated that the collagen triple helix is denatured at about the same temperature at which coacervation takes place in concentrated FSc extract solutions under optical and confocal microscopy.

We started a preliminary sequence analysis of FSc collagen from *Salmo salar* to provide a potential explanation for the observed LLPS behavior. We found that salmon collagen had a striking increase in GG repeats and a compositional preference for hydroxylated residues, serine and threonine, compared with bovine collagen. These substitutions replaced primarily proline, hydroxyproline and alanine residues in the X and Y positions of the bovine collagen sequence. We propose that compositional biases for these residues contribute to the predilection of forming the two-phase state, at least in part, by increasing collagen flexibility and stabilizing water-evicting interactions between gelatin molecules.

TABLE OF CONTENTS

LIST OF FIGURES

Introduction

Type I collagen

Type I collagen is one of the largest and most abundant proteins in the animal body. This fibrillar collagen is found in many extracellular materials such as tendons, bones, skin, and fish scales. These materials depend on this structural protein to maintain important functions such as elasticity, strength, and durability. In a historical context, the elasticity and durability of bovine tendon were critical factors in the impressive quality of composite bows as used by Scythian horse archers [1]. In biological structured materials, collagen is thought to be processed as a liquid crystal prior to forming structural assemblies[2] and exhibits birefringent properties *in vitro* at high concentrations.

The tropocollagen molecule is made up of three monomer polypeptides twisted into left-handed polyproline type II helices that wrap around one another to form right-handed triple superhelix. This structured tropocollagen molecule is also termed native collagen. When exposed to sufficient heat, the collagen helix denatures into three monomeric components (some of these are dimer or trimeric due to covalent interchain crosslinks), also referred to as gelatin or gelatin monomers. Soluble gelatins take on intrinsically disordered states, such as a random coil, and lose both the persistence and anisotropy of the triple-helical native collagen. Upon cooling, gelatin molecules come together to form solid-like gels, which are celebrated for their variety of applications in industrial and medical materials, pharmaceuticals, food, and alcoholic beverage clarification[3].

The properties of collagen are largely derived from its monomer sequence, which is characterized by Glycine-X-Y (X = pro, Y = trans-4-hydroxyproline) repeats, referred to as GXY. Although type I collagen is considered a low complexity protein, its large size (≈ 300

1

kDa) creates challenges to comprehensively probing the effect of sequence on collagen structure, function, and phase behavior in a biologically relevant manner. The difficulty in studying this molecule is driven by several factors such as the current challenges in synthesizing large mimetic collagen peptides [4], designing comprehensive *in silico* simulations to predict a variety of behaviors [5], and limited options for expressing large fully modified collagen molecules in recombinant systems [6]. Native collagens sourced from bovine and rat tissue are a common choice for physiological and cellular studies as they have thermal stabilities akin to ours [7]. However, due to the risk of severe disease transfer [8] such as bovine spongiform encephalopathy and high cost of quality material, fish scales have gained attention as a low-cost collagen source [9] as fish collagen is not known to transfer disease and scales are a common waste by-product of fisheries [9]. Additionally, scales are abundant in collagen: collagen comprises 50 % to 70 % of the scale, with the remaining material being primarily inorganic hydroxyapatite. We find FSc type I collagen an interesting model for studying collagen stability and dynamics in a biologically relevant context.

<u>Protein Liquid-liquid Phase Separation or Coacervation</u>

A cohort of polymer solutions undergoes liquid-liquid demixing under a set of conditions where polymers interact to form an immiscible polymer-rich fluid phase within the polymer-poor solvent phase, referred to as coacervation[10]. Protein coacervates are observed in intracellular compartments [11–13], extracellular structures such as the mussel byssus [14], and polymer solutions [10,15,16]. The two-phase state can be stabilized in polyelectrolyte proteins by neutralization with an oppositely charged polymer or molecule as seen in gelatin and gum-arabic [17] and in *Glycera* jaw multi-tasking protein and Cu^{2+} [18]. Coacervation of charged proteins is not confined to electrostatic interactions alone but can

also be induced via hydrophobic interactions [19], π-cation interactions [20], and repulsion in poor-solvent solutions [10].

Thesis purpose and aim

The aim of this work is two-fold: 1) to compare the biochemical and thermal properties of fish scale and bovine skin type I collagens; 2) to explore the liquid-liquid phase separation behavior of fish scale collagen contrasted with bovine collagen solution behavior.

References

(1)   McEwen, E.; Miller, R. L.; Bergman, C. A. Early Bow Design and Construction. *Sci. Am.* **1991**, *264* (6), 76–83.
(2)   Canty, E. G.; Kadler, K. E. Procollagen Trafficking, Processing and Fibrillogenesis. *J. Cell Sci.* **2005**, *118* (Pt 7), 1341–1353. https://doi.org/10.1242/jcs.01731.
(3)   Ockerman, H. W.; Hansen, C. L. *Animal By-Product Processing*; Ellis Horwood series in food science and technology; VCH: Cambridge New York, NY Basel (Switzerland) Weinheim, 1988.
(4)   Bella, J. A New Method for Describing the Helical Conformation of Collagen: Dependence of the Triple Helical Twist on Amino Acid Sequence. *J. Struct. Biol.* **2010**, *170* (2), 377–391. https://doi.org/10.1016/j.jsb.2010.02.003.
(5)   Cutini, M.; Bechis, I.; Corno, M.; Ugliengo, P. Balancing Cost and Accuracy in Quantum Mechanical Simulations on Collagen Protein Models. *J. Chem. Theory Comput.* **2021**, *17* (4), 2566–2574. https://doi.org/10.1021/acs.jctc.1c00015.
(6)   Stein, H.; Wilensky, M.; Tsafrir, Y.; Rosenthal, M.; Amir, R.; Avraham, T.; Ofir, K.; Dgany, O.; Yayon, A.; Shoseyov, O. Production of Bioactive, Post-Translationally Modified, Heterotrimeric, Human Recombinant Type-I Collagen in Transgenic Tobacco. *Biomacromolecules* **2009**, *10* (9), 2640–2645. https://doi.org/10.1021/bm900571b.
(7)   Leikina, E.; Mertts, M. V.; Kuznetsova, N.; Leikin, S. Type I Collagen Is Thermally Unstable at Body Temperature. *Proc. Natl. Acad. Sci.* **2002**, *99* (3), 1314–1318. https://doi.org/10.1073/pnas.032307099.
(8)   Subhan, F.; Ikram, M.; Shehzad, A.; Ghafoor, A. Marine Collagen: An Emerging Player in Biomedical Applications. *J. Food Sci. Technol.* **2015**, *52* (8), 4703–4707. https://doi.org/10.1007/s13197-014-1652-8.
(9)   Berillis, P. Marine Collagen: Extraction and Applications. 13.
(10)   Booij, H. L.; Bungenberg de Jong, H. G. *Biocolloids and their Interactions*; Springer Vienna: Vienna, 1956. https://doi.org/10.1007/978-3-7091-5456-4.
(11)   Rai, A. K.; Chen, J.-X.; Selbach, M.; Pelkmans, L. Kinase-Controlled Phase Transition of Membraneless Organelles in Mitosis. *Nature* **2018**, *559* (7713), 211–216. https://doi.org/10.1038/s41586-018-0279-8.

(12)   Watanabe, K.; Morishita, K.; Zhou, X.; Shiizaki, S.; Uchiyama, Y.; Koike, M.; Naguro, I.; Ichijo, H. Cells Recognize Osmotic Stress through Liquid–Liquid Phase Separation Lubricated with Poly(ADP-Ribose). *Nat. Commun.* **2021**, *12* (1), 1353. https://doi.org/10.1038/s41467-021-21614-5.

(13)   Zhang, X.; Vigers, M.; McCarty, J.; Rauch, J. N.; Fredrickson, G. H.; Wilson, M. Z.; Shea, J.-E.; Han, S.; Kosik, K. S. The Proline-Rich Domain Promotes Tau Liquid–Liquid Phase Separation in Cells. *J. Cell Biol.* **2020**, *219* (11). https://doi.org/10.1083/jcb.202006054.

(14)   Wei, W.; Tan, Y.; Martinez Rodriguez, N. R.; Yu, J.; Israelachvili, J. N.; Waite, J. H. A Mussel-Derived One Component Adhesive Coacervate. *Acta Biomater.* **2014**, *10* (4), 1663–1670. https://doi.org/10.1016/j.actbio.2013.09.007.

(15)   McCarty, J.; Delaney, K. T.; Danielsen, S. P. O.; Fredrickson, G. H.; Shea, J.-E. Complete Phase Diagram for Liquid–Liquid Phase Separation of Intrinsically Disordered Proteins. *J. Phys. Chem. Lett.* **2019**, *10* (8), 1644–1652. https://doi.org/10.1021/acs.jpclett.9b00099.

(16)   Levenson, R.; Bracken, C.; Bush, N.; Morse, D. E. Cyclable Condensation and Hierarchical Assembly of Metastable Reflectin Proteins, the Drivers of Tunable Biophotonics*. *J. Biol. Chem.* **2016**, *291* (8), 4058–4068. https://doi.org/10.1074/jbc.M115.686014.

(17)   Tiebackx, F. W. Gleichzeitige Ausflockung zweier Kolloide. *Z. Für Chem. Ind. Kolloide* **1911**, *8* (4), 198–201. https://doi.org/10.1007/BF01503532.

(17)   Wonderly, W. R.; Nguyen, T.; Malollari, K.; DeMartini, D.; Delparastan, P.; Valois, E.; Messersmith, P.B.; Helgeson, M.; Waite, J.H.; Multi-Tasking Polypeptide from Bloodworm Jaws: Catalyst, Template, and Copolymer in Film Formation. Unpublished manuscript, **2021**

(19)   Lin, Y.; Fichou, Y.; Longhini, A. P.; Llanes, L. C.; Yin, P.; Bazan, G. C.; Kosik, K. S.; Han, S. Liquid-Liquid Phase Separation of Tau Driven by Hydrophobic Interaction Facilitates Fibrillization of Tau. *J. Mol. Biol.* **2021**, *433* (2), 166731. https://doi.org/10.1016/j.jmb.2020.166731.

(20)   Kim, S.; Yoo, H. Y.; Huang, J.; Lee, Y.; Park, S.; Park, Y.; Jin, S.; Jung, Y. M.; Zeng, H.; Hwang, D. S.; Jho, Y. Salt Triggers the Simple Coacervation of an Underwater Adhesive When Cations Meet Aromatic π Electrons in Seawater. *ACS Nano* **2017**, *11* (7), 6764–6772. https://doi.org/10.1021/acsnano.7b01370.

Methods

Fish Scale Collagen Purification

The FSc extract purification protocol is presented chronologically in Diagram 1.

Step 1. Rinse scales thoroughly with water
All subsequent steps use pre-chilled solutions and are done at 4ºC.
Stir bars are used for all incubation steps

Step 2. Incubate in 10% NaCl for **2 days**
(1 L / 50 g scales)

Step 3. Rinse thoroughly with water

Step 4. Incubate 0.5M EDTA 0.05 Tris pH 7.5 for **4 hrs**
(1 L / 50 g scale)

Repeat **Step 3. (**Rinse thoroughly with water)

Step 5. Incubate in 5% acetic acid for **3 days**
(250 mL / 50 g scales)

Step 6. Blend/homogenize for **10 minutes** and incubate
for another **1 hour** using stir bar

Step 7. NaCl added to 0.7M by adding 4.2 M NaCl
(1:5 NaCl:scale solution)

Step 8. Rinse pellet with cold water
Resuspend in 5% acetic acid

Repeat Steps 6-7 (twice)

Repeat Steps 6-7 except: resuspend in 10 mM HCl
instead of 5% acetic acid

Step 8: Dialyze against 10 mM HCl (total dialysis
volume ≈36,000X sample volume)

Step 9: Concentrate using Amicon™ Ultra-15
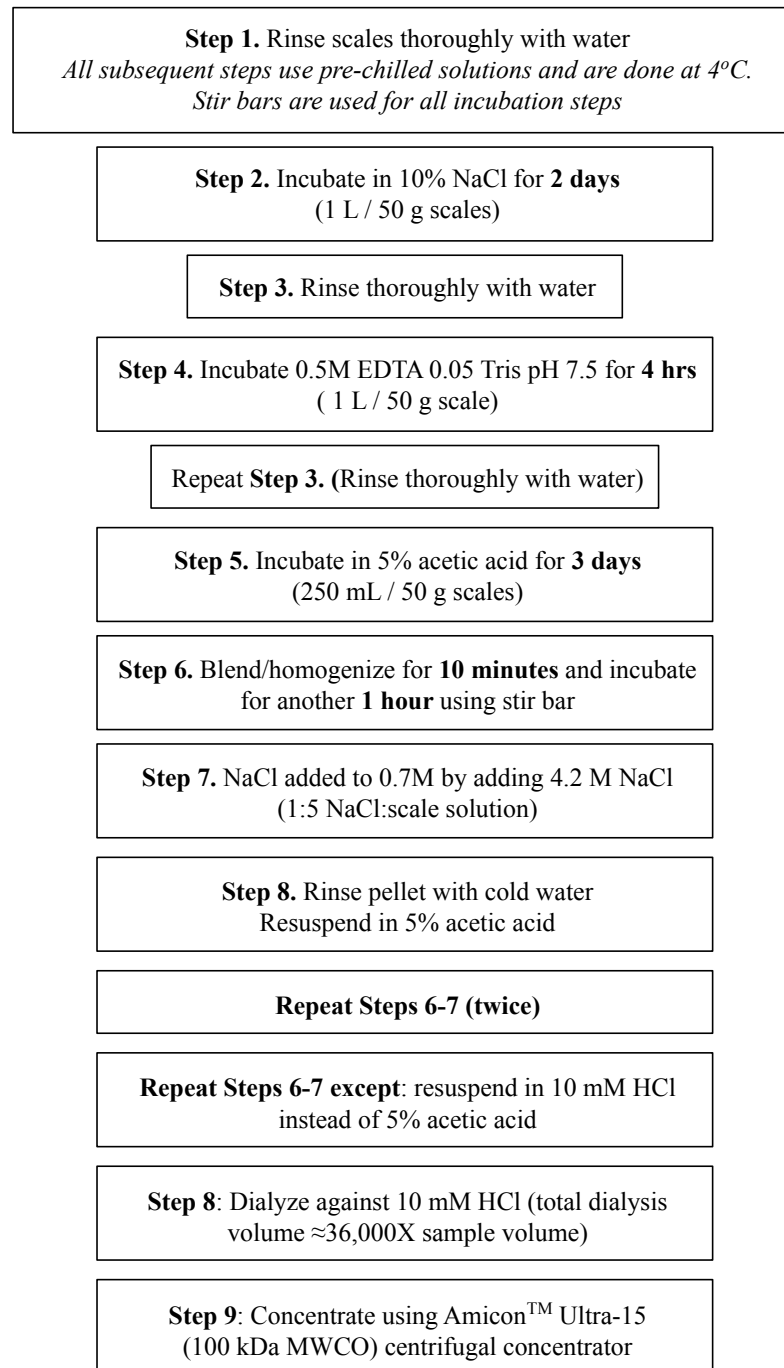(100 kDa MWCO) centrifugal concentrator

Diagram 1. Outline of purification protocol. Long term storage of purified collagen solutions was accomplished by freezing dilute collagen solutions (3 mg/mL) from **Step 8** at -80 ˚C.

Polyacrylamide Gel Electrophoresis

Sodium dodecyl sulfate (SDS) gels were made to 10% acrylamide. Samples were diluted 1:1 with 2X Laemmli sample buffer (Bio-Rad, #1610737) and run at 90 V for approximately 1.5 hours. After which, gels were incubated with staining solution containing 0.125% Coomassie brilliant blue G-250 (Bio-Rad, #1610406) for 30 minutes and then transferred to destaining solution (50% methanol 10% acetic acid) and incubated overnight.

Acetic acid-urea (AU) gels were made to 7.5% and pre-equilibrated before being stored in 5% acetic acid until use. Protein samples were diluted 1:1 with 2X AU sample buffer (5% acetic acid, 8M urea) and run for 1.5-2 hours at 10mA with reverse current direction in 5% acetic acid as the running buffer. Gels were then stained with 0.1% Coomassie brilliant blue G-250 in 40% methanol 20% acetic acid for 30 minutes to 1 hour and destained overnight in 20% methanol 10% acetic acid.

Amino Acid Analysis

Approximately 3 mg of freeze-dried sample (Labconco Lyph-lock 6, Kansas City, MO) from step 8 (purification chart) was transferred to a Wheaton Glass 1 mL ampule (#651502) and incubated with 190 µL 6 N constant boiling HCl and 10 µL phenol for 24 hours at 110 ˚C. The hydrolyzed solution was then transferred to a centrifuge tube. The remaining amino acid solution in the glass ampule was recovered by consecutive rinsing with 200 µL MilliQ H_2O which was then transferred to the respective sample centrifuge tube. The sample liquid was then flash evaporated using a Savant SpeedVac vacuum concentrator (Holbrook, NY). The dried residue was then resuspended in 50 µL H_2O and flash evaporated (3x), which was then repeated with 50 µL ethanol (3x). Once dry, amino acids were resuspended in 100 µL 0.02 N HCl and centrifuged to remove any insoluble. The clarified solution was then loaded into a

Hitachi High-Technologies model L8900 analyzer.

<u>LC-MS/MS Sample Preparation and</u>

Purified Atlantic salmon scale extract was mixed 1:1 with 2X Laemmli buffer (100 mM DTT) and incubated for 10 minutes at 70˚C. 12% Bolt[TM] Bis-Tris gels (Invitrogen) were run at 110V for 1.5 hours in MOPS running buffer (Invitrogen). The gel was then incubated with staining solution containing 0.1% Coomassie Brilliant Blue R-250 (Bio-Rad), then destained in 30% methanol followed by water. Protein bands associated with the α1 and α2 monomers were excised using a sterile scalpel blade, covered with 0.1% acetic acid. Gel samples were shipped to Bioproximity for tryptic digestion and LC-MSMS (Thermo Q-Exactive HF-X Orbitrap mass spectrometer). Note: acetic acid and water used were LC-MS grade (Optima).

Identification of protein sequences was accomplished by using PEAKS Studio XPro (Bioinformatics Solutions) to search open-reading frames of the Atlantic Salmon genomic library (Uniprot[1] genome accession: GCF_000233375) and de novo assemble LC-MS/MS results. Results were filtered with a 1% false discovery rate threshold, and a minimum of 5% ion intensity for PTMs.

<u>Confocal Microscopy Sample Preparation and Image Acquisition</u>

Images were acquired on a Leica SP8 FLACON Resonant Scanning Confocal Microscope (40× water objective) equipped with a cryo-incubator stage. Undesired sample heating, which could initiate LLPS before measurement, is posed by two risks: (1) heat transfer during mounting of the samples onto the microscope stage and (2) sample exposure to ambient room temperature and the microscope objective. Samples were pipetted into flat capillary tubes (Friedrich & Dimmock, BMS-020-2-15-50, 0.2MMX2.0MMIDX50MM).

Each capillary was fixed to the outer bottom of a petri dish using fast curing optical adhesive (Norland, #NOA81) at both ends, consequently sealing the capillary tube ends to preventing sample leaking. In order to mitigate fast sample heating, petri dishes were filled with ice water and stored on ice before finally being mounted onto the pre-cooled microscope stage (15 ˚C). A thin-wired temperature probe was attached flush with the capillary tube during imaging to track temperature. Image acquisition started when temperature equilibrated to 10 ˚C (heating rate ≈ 1˚C/min), after which temperature was stable and could be sustained by addition of ice to petri dish. The incubator stage temperature was then increased to 20 ˚C which resulted in < 1 ˚C/min temperature increase from 10 ˚C. For images taken above 20 ˚C, the temperature stage was set to 30 ˚C and exhibited the same heating rate < 1 ˚C/min.

Particle Tracking and Microrheology

500 nm fluorescent polystyrene beads (Sigma-Aldrich) in 10 mM HCl were sonicated in a water bath for 15 minutes to disperse and aggregated beads. Sonicated bead solutions were added to several concentrations of purified Atlantic salmon scale extract solutions at a 0.2% volume fraction. Capillary tubes and petri dishes were prepared as described in confocal image acquisition. Images were acquired at a frame rate of 1 frame every 0.04 seconds (500 frames total) and contained 10-40 beads within the microscope field of view. Leica image files were converted to compatible particle tracking algorithm files in ImageJ.

Circular Dichroism

Collagen solutions were made to 0.1 mg/mL protein in 10 mM HCl and stored on ice. Collagen concentration was confirmed by UV-Vis absorbance at 230 nm and a collagen standard of known concentration. Circular dichroism (CD) measurements were collected on a Jasco J-1500 spectropolarimeter equipped with a temperature-controlled holder. Keeping the

sample on ice, 200 μL of solution was loaded directly into a 1mm quartz cell (Starna, #1-Q-1) which is pre-equilibrated to 4 ˚C inside the sample holder. Temperature was tracked by a thin probe placed inside the cell and in contact with the sample solution. CD spectra were measured from 190 nm to 260 nm with a 50 nm/min scanning speed and 1 nm bandwidth. Thermal stability measurements were taken by following the intensity at 220 nm at a 1 ˚C/min heating rate with temperature ramping paused during acquisitions.

Stability Heatmap

To visualize and compare regions of triplet stability in *Salmo salar* and *Bos taurus* monomer sequences, a 1D heatmap was generated using Seaborn Plugin in Python 3.7 and triplet stability values from an online tool called *collagen stability calculator*[2]. The stability values in this algorithm are derived from melting experiments of short synthetic peptides at pH = 7 and a heating rate of 0.1˚C/min. Sequences were analyzed starting from the first GXY repeat and ending at the last GXY repeat, excluding non-GXY sequences at the C- and N-terminus of the monomer chain.

Materials

Acid-solubilized bovine type I collagen (#C4243) from Sigma-Aldrich (St. Louis, MO). Fish scales for FSc preparations were obtained from Santa Barbara Fish Market, Santa Barbara, CA.

Results

Polyacrylamide Gel Electrophoresis of Fish Scale Extract
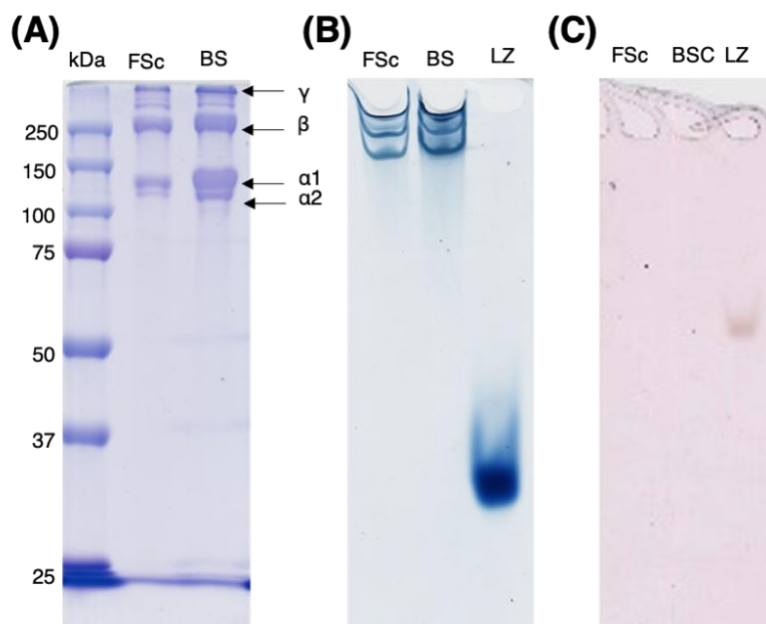
Figure 1. Polyacrylamide Gel electrophoresis shows similar banding of BS and FSc samples (**A**) SDS PAGE of FS (lane 2) and BS (lane 3) collagens stained with Coomassie dye with a standard molecular weight ladder in lane 1. (**B**) Coomassie and (**C**) PAS stained acetic acid-urea polyacrylamide gels of FSc extract (lane 1) and BS collagen (lane 2) and lysozyme (LZ, lane 3) Note: PAS stained gel (**C**) was run for a shorter period of time, which resulted in the reduced migration of the lysozyme band

Our first approach to identify the composition of the fish scale extracts was utilizing sodium dodecyl sulfate (SDS) polyacrylamide gel electrophoresis (PAGE) and using Coomassie dye to stain all protein bands present. SDS PAGE shows characteristic α1 and α2 monomers, β dimers, and γ multimer banding patterns (Fig 1A) for both fish scale (FSc, lane 2) and bovine skin (BS, lane 3) collagen samples. The relative intensity of the α1 bands are higher than the α2 bands intensity for both samples. The BS collagen is a heterotrimer with a (α1)$_2$α2 mixed monomer composition, which results in a higher α1 monomer staining intensity.

In order to screen for significant variations in charge between BS and FSc proteins, acetic acid-urea (AU)-PAGE was utilized to separate proteins based on both size and effective charge[3]. The acidic conditions of AU-PAGE (pH ≈ 3) allows basic proteins to migrate through the gel when the current is run in the reverse direction from SDS-PAGE (basic proteins will migrate towards the cathode). This technique is useful for detecting variations in positive charge density and post-translational modifications of a protein, such as levels of acetylation, which will result in different migration patterns in otherwise identical proteins. Both the BS and FSc samples show three bands with nearly identical migration patterns when using acetic acid-urea PAGE (Fig. 1B). Periodic Schiff Staining (PAS) of native gels shows no detectable levels of polysaccharide staining, suggesting negligible glycosylation. For reference, a sample of lysozyme (LZ), which has N-linked Glc and GlcNac, is included in lane 3 (Fig. 1B-C).

Amino Acid Analysis and Sequence Determination

Figure 2. Amino acid analysis of Atlantic salmon (orange dashed line), black rockfish (yellow solid), and ocean whitefish (blue dots) scale extracts along with bovine skin type I collagen (Sigma-Aldrich, #C4243).
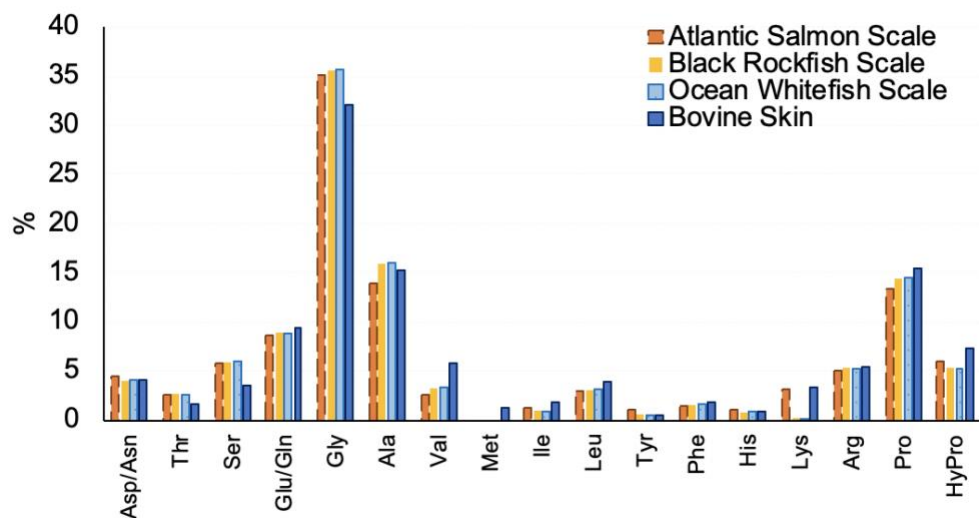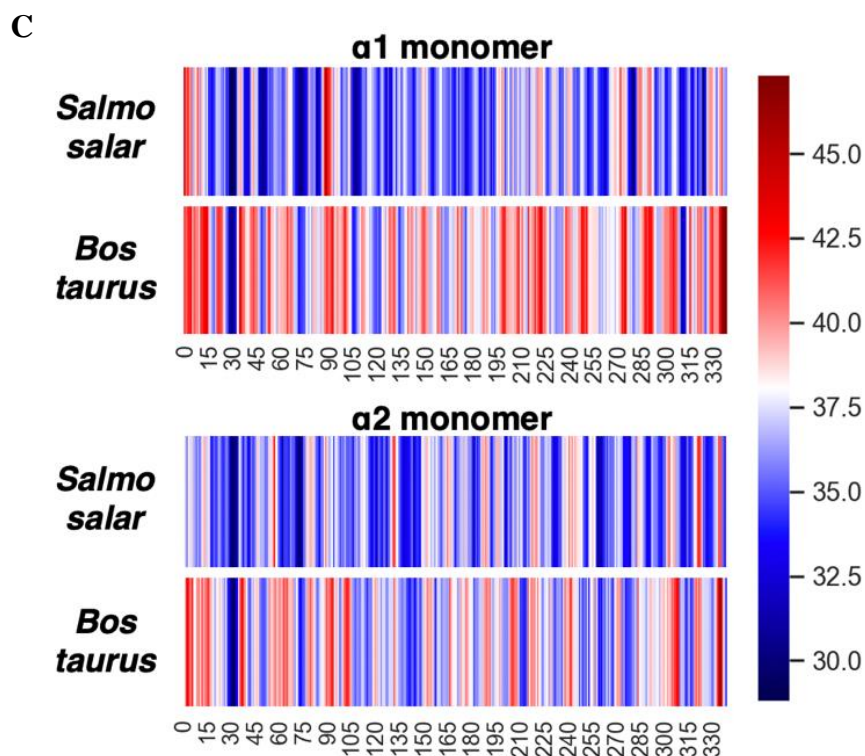
A

```
                        172       182       192       202       212       222       232       242
CO1A1_BOVIN       163 LSYGYDEKSTG--ISVPGPMGPSGPRGLPGPPGAPGPQGFQGPPGEPGEPGASGPMGPRGPPGPPGKNGDDGEACKPGRPGERGPPGPQG
A0A1S3Q7E3_SALSA  151 MHHGYDQKSGGGQAQMGPMGPMGPRGLAGPPGNTGPPGHHGPPGEPGEPGSSGPMGERGAMGPHGKNGDNGEPGRSGRPGERGSAGPQG

                        252       262       272       282       292       302       312       322       332
CO1A1_BOVIN       251 ARGLPGTAGLPGMKGHRGFSGLDGAKGDAGPAGPKGEPGSPGENGAPGQMGPRGLPGERGRPGAPGPAGARGNDGATGAGPPGPTGPA
A0A1S3Q7E3_SALSA  240 GRGTPGTPGLPGIKGHRGFNGMDGAKGDGGPAGPKGEGGAPGSNGNSGAMGARGLPGERGRPGPAGSSGARGNDGNSGPAGSNGPTGPS

                        342       352       362       372       382       392       402       412       422
CO1A1_BOVIN       340 GPPGFPGCAVGAKGEGGPQQGPRGSEGPQGVRGEPGPPGPAGAAGPAGNPGADGQPGAKGANGAPGIAGAPGFPGCARGPSGPQGPSGPPGP
A0A1S3Q7E3_SALSA  329 GAPGFPGGAGAKGETGPQGCRGSDGPQGSRGEPGNPGPAGAAGPSGAPGSDGQAGGKGSTGAAGISGAPGFPGTRGPAGPPGSGGGPGP

                        432       442       452       462       472       482       492       502       512
CO1A1_BOVIN       429 KGNSGEPGAPGSKGDTGAKGEPGPTGIQGPPGPAGEEGKRGARGEPGPAGLPGPPGERGGPGSRGFPGADGVAGPKGPAGERGAPGPAG
A0A1S3Q7E3_SALSA  418 KGNHGEHGAQGPKGCEPGVKGEPGSTGPGGPPGPSGEEGKRGARGEPGGAGGVGPPGERGMPGCMRGFPGSDGAAGCKGGPGERGANGCPMG

                        522       532       542       552       562       572       582       592       602
CO1A1_BOVIN       518 PKGSPGEAGRPGEAGLPGAKGLTGSPGSPGPDGKTGPPGPAGQDGRPGPPGPPGARGQAGVMGFPGPKGAAGEPGKAGERGVPGPPGCAV
A0A1S3Q7E3_SALSA  507 AQGCATGEAGRPGEPGMPGSKGMTGSPGSPGPDGKVGPAGVTGNDGRSGPPGSSGARGQPGVMGFPGPKGPAGENGKPGERGEAGAVGGL

                        612       622       632       642       652       662       672       682       692
CO1A1_BOVIN       607 GPAGKDGEAGAQGPPGPAGPAGERGEQGPAGSPGFQGLPGPAGPPGEAGKPGEQGVPGDLGAPGPSGARGERGFPGCERGVQGPPGPAGP
A0A1S3Q7E3_SALSA  596 GAPGKDGDSGAPGAQGPAGAQGEKGEGGPAGPNGFQGLPGCTQGATGEPGKAGEQGVPGEVGGPGPSGSRGDRGFPGCERGGLGSAGPTGP

                        702       712       722       732       742       752       762       772       782
CO1A1_BOVIN       696 RGANGAPGNDGAKGDAGAPGAPGSQGAPGLQGMPGERGAAGLPGPKGDRGDAGPKGADGAPGKDGVRGLTGPIGPPGPAGAPGDKGEAG
A0A1S3Q7E3_SALSA  685 RGANGSPGNDGARGESGAAGAPGGMGAPGLQGMPGERGSSGNSGAKGERGDGGPKGADGGPGKDGMRGMTGPIGPPGPTGAHGEKGEGG

                        792       802       812       822       832       842       852       862       872
CO1A1_BOVIN       785 PSGPAGPTGARGAPGDRGEPGPPGPAGFAGPPGADGQPGAKGEPGDAGAKGDAGPPGPAGPAGPPGPIGNVGAPGPKGARGCSAGPPGAT
A0A1S3Q7E3_SALSA  774 LGGPPGPTGGRGSPGERGEHGAPGPAGFAGPPGADGQPGNKGEAGNNGPKGEAGAPGPGVGAPGPQGPAGNSGAKGTRGAPGPPGAS

                        882       892       902       912       922       932       942       952       962
CO1A1_BOVIN       874 GFPGAAGRVGPPGPSGNAGPPGPPGPAGKEGSKGPRGETGPAGRPGEVGPPGPPGPPGAGEKGAPGADGPAGAPGTPGPQGIAGQRGVVGL
A0A1S3Q7E3_SALSA  863 GMPGPGGRVGPPGGSGAPGSAGPPGPAGKEGQRGGRGETGNAGRPGEAGAAGPPGPSGPSGSGKGNDGPMGAPGTPGPGGIAGQRGIVGG

                        972       982       992       1002      1012      1022      1032      1042
CO1A1_BOVIN       963 PGQRGERGFPGLPGPSGEPGKQGPSGASGERGPPGPMGPPGLAGPPGESGREGAPGAEGSPGRDGSPGAKGDRGETGPAGPPGAPGAPG
A0A1S3Q7E3_SALSA  952 PGGRGPSGTAGTPGPAGEPGKQGPGGPVGERGSPGPMGPPGLSGAPGEAGREGSAGHDGVSGRDGPPGPKGDRGENGNAGSPGAPGPPG

                        1062      1072      1082      1092      1102      1112      1122      1132
CO1A1_BOVIN      1052 APGPVGPAGKSGDRGETGPAGPAGPIGPVGARGPAGPQGPRGDKGETGEQGDRGIKGHRGFSGLQGPPGPPGSPGEQGPSGASGPAGPR
A0A1S3Q7E3_SALSA 1041 APGPSGASGKTGSRGESGPAGPAGPNGPAGGRGGLGPAGARGDKGEAGEGGERGHKGHRGLNGMPGLPGPSGSPGESGPAGLNGPSGPR

                        1142      1152      1162      1172      1182      1192      1202      1212      1222
CO1A1_BOVIN      1141 GPPGSAGSPGKDGLNGLPGPIGPPGPRGRTGDAGPAGPPGPPGPPGPPGPPGPSGGYDLSFLPQP---PQEKAHDGGRYYRAD
A0A1S3Q7E3_SALSA 1130 GPHGSSGSAGKDGMNGLPGPIGPPGPRGRNGEMGAAGAPGSPGLPGPPGPAGGGFHVAPIYNQEKGPDPMRGGGGYHYRAD
```

B

```
                        89        99        109       119       129       139       149       159
CO1A2_BOVIN        80 QFDAKG---GGPGPMGLMGPRGPPGASGAPGPQGFQGPPGEPGEPGQTGPAGARGPPGPPGKAGEDGHPGKPGRPGERGVVGPQGAR
A0A1S3M538_SALSA   73 QFDGGKGSDPGPGPMGLMGSRGPNGPPGSPGPQGFTGHAGEPGEPGQTGAIGARGPPGSAGKPGEDGNNGRPGKPGDRGGPGTQGAR

                        169       179       189       199       209       219       229       239       249
CO1A2_BOVIN       164 GFPGTPGLPGFKGIRGHNGLDGLKGQPGAPGVKGEPGAPGENGTPGQTGARGLPGERGRVGAPGPAGARGSDGSVGPVGPAGPIGSA
A0A1S3M538_SALSA  160 GFPGTPGLPGMKGHRGYNGLDGRKGESGTAGAKGETGAHGANGNPGPAGSRGLNGERGRAGPAGPAGARGADGSTGPAGPAGPLGAA

                        259       269       279       289       299       309       319       329       339
CO1A2_BOVIN       251 GPPGFPGPAPGPKGELGPVGNPGPAGPAGPRGEVGLPGCLSGPVGPPGNPGANGLPGAKGAAGLPGVAGAPGLPGPRGIPGPVGAAGAT
A0A1S3M538_SALSA  247 GPPGFPGAPGPKGEIGGVGSNGPSGPQGGRGEPGINGAVGPVGPVGNPGNNGINGAKGAAGLPGVAGAPGFPGPRGGPGPQGPQGST

                        349       359       369       379       389       399       409       419
CO1A2_BOVIN       338 GARGLVGEPGPAGSKGESGNKGEPGAVCQPGPPGPSGEEGKRGSTGEIGPAGPPGPGLRGNPGSRGLPGADGRAGVMGPAGSRGAT
A0A1S3M538_SALSA  334 GARGLGGDPGPSGQKGDSGAKGEPGHSGVQGPAGPAGEEGKRGSTGEAGATGPAGLRGARGGAGTRGLPGLEGRGGPIGMPGARGAT

                        429       439       449       459       469       479       489       499       509
CO1A2_BOVIN       425 GPAGVRGPNGDSGRPGEPGLMGPRGPPGSPGNIGPAGKEGPVGLPGIDGRPGPIGPAGARGEPGNIGFPGPKGPSGDPGKAGEKGHA
A0A1S3M538_SALSA  421 GPAGIRGAPGDAGRAGESGLTGARGLPGNSGQGGPPGKEGPSGAAGLDGRTGPPGPTGPRGQPGNIGFPGPKGPGGEAGKGGDKGPT

                        519       529       539       549       559       569       579       589       599
CO1A2_BOVIN       512 GLAGARGAPGPDGNNGAQGPPGLQGVQGGKGEQGPAGPPGFQGLPGPAGTAGEAGKPGERGIPGEFGLPGPAGARGERGPPGESGAA
A0A1S3M538_SALSA  508 GATGLRGGPGADGNNGAPGPAGVVGNAGEKGEQGPSGAPGFQGLPGPAGPAGEAGKAGNQGMPGDQGLPGPAGVKGERGNSGPAGSA

                        609       619       629       639       649       659       669       679
CO1A2_BOVIN       599 GPTGPIGSRGPSGPPGPDGNKGEPGVVGAPGTAGPSGPSGLPGERGAAGIPGGKGEKGETGLRGDIGSPGRDGARGAPGAICAPGPA
A0A1S3M538_SALSA  595 GSQGAIGARGPAGTPGPDGGKGEPGSVGIVGAAGHQGPGGMPGERGAGGTPGPKGEKGEGGHRGLEGNMGRDGARGAPGPSGPPGPS

                        689       699       709       719       729       739       749       759       769
CO1A2_BOVIN       686 GANGDRGEAGPAGPAGPAGPRGSPGERGEVGPAGPNGFAGPAGAAGQPGAKGERGTKGPKGENGPVGPTGPVGAAGPSGPNGPPGPA
A0A1S3M538_SALSA  682 GANGEKGESGSFGPAGPAGLRGPSGERGEGGPAGPPGFAGPPGSDGQSGPRGEKGPAGGKGDVGPAGPAGPAGQSGPSGASGPAGPP

                        779       789       799       809       819       829       839       849       859
CO1A2_BOVIN       773 GSRGDGGPPGATGFPGAAGRTGPPGPSGISGPPGPPGPAGKEGLRGPRGDGQPGVGRSGETGASGPPGFVGEKGPSGEPGTAGPPGTP
A0A1S3M538_SALSA  769 GGRGDAGPSGLTGFPGAAGRVGGPGPAGIAGPPGSAGPAGKDGPRGLRGDPGPGGPQGEPGEGAVGPAGITGDKGPSGESGPPGPPGTA

                        869       879       889       899       909       919       929       939
CO1A2_BOVIN       860 GPQGLLGAPGFLGLPGSRGERGLPGVAGSVGEPGPLGIAGPPGARGPPGNVGNPGVNGAPGEAGRDGNPGNDGPPGRDGQPGHKGER
A0A1S3M538_SALSA  856 GPQGVLGHAGFVGLPGSRGDKGLPGGPGAVGEPGRLGPAGASGPRGPSGNIGMPGMTGTQGEAGREGNPGNDGPPGRPGAAGFKGDR

                        959       969       979       989       999       1009      1019      1029
CO1A2_BOVIN       947 GYPGNAGPVGAAGAPGPQGPVGPVGKHGNRGEPGPAGAVGPAGAVGPRGPSGPQGIRGDKGEPGDKGPRGLPGLKGHNGLQGLPGLA
A0A1S3M538_SALSA  943 GEPGSPGALGSSGQPGPNGPAGSAGRPGNRGESGPTGNGGPVGAAGARGAPGPAGPRGEKGGACEKGDRGMKGLRGHGGLQGMPGPN

                        1039      1049      1059      1069      1079      1089      1099      1109      1119
CO1A2_BOVIN      1034 GHHGDQGAPGAVGPAGPRGPAGPSGPAGKDGRIGQPGAVGPAGPAGIRGSQGSQGPAGPPGPPGPPGPPGPSGGGYEFGFDGDFYRA
A0A1S3M538_SALSA 1030 GPSGETGSAGITGPAGPRGPAGPHGPPGKDGRAGGHGAIGPVGHRGSPGHLGPAGPPGSPGLPGPAGPAGGGYDQSGGYDEYRA
```

C

Figure 3. (**A**) Alignment of α1 monomers with _% identity and (**B**) α2 monomer sequences with _% identity (**C**) Heatmap of relative triplet stability of aligned sequences using values predicted from the collagen stability calculator[11]

Uniprot identifier and sequence ranges: * P02453|163-1217 and P02465|80-1117; ** A0A1S3Q7E3|162-1210 and A0A1S3M538|73-1112

LC-MS/MS analysis of peptides produced by in-gel trypsin digests had hits for several isomers of *Salmo salar* type I collagen monomer Uniprot sequences. The UniProtKB entry names with the highest coverage (≈ 50%) were salsa_A0A1S3Q7E3 for the α1 monomer and salsa_A0A1S3M538 for the α2 monomer. Shown in the sequence alignment are the helical regions of bovine monomers and the aligned *Salmo salar* sequence (Fig 3A). Supplementary table 1 contains molecular weight values (not including post-translational modification), isoelectric points, and amino acid preferences of the monomer sequences. The *Salmo salar* and *Bos taurus* triple helix domain sequences show a preference for glycine, proline, and alanine residues (Fig. 3C). It is worth noting that a significant portion of prolines will be post-translationally modified to 4-hydroxyproline when found in the Y position of the

13

GXY repeat, as can be evidenced by the hydroxyproline content from amino acid analysis (Fig. 2). The most notable difference between the two collagen species is the 23% decrease in Pro composition in the Atlantic salmon collagen, where Pro is most commonly substituted by Gly and Ser residues.

To visualize and compare regions of triplet stability in *Salmo salar* and *Bos taurus* monomer sequences, a 1D heatmap was generated using Seaborn Plugin in Python 3.7 and triplet stability values from an online tool called *collagen stability calculator*[2]. The stability values in this algorithm are derived from melting experiments of short synthetic peptides at pH = 7 and a heating rate of 0.1˚C/min. Triplet sequences were analyzed sequentially starting from the first GXY repeat and ending at the last GXY repeat, therefore excluding non-GXY sequences at the C- and N-terminus of the monomer chain.

The relative temperature stability values are represented as colors where blue indicates low temperature stability and red indicates high temperature stability. The salmon monomers have significantly lower stability than the bovine monomers as seen by the predominantly blue color of the heat map with sparse, narrow pink regions.

Thermal Stability of Atlantic salmon collagen
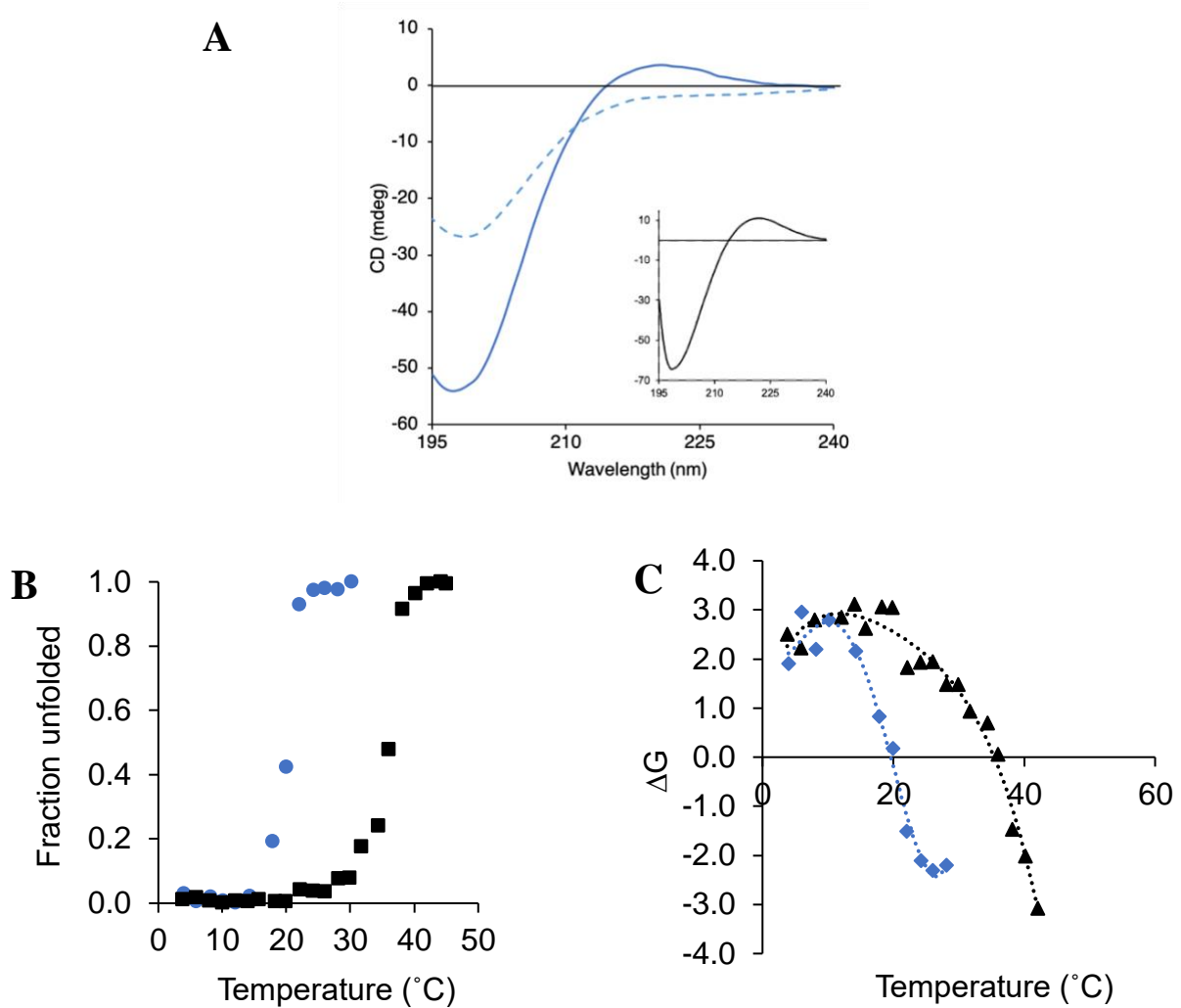
Figure 4. (**A**) CD spectra of FS collagen at 4 ˚C (solid line) and 25 ˚C (dashed line). CD spectrum of BS collagen at 4 °C with known triple helical structure (inset) is provided for comparison (**B**) Thermal stability of the BS and FSc collagen helices based on ellipticity at 220 nm and at 1 ˚C/min heating rate. Bovine skin collagen (black squares) has a $T_m \approx 36$ ˚C and Atlantic salmon scale collagen (blue circles) has a $T_m \approx 20$ ˚C. (**C**) Free energy (ΔG) calculated from CD thermal stability values in (B) indicates lower stability in FSc collagen. All solutions are 0.1 mg/mL collagen in 10 mM HCl at pH = 2.

The circular dichroism spectrum of FSc extract resembles a type II polyproline helix

at 4 ˚C (Fig 4A, line). The positive and negative peaks have lower intensity than the bovine

skin collagen solution with known triple helical structure (Fig. 4A, inset). The degree of

helicity can be described using the ratio of positive to negative peak (Rpn) which we

15

quantified by taking the absolute ratio of intensities of the 220 nm peak and the 199 nm peak. The Rpn for both samples falls within range of standard triple helix values[4], with bovine (Rpn = 0.16) being $\approx$ 2.5 times more helical than salmon (Rpn = 0.067) at 4 ˚C. Once temperature is increased incrementally to 25 ˚C the fish scale collagen spectrum resembles a random-coil signature (Fig 4A, dashed line).

The effect of temperature on conformational stability is determined by following the CD intensity at a single wavelength (220 nm) as temperature is incrementally increased and converting it to fraction unfolded as done previously.[5,6] The transition to random coil occurs between 20-22 ˚C for FSc collagen and 36-38 ˚C for BS collagen.

Microscopy
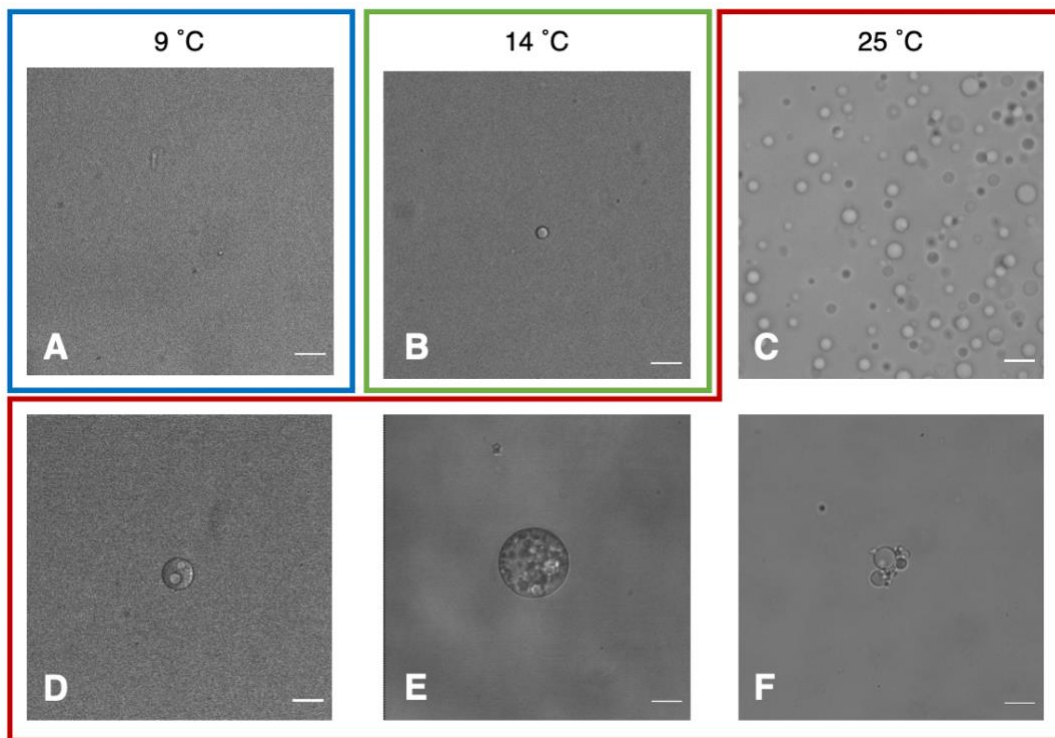
Figure 5. A FSc phase behavior. (blue box) Brightfield images of FS extract solutions in the single-phase state at 9 ˚C. **B** (green box) the solutions at an intermediate temperature (14 ˚C) between folded and unfolded collagen. **C-F** (red box) the solutions when significant phase separation is present above T_m temperature (25 ˚C). Scale bar is 10 μm. All solutions are 30 mg/mL collagen in 10 mM HCl at pH = 2.

Above its melting temperature, BS collagen solutions show some indication of precipitation but mostly appeared to remain soluble (Fig S1). FSc collagen solutions, in the same solvent conditions, appeared to undergo liquid phase separation as temperature approaches FSc collagen T_m (Fig. 5C-F). At temperatures below 10 ˚C, the solution appeared to be a single-phase solution of soluble native collagen molecules (Fig. 5A). At 14 ˚C, some small droplets begin to appear (Fig. 3B). At 25 ˚C, polydisperse droplets ranging from of 1 µm to 15 µm in diameter in size appeared and were observed to coalesce over time (Fig 5C, SI video1). Various hierarchies of multiphase species were detected in the samples (Fig. 3D-E) and they remained stable throughout image acquisition time frame (up to 1hr). The small inner droplet species moved significantly slower than droplets in the surrounding dilute phase, indicative of high viscosity of the condensed phase inside the droplets.

Microrheology of 30 mg/mL FSc extract solutions
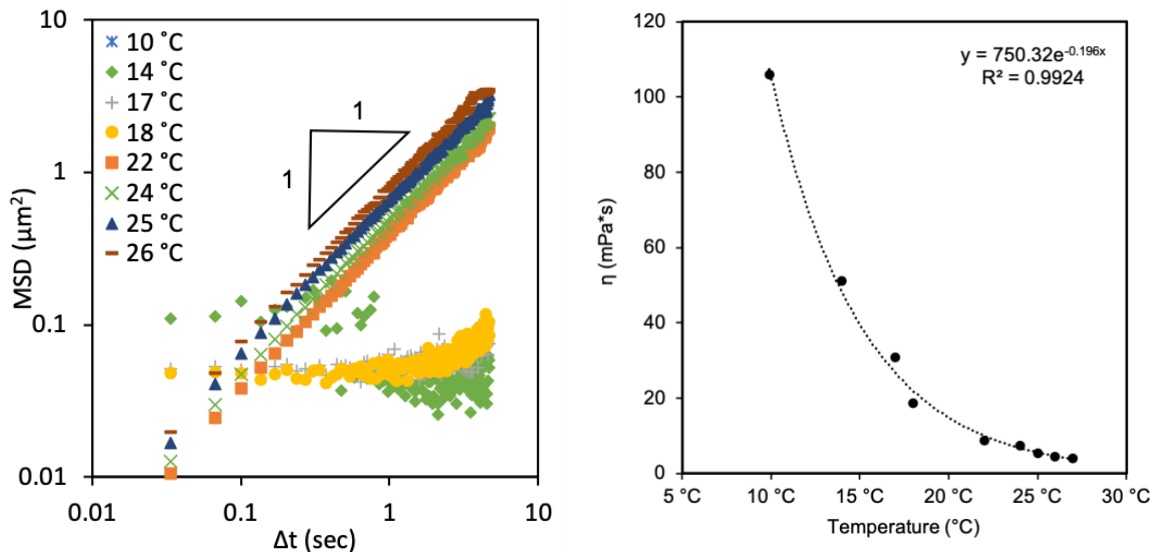
Figure 6. (**A**) Mean square displacement (MSD) values from multiple particle tracking (MPT) experiments of 30 mg/mL FSc collagen solutions in 0.01 M HCl. The logarithmic slope = 1 represents diffusive motion of particles. As temperature is increased, MSD decreases and becomes flat (**B**) Viscosity of 30mg/mL FS extract at increasing temperatures based on. The assigned fit to exponential decay serves to guide the eye.

Native fish scale collagen solutions are visibly viscous at high concentrations (15 mg/mL) and low temperature (4 ˚C) and become less viscous as the solution equilibrates with room temperature. Microrheology was employed to elucidate the phase behavior and quantify the viscosity of the bulk material (FSc collagen solution) by averaging mean-squared displacement (MSD) of fluorescent 500nm polystyrene beads and time (Fig. 6). MSD quantifies the average distance squared a particle travel for a given lag time, which is the time elapsed between two frames for which the displacement is computed. During this process, a thermocouple was attached to the sample chamber to give instantaneous readout of the sample temperature. After the sample reached room temperature (≈ 22 ˚C), it was further heated up to 30 ˚C using the incubation chamber attached to the microscope. At each temperature, a video of particle movement was recorded. The tracking procedure was carried out by using an open-source tracking algorithm[7]. Briefly in each frame, particle center was determined by fitting a smooth Gaussian profile. Thereafter, the trajectories were linked across frames by searching around each particle in the subsequent frame. In the end, MSD values were averaged across particles and time. The large number of particle trajectories ensured that the probe particles was exploring the statistically relevant average properties of the material. The MSD is a common measure of the space explored by the random motion of particles, and can be correlated with diffusion coefficient (D) using:

$$MSD = 4D\Delta t$$

Viscosity (η) can then be calculated using the Stokes-Einstein equation:

$$\eta = \frac{k_B T}{6\pi R \text{D}}$$

where k_B is the Boltzmann constant, T is the temperature in Kelvin, and R is the particle radius. The viscosity values at increasing temperatures can be seen in Fig. 6. For acquisition

of image stacks at low temperatures (10 – 20 ˚C), it was necessary to take image relatively quickly. In several instances, the resulting image stacks have a low signal-to-noise ratio and contained image artifacts introduced by the confocal line scans. In these cases, MPT was not able to properly track and link the trajectories to extract reliable information and the image stack was instead manually analyzed by particle tracking in ImageJ. Consequently, many data points below 20 ˚C had to be excluded from the current work.

The increase in temperature resulted in an increase in MSD slope values starting from 0.03 μm^2 at 10 ˚C and increased exponentially with rising temperature to a final slope value of 0.87 μm^2 at 27 ˚C (Fig. 6A), an approximately 26-fold total decrease in viscosity (Fig. 6B). The exponential decay fit assigned to temperature-dependent viscosity serves to guide the eye.

Discussion

Fish Scale Extract Purification and Compositional Analysis

The fish scale extract shows identical SDS and AU gel banding patterns as bovine skin collagen, suggesting that the major protein in FSc extract is collagen that is a $(\alpha 1)_2 \alpha 2$ heterotrimer (Fig. 1A-B). The absence of PAS-stained protein bands in both FSc and BS samples indicates that, within limits of detection, FSc collagen is not significantly more glycosylated than BS collagen. The minimal FSc collagen glycosylation further corroborates that polypeptide chains alone are sufficient for phase separation. If either additional proteins or polysaccharides were present in the FSc extract, besides collagen type I and its monomers, they were not detected on SDS PAGE or acid urea gel electrophoresis and resisted consecutive salting out steps and extensive dialysis (Chart 1). This leads us to the conclusion that there is only one macromolecular component in the extract. These assertions are further

supported by both AAA and LC-MS/MS (Fig. 2) which indicated compositional bias for

glycine, proline, and hydroxyproline as well as partial sequence that matched the Uniprot

database sequences for Atlantic salmon type I collagen monomers. Finally, CD also confirms

that the extract from fish scales is collagenous (Fig. 4A). FSc collagen had a low melting

temperature which is typical of other collagens extracted from fish [8].

Fish Scale Gelatin Coacervation

Of particular interest was our observation of multi-hierarchical and polydisperse

droplet formation and coalescence from concentrated FSc collagen solutions extracted from

scales of *Salmo salar,* but also *Caulolatilus princeps, Sebastes entomales, Sebastes mystinus,*

Sebastes melanops, and *Sebastes nigrocinctus*. Fish scale gelatin LLPS is consistent with

protein coacervation by virtue of the spherical droplet morphology and droplet coalescence

which together indicate that it is indeed in a liquid phase. Furthermore, our gel

electrophoresis (Fig. 1) and amino acid analysis (Fig. 2) results indicate the LLPS to be a

single-component, self-coacervating system.

Gelatin-based coacervate systems have been extensively described by Bungenberg de

Jong [9] in 1931 as well as more recently [10]. These studies differed significantly from the

present study in that they all relied on salts, crowding agents such as ethanol, or oppositely

charged polymers to induce phase separation. By contrast, our FSc gelatin coacervates appear

to be temperature induced, with a coacervation temperature that coincided with the FSc

collagen melting temperature (T_m). To our knowledge temperature-induced gelatin

coacervation, without mixing of LLPS-triggering agents, has not been previously reported. A

significant portion of previous LLPS studies primarily used temperature stable collagens

and/or high bloom gelatins [8-10], features that most fish scale collagens lack. Although

20

previously considered a disadvantage, it is possible that the instability of fish scale collagens in fact provided a functional purpose for both material applications and biological collagen processing which we will discuss in a following section.

<u>Collagen Stability and Phase Transition</u>

The fact that the coacervation temperature coincides with the melting temperature suggests unfolding of the coiled coil domain is necessary for phase separation. Further structural analyses of coacervating solutions are necessary to screen for transitions to alternative secondary structures such as β-structures, which have been associated with low complexity sequences [11]. Together, the fish scale extract compositional analysis, thermal stability measurements, and microscopy observations lead us to propose the model in Fig. 7 of a single component coacervate phase that is triggered by increasing temperature.

Microrheology of FSc solutions at increasing temperatures helped us quantify temperature-dependent viscosity of solutions, but also indicated a change in phase behavior at temperatures above the FSc collagen T_m. The slope value < 1 at 10 ˚C to 18 ˚C indicates a gel-like, viscoelastic behavior in the sample [12]. Concentrated bovine collagen solutions below denaturing temperatures have liquid crystal (LC)-like phase behavior and exhibit birefringence under cross polarized light (Fig. S2). FSc collagen solutions may also exhibit a LC-like state below the collagen T_m temperature or have the propensity to form LC via shear alignment [13,14]. For all temperatures above the FSc collagen T_m temperature, MSD slope = 1 and suggests Brownian motion of particles in a Newtonian fluid [12,15]. Additional rheological measurements with respect to shear rate can directly provide information on the storage and loss moduli, which can further asses unfolding and phase behavior at increasing temperatures.

Although the polystyrene beads remained in the depleted phase, the movement of small droplets within droplets provides the opportunity to assess rheological properties inside the multi-hierarchical coacervates (Fig. 5D, 5E) *(in progress)*. Preliminary observations showed contrastingly slow dispersion of droplets within the multi-hierarchical coacervate phase compared to Brownian motion of coacervates in the surrounding solvent.

Collagen melting temperatures decrease with decreasing heating rates and never experimentally reach a thermodynamic T_m [16,17]. The dependence of thermal stability on heating rate may provide an explanation for the unexpected occasional appearance of droplets at temperatures below the T_m (Fig. 5B). Uneven heat absorption or minor unintended sample heating could lead to local micro-unfolding of some collagen helices which nucleate further unfolding of nearby helices. Local unfolding of one collagen trimer then cooperatively unfolds neighboring collagen trimers by providing the opportunity for new favorable interactions and consequently lowering the necessary energy barrier to break trimer intra-strand interactions. Small droplet formation below the predicted T_m temperature could be explained by sufficient levels of cooperative unfolding interactions. It leads us to believe micro-unfolding events may precede and induce phase separation (Fig. 7B). It is worth considering complete denaturation of the collagen coiled-coil may not be necessary to initiate coacervation and some secondary structure could remain in these premature droplets at 14°C.
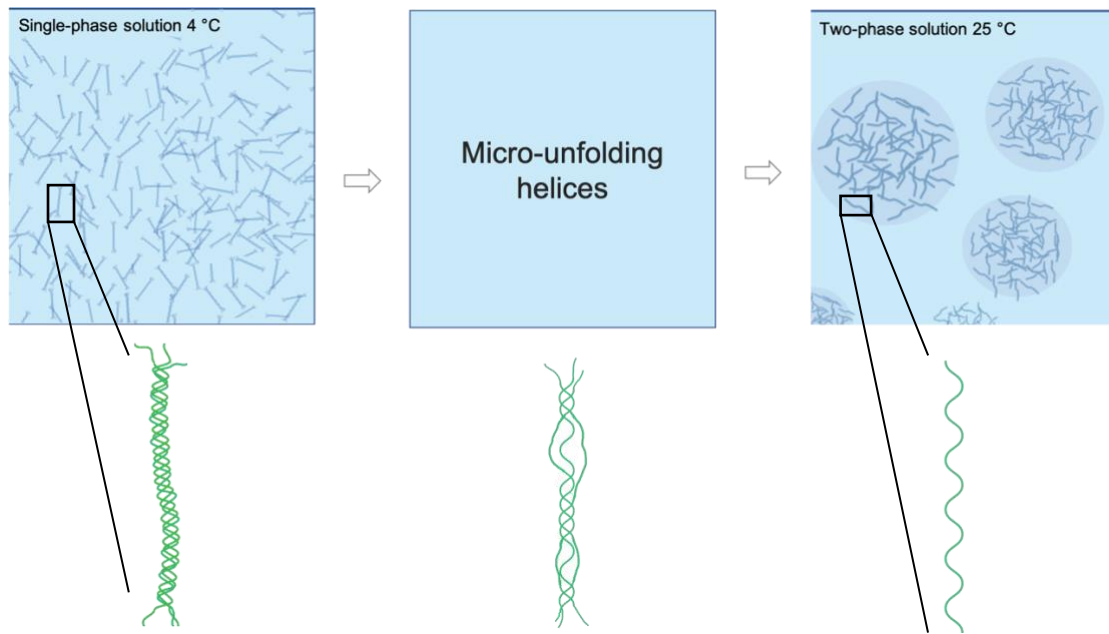
Figure 7. Cartoon model of proposed structures in FSc phase at increasing temperatures. (**A**) flexible collagen trimers in the single-phase solution at 4 ˚C; (**B**) micro-unfolding occurs at intermediate temperature; (**C**) intrinsically disordered gelatin monomers predominate in the LLPS droplets

Fish Scale Collagen Characteristics That May Facilitate LLPS

The flexibility conferred by salmon collagen sequence may be an important component in fish scale gelatin coacervation. Two points that indicate the FSc collagens are more flexible: (1) There are 102 instances of GG repeats total in the salmon collagen trimer compared to only 13 instances in the bovine collagen trimer. Additionally, there are five more instances of GGG triplets in the salmon collagen trimer and a GGGG sequence that is absent in bovine collagen trimer. Glycine, with its hydrogen atom side chain, provides the least steric hindrance for chain mobility. (2) Bovine collagen contains more alanine substitutions in the X and Y positions than salmon collagen, which were shown to result in decreased local flexibility [18]. The increased chain flexibility would make the formation rigid helix-helix interactions entropically unfavorable and as a result the interactions may be more

transient and fluid-like.

Amino acid composition of the FSc extracts exposes another possible factor contributing to coacervate formation. Each extract has consistently higher serine and threonine content when compared to BS collagen (Fig. 2). Elevated serine content is correlated with lower T_m values across several species [19] and in short, synthetic collagen-like molecules [20]. Furthermore, the hydroxylated amino acid side chains may promote interactions between collagen monomers that initiate and/or stabilize the LLPS state once liberated from the coiled-coil conformation. There are additional differences in compositional preference for certain residues (such as Asn, Met, His, and Glu), but these biases are either very subtle, or are not conserved across all tested fish species. It is therefore it is unclear how they may contribute to LLPS, unless the specific order of residues in the X and Y position of GXY triplets contributes to LLPS state formation, much as they contribute to triple helix stability in model peptides [2]. Further sequence analysis for presence of charge or steric biases for the X and Y positions in FSc collagen sequences may provide preliminary insights.

The charge of collagen α1 and α2 monomers changes with the pH of solvent. In acidic conditions, the proteins have a net positive surface charge due to their high isoelectric points (pI) (Table 1). The AU-gel protein band migration toward the negative electrode (Fig. 1) confirms the predicted net positive protein charge at acidic pH for both BS and FSc collagen. Collagen molecules in solution remain soluble and stable in acidic solutions at dilute concentrations due to the charge repulsion between positive residues, therefore the soluble protein is often stored and sold in 0.01 M HCl or 0.833 M acetic acid. However, the predicted repulsion from net positive charge can be overcome by attractive forces at sufficiently high protein concentrations via molecular crowding [21, 22]. The association of

24

collagen molecules at high ionic strengths also suggests that hydrophobic and non-ionic forces play a role in association and assembly, and not electrostatic forces alone [23]. This is particularly well exemplified by the coacervation of mfp1 (pI 10.5) driven by cation-π interactions [24]. It is possible the dense packing in concentrated FSc gelatin solutions results in eviction of water from the protein by overcoming the predicted charge repulsion via molecular crowding and attractive forces. Water depletion has been shown to stabilize entropically driven LLPS in small molecules with predicted charge repulsion[25] and dramatically increase coacervate yield in polyelectrolyte solutions [26].

The low T_m temperature of FSc collagen may have an adaptive value in the living fish. Collagen hydroxyproline content is directly related to physiological temperature of the organism in which the collagen is found [27,28], acting as a fine-tuning mechanism for structure, thermal stability, and mechanical properties. The T_m temperature of collagen molecules is most often near physiological temperature, which leads to regional micro-unfolding of the coiled coil that is believed to be important for collagen strand maintenance, turnover and protein binding [16,29]. The decreased collagen T_m value in cold-blooded organisms, such as fish, allows sufficient molecule flexibility and allows for collagen fibril maintenance in tissue and scales. It is not clear if LLPS of purified fish scale collagen is biologically relevant to the fish scale or if it is an unintended result of the changes in collagen sequence for thermal compatibility with cold environments. Transient increase in body temperature has been reported in fast swimming fish at maximum sustainable swimming speeds [30,31], this could result in thermal damage to collagens that has to be repaired. The LLPS of unfolded collagen molecules could lend itself to that repair if physiological scale conditions permit phase separation.

Conclusion

This work provides an account of the temperature-induced coacervation phenomenon from a simple, scalable fish scale extraction. Compositional and structural analysis of purified fish scale extracts indicate that it is composed of native type I collagen molecules in the single-phase state at 4 ˚C which denature at coacervate-permitting temperatures. These results imply the potential existence of novel collagen phases where flexible gelatin molecules assemble into metastable fluid aggregates rather than solid precipitates. Necessary additional analysis includes the following:

(1)   Amino acid analysis of partially hydrolyzed of FSc collagen (2hrs in 6N HCl at 100 ˚C) to screen for presence of phosphorylation of serine and threonine residues [32,33]. Detection of phosphoserine and phosphothreonine residues in FSc collagen could provide alternative insights to the mechanism of FSc collagen coacervation.

(2)   Screen for presence of free phosphate ions using phosphorus and phosphate detection test strips (Bartovation) sensitive to 1 ppm.

(3)   Quantify diffusion of small droplets within multi-phase coacervates using differential dynamic microscopy (DDM) [34]

Species:	Salmo salar		Bos taurus	
Monomer and	α1 (162-1210)	α2 (73-1112)	α1 (163-1217)	α2 (80-1117)
# residues	1049	1040	1055	1038
Predicted MW (no PTM)	94523.32	92614.96	94544.85	93415.33
Predicted pI (no PTM)	9.53	9.93	9.28	9.93
Total - charge (Asp + Glu)	78	69	83	69
Total + charge (Arg + Lys)	89	87	91	86
Top 5 residue % composition	Gly 36.3% Pro 16.9% Ala 9.4% Ser 5.9% Arg 5.1%	Gly 35.8% Pro 16.3% Ala 10.7% Arg 5.3% Ser 4.7%	Gly 32.8% Pro 22.8% Ala 11.7% Arg 5% Ser 3.6%	Gly 33.6% Pro 20% Ala 10.2% Arg 5.3% Ser 3.4%

Table 1. Uniprot monomer sequence length and position in full procollagen sequence, molecular weights (MW) not including post-translational modifications (e.g. hydroxyproline), isoelectric point (pI), total number of negative and positive side chain residues, and 5 most common residues in sequence [ref].
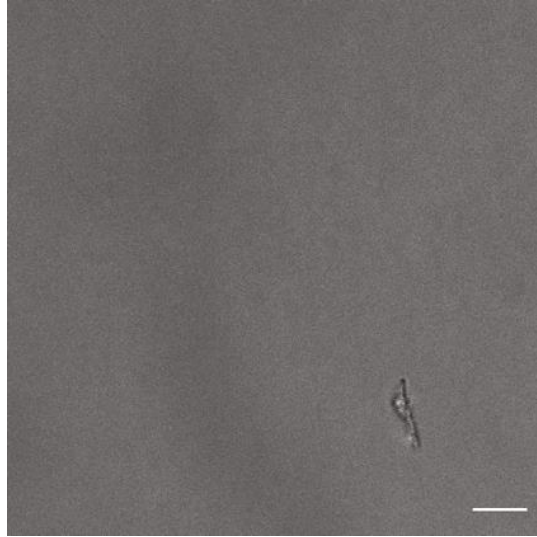
Figure S1. BS collagen solution at 38 ˚C. 10 µm scale bar
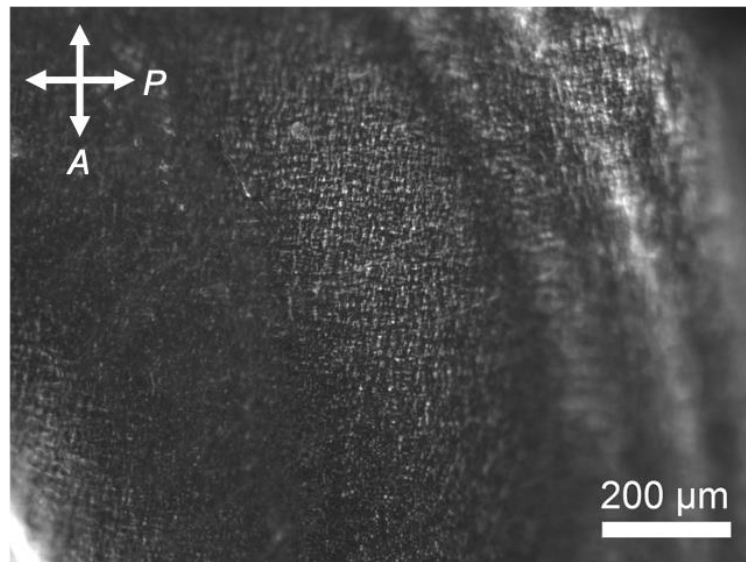
Figure S2. Polarized light microscopy of <50 mg/mL BS collagen solution in 10mM HCl at 25 ˚C shows birefringent patterns in solution indicating organized structures that are light polarizing. Direction of cross polarizers are indicated by white arrows.

References

(1)   The UniProt Consortium. UniProt: The Universal Protein Knowledgebase in 2021. *Nucleic Acids Res.* **2021**, *49* (D1), D480–D489. https://doi.org/10.1093/nar/gkaa1100.

(2)   Persikov, A. V.; Ramshaw, J. A. M.; Brodsky, B. Prediction of Collagen Stability from Amino Acid Sequence. *J. Biol. Chem.* **2005**, *280* (19), 19343–19349. https://doi.org/10.1074/jbc.M501657200.

(3)   Waterborg, J. H. Acetic Acid-Urea Polyacrylamide Gel Electrophoresis of Basic Proteins. In *The Protein Protocols Handbook*; Walker, J. M., Ed.; Springer Protocols Handbooks; Humana Press: Totowa, NJ, 2002; pp 103–111. https://doi.org/10.1385/1-59259-169-8:103.

(4)   Feng, Y.; Melacini, G.; Taulane, J. P.; Goodman, M. Acetyl-Terminated and Template-Assembled Collagen-Based Polypeptides Composed of Gly-Pro-Hyp Sequences. 2. Synthesis and Conformational Analysis by Circular Dichroism, Ultraviolet Absorbance, and Optical Rotation. *J. Am. Chem. Soc.* **1996**, *118* (43), 10351–10358. https://doi.org/10.1021/ja961260c.

(5)   Chan, V. C.; Ramshaw, J. A. M.; Kirkpatrick, A.; Beck, K.; Brodsky, B. Positional Preferences of Ionizable Residues in Gly-X-YTriplets of the Collagen Triple-Helix. *J. Biol. Chem.* **1997**, *272* (50), 31441–31446. https://doi.org/10.1074/jbc.272.50.31441.

(6)   Engel, J.; Chen, H.-T.; Prockop, D. J.; Klump, H. The Triple Helix ⇌ Coil Conversion of Collagen-like Polytripeptides in Aqueous and Nonaqueous Solvents. Comparison of the Thermodynamic Parameters and the Binding of Water to (L-Pro-L-Pro-Gly)n and (L-Pro-L-Hyp-Gly)n. *Biopolymers* **1977**, *16* (3), 601–622. https://doi.org/10.1002/bip.1977.360160310.

(7)   Crocker, J. C.; Grier, D. G. Methods of Digital Video Microscopy for Colloidal Studies. *J. Colloid Interface Sci.* **1996**, *179* (1), 298–310. https://doi.org/10.1006/jcis.1996.0217.

(8)   Moreno, H. M.; Montero, M. P.; Gómez-Guillén, M. C.; Fernández-Martín, F.; Mørkøre, T.; Borderías, J. Collagen Characteristics of Farmed Atlantic Salmon with Firm and Soft Fillet Texture. *Food Chem.* **2012**, *134* (2), 678–685. https://doi.org/10.1016/j.foodchem.2012.02.160.

(9)   Booij, H. L.; Bungenberg de Jong, H. G. *Biocolloids and their Interactions*; Springer Vienna: Vienna, 1956. https://doi.org/10.1007/978-3-7091-5456-4.

(10)   Pei, Y.; Zheng, Y.; Li, Z.; Liu, J.; Zheng, X.; Tang, K.; Kaplan, D. L. Ethanol-Induced Coacervation in Aqueous Gelatin Solution for Constructing Nanospheres and Networks: Morphology, Dynamics and Thermal Sensitivity. *J. Colloid Interface Sci.* **2021**, *582*, 610–618. https://doi.org/10.1016/j.jcis.2020.08.068.

(11)   Martin, E. W.; Mittag, T. Relationship of Sequence and Phase Separation in Protein Low-Complexity Regions. *Biochemistry* **2018**, *57* (17), 2478–2487. https://doi.org/10.1021/acs.biochem.8b00008.

(12)   McGlynn, J. A.; Wu, N.; Schultz, K. M. Multiple Particle Tracking Microrheological Characterization: Fundamentals, Emerging Techniques and Applications. *J. Appl. Phys.* **2020**, *127* (20), 201101. https://doi.org/10.1063/5.0006122.

(13)   Kim, H.; Jang, J.; Park, J.; Lee, K.-P.; Lee, S.; Lee, D.-M.; Kim, K. H.; Kim, H. K.; Cho, D.-W. Shear-Induced Alignment of Collagen Fibrils Using 3D Cell Printing for

Corneal Stroma Tissue Engineering. *Biofabrication* **2019**, *11* (3), 035017. https://doi.org/10.1088/1758-5090/ab1a8b.

(14) Malladi, S.; Miranda-Nieves, D.; Leng, L.; Grainger, S. J.; Tarabanis, C.; Nesmith, A. P.; Kosaraju, R.; Haller, C. A.; Parker, K. K.; Chaikof, E. L.; Günther, A. Continuous Formation of Ultrathin, Strong Collagen Sheets with Tunable Anisotropy and Compaction. *ACS Biomater. Sci. Eng.* **2020**, *6* (7), 4236–4246. https://doi.org/10.1021/acsbiomaterials.0c00321.

(15) Cheng, L.-C.; Hsiao, L. C.; Doyle, P. S. Multiple Particle Tracking Study of Thermally-Gelling Nanoemulsions. *Soft Matter* **2017**, *13* (37), 6606–6619. https://doi.org/10.1039/C7SM01191A.

(16) Leikina, E.; Mertts, M. V.; Kuznetsova, N.; Leikin, S. Type I Collagen Is Thermally Unstable at Body Temperature. *Proc. Natl. Acad. Sci.* **2002**, *99* (3), 1314–1318. https://doi.org/10.1073/pnas.032307099.

(17) Persikov, A. V. Equilibrium Thermal Transitions of Collagen Model Peptides. *Protein Sci.* **2004**, *13* (4), 893–902. https://doi.org/10.1110/ps.03501704.

(18) Chow, W. Y.; Forman, C. J.; Bihan, D.; Puszkarska, A. M.; Rajan, R.; Reid, D. G.; Slatter, D. A.; Colwell, L. J.; Wales, D. J.; Farndale, R. W.; Duer, M. J. Proline Provides Site-Specific Flexibility for in Vivo Collagen. *Sci. Rep.* **2018**, *8* (1), 13809. https://doi.org/10.1038/s41598-018-31937-x.

(19) Rigby, B. J. Correlation between Serine and Thermal Stability of Collagen. *Nature* **1967**, *214* (5083), 87–88. https://doi.org/10.1038/214087a0.

(20) Shu, F.; Dai, C.; Wang, H.; Xu, C.; Wie, B.; Zhang, J.; Xu, Y.; He, L.; Li, S. Formation, Stability and Self-Assembly Behaviour of the Collagen-Like Triple Helix Confirmation: The Role of Ser, Ala and Arg/Glu. *ChemistrySelect* **2019**, *4* (45), 13370–13379. https://doi.org/10.1002/slct.201903500.

(21) Gobeaux, F.; Mosser, G.; Anglo, A.; Panine, P.; Davidson, P.; Giraud-Guille, M.-M.; Belamie, E. Fibrillogenesis in Dense Collagen Solutions: A Physicochemical Study. *J. Mol. Biol.* **2008**, *376* (5), 1509–1522. https://doi.org/10.1016/j.jmb.2007.12.047.

(22) Gobeaux, F.; Belamie, E.; Mosser, G.; Davidson, P.; Panine, P.; Giraud-Guille, M.-M. Cooperative Ordering of Collagen Triple Helices in the Dense State. *Langmuir* **2007**, *23* (11), 6411–6417. https://doi.org/10.1021/la070093z.

(23) Ripamonti, A.; Roveri, N.; Braga, D.; Hulmes, D. J. S.; Miller, A.; Timmins, P. A. Effects of PH and Ionic Strength on the Structure of Collagen Fibrils. *Biopolymers* **1980**, *19* (5), 965–975. https://doi.org/10.1002/bip.1980.360190503.

(24) Kim, S.; Yoo, H. Y.; Huang, J.; Lee, Y.; Park, S.; Park, Y.; Jin, S.; Jung, Y. M.; Zeng, H.; Hwang, D. S.; Jho, Y. Salt Triggers the Simple Coacervation of an Underwater Adhesive When Cations Meet Aromatic π Electrons in Seawater. *ACS Nano* **2017**, *11* (7), 6764–6772. https://doi.org/10.1021/acsnano.7b01370.

(25) Straub, J. S.; Nowotarskib, M. S.; Jiaqi Luc; Tanvi Shethd; Matthew P.A. Fishera; Matthew E. Helgesond; Alexej Jerschowc; Songi Han. Phosphates Form Spectroscopically Dark State Assemblies in Common Aqueous Solutions. 2021.

(26) Park, S.; Barnes, R.; Lin, Y.; Jeon, B.; Najafi, S.; Delaney, K. T.; Fredrickson, G. H.; Shea, J.-E.; Hwang, D. S.; Han, S. Dehydration Entropy Drives Liquid-Liquid Phase Separation by Molecular Crowding. *Commun. Chem.* **2020**, *3* (1), 83. https://doi.org/10.1038/s42004-020-0328-8.

(27) Thengiz V. Burjanadze. New Analysis of the Phylogenetic Change of Collagen Thermostability. *Biopolymers* **2000**, *53*, 523–528.

(28) Gauza-Włodarczyk, M.; Kubisz, L.; Włodarczyk, D. Amino Acid Composition in Determination of Collagen Origin and Assessment of Physical Factors Effects. *Int. J. Biol. Macromol.* **2017**, *104*, 987–991. https://doi.org/10.1016/j.ijbiomac.2017.07.013.

(29) Persikov, A. V.; Brodsky, B. Unstable Molecules Form Stable Tissues. *Proc. Natl. Acad. Sci.* **2002**, *99* (3), 1101–1103. https://doi.org/10.1073/pnas.042707899.

(30) Sepulveda, C.; Dickson, K. A. Maximum Sustainable Speeds and Cost of Swimming in Juvenile Kawakawa Tuna (Euthynnus Affinis) and Chub Mackerel (Scomber Japonicus). *J. Exp. Biol.* **2000**, *203* (20), 3089–3101. https://doi.org/10.1242/jeb.203.20.3089.

(31) Barrett, I.; Hester, F. J. Body Temperature of Yellowfin and Skipjack Tunas in Relation to Sea Surface Temperature. *Nature* **1964**, *203* (4940), 96–97. https://doi.org/10.1038/203096b0.

(32) Cox, K. Analysis of Phosphorylated Amino Acids Using the Hitachi L-8900 Amino Acid Analyzer. 1.

(33) Allerton, S. E.; Perlmann, G. E. Chemical Characterization of the Phosphoprotein Phosvitin. *J. Biol. Chem.* **1965**, *240* (10), 3892–3898. https://doi.org/10.1016/S0021-9258(18)97126-7.

(34) Bayles, A. V.; Squires, T. M.; Helgeson, M. E. Probe Microrheology without Particle Tracking by Differential Dynamic Microscopy. *Rheol. Acta* **2017**, *56* (11), 863–869. https://doi.org/10.1007/s00397-017-1047-7.