Evaluation of the Vitamin K Dependence of Human Osteocalcin Metal Binding and Self-Assembly Properties

by

Stephanie Michelle Thibert

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved July 2021 by the
Graduate Supervisory Committee:

Chad R. Borges, Chair
Joshua LaBaer
Po-Lin Chiu

ARIZONA STATE UNIVERSITY

August 2021

ABSTRACT

Osteocalcin (Oc) is the most abundant non-collagen protein found in the bone, but its precise function is still not completely understood. Three glutamic acid (Glu) residues within its sequence are sites for vitamin K-dependent post-translational modification, replacing a hydrogen with a carboxylate located at the γ-carbon position, converting these to γ-carboxyglutamic acid (Gla) residues. This modification confers increased binding of Oc to Ca^{2+} and hydroxyapatite matrix. Presented here, novel metal binding partners Mn^{2+}, Fe^{3+}, and Cr^{3+} of human Oc were determined, while the previously identified binders to (generally) non-human Oc, Ca^{2+}, Mg^{2+}, Pb^{2+} and Al^{3+} were validated as binders to human Oc by direct infusion mass spectrometry with all metals binding with higher affinity to the post-translationally modified form (Gla-Oc) compared to the unmodified form (Glu-Oc). Oc was also found to form pentamer (Gla-Oc) and pentamer and tetramer (Glu-Oc) homomeric self-assemblies in the absence of NaCl, which disassembled to monomers in the presence of near physiological Na^{+} concentrations. Additionally, Oc was found to form filamentous structures *in vitro* by negative stain TEM in the presence of increased Ca^{2+} titrations in a Gla- and pH-dependent manner. Finally, by combining circular dichroism spectroscopy to determine the fraction of Gla-Oc bound, and inductively-coupled plasma mass spectrometry to quantify total Al concentrations, the data were fit to a single-site binding model and the equilibrium dissociation constant for Al^{3+} binding to human Gla-Oc was determined (K_d = 1.0 ± 0.12 nM). Including citrate, a known competitive binder of Al^{3+}, maintained Al in solution and enabled calculation of free Al^{3+} concentrations using a Matlab script to solve the complex set of linear equations. To further improve Al solubility limits, the pH of the system was lowered to 4.5, the pH during bone resorption. Complementary binding experiments with Glu-Oc were not possible due to the observed precipitation of Glu-Oc at pH 4.5, although qualitatively if Glu-Oc binds Al^{3+}, it is with

much lower affinity compared to Gla-Oc. Taken together, the results presented here further support the importance of post-translational modification, and thus adequate nutritional intake of vitamin K, on the binding and self-assembly properties of human Oc.

DEDICATION

To my mom, Ilona, my greatest female role model who taught me that hard work and perseverance will ultimately lead to success.

To my dad, Robert, who gave me every opportunity for success and encouraged me to follow my dreams, even if it meant moving thousands of miles away.

To Robayet, for teaching me the value of balance in life.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Dr. Chad Borges, for his exceptional guidance throughout my PhD career. It is because of his knowledge, expertise, and willingness to allow me to spread my "research wings" that the work presented here was even possible.

I would also like to thank my committee members, Dr. Joshua LaBaer and Dr. Po-Lin Chium, who gave valuable feedback on my research throughout my PhD.

I would like to thank the School of Molecular Sciences for the financial support I have received over the years, and the diversity in research, courses, and individuals I have had the pleasure of interacting with. Thank you to the Center for Personalized Diagnostics at the Biodesign Institute at Arizona State University for giving me the opportunity to interact and collaborate with stellar scientists in state-of-the-art facilities. I would also like to thank my many collaborators, including, but not limited to, Dr. Stephen Romaniello, who saw my potential long before I did, and Dr. Trevor Martin and Dr. Gwyneth Gordon from the METALs facility at ASU, Chloe Truong and Kazi Waheeda from Dr. Chiu's lab for their countless hours of both experimental and emotional support, Dr. Guanhong Bu and Dr. Brent Nannenga for their help with SEC FPLC work, and Dr. Megan Maurer and Dr. Timothy Karr from the MS Core facility at the Biodesign Institute for their help with SEC HPLC work.

I would also like to thank my lab members, both past and present, including Dr. Shadi Ferdosi, Dr. Yueming Hu, Dr. Joshua Jeffs, Dr. Jesús Aguilar Diaz De Leon, Erandi Kapuruge, Nilojan Jehanathan, and Dr. Stephen Rogers for their guidance and support.

I would like to thank Dr. Robayet Chowdhury for his love and support throughout my PhD, and for listening to my 1 am hypotheses.

Finally, I would like to thank my parents for introducing me to science and encouraging me as I focused on my career.

TABLE OF CONTENTS

Page

CHAPTER Page

CHAPTER Page

APPENDIX Page

LIST OF TABLES

Table Page

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

LIST OF ABBREVIATIONS

1-Gla-Oc	Gla-Oc modified at one of three glutamic acid residues that are sites for post-translational modification to form a γ-carboxy glutamic acid residue
2-Gla-Oc	Gla-Oc modified at two of three glutamic acid residues that are sites for post-translational modification to form γ-carboxy glutamic acid residues
2x 3-Gla-Oc	3-Gla-Oc homodimer
2x Gla-Oc	Gla-Oc dimer
3-Gla-Oc	Gla-Oc modified at all three glutamic acid residues that are sites for post-translational modification to form γ-carboxy glutamic acid residues
ACP	Amorphous calcium phosphate
ADH	Adipic acid dihydrazide
AGE	Advanced glycation endproduct
Al	Aluminum
AlOc	Al^{3+}-bound osteocalcin (1:1 binding)
ATP	Adenosine triphosphate
β-TCP	β-tricalcium phosphate
BGP	Bone Gla-protein
BMD	Bone mineral density
BRCA1	Breast cancer type 1 susceptibility protein
BRCT	BRCA1 C-terminal
BSA	Bovine serum albumin
cAl	Total aluminum
CAS	Chrome azurol S
cCit	Total citrate concentration
CD	Circular dichroism

CKD	Chronic kidney disease
cOc	Total osteocalcin concentration
DCPD	Dicalcium hydrogen phosphate dehydrate
ECF	Extracellular fluid
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ESI	Electrospray ionization
FPLC	Fast protein liquid chromatography
FT-IRM	Fourier transform infrared microspectroscopy
GI	Gastrointestinal
Gla	γ-carboxy glutamic acid
Gla-Oc	Osteocalcin modified at one, two or all three glutamic acid residues that are sites for post-translational modification, forming γ-carboxy glutamic acid residues
Glu	Glutamic acid
Glu-Oc	Osteocalcin containing unmodified glutamic acid residues at the three sites that can be post-translationally modified to form γ-carboxy glutamic acid residues
GPRC6A	G protein-coupled receptor family C group 6 member A
HBS	HEPES buffered saline
HBS-N	HEPES buffered NaCl
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HexNAc-Hex-NANA	N-acetylglucosamine-hexose-N-acetylneuraminic acid
HPLC	High performance liquid chromatography
HSA	Human serum albumin
ICP-MS	Inductively-coupled plasma mass spectrometry
IGFR	Insulin-like growth factor receptor

Ins-Ins	Insulin homodimer
IR	Infrared radiation
JNK	c-Jun N-terminal kinase
LC	Liquid chromatography
MRE	Mean residue ellipticity
MS	Mass spectrometry
MSIA	Mass spectrometric immunoassay
ND	Not detected
NLR	Non-linear regression
NMR	Nuclear magnetic resonance
Nnt	Nicotinamide nucleotide transhydrogenase
Oc	Osteocalcin
OCP	Octacalcium phosphate
Opn	Osteopontin
PAGE	Polyacrylamide gel electrophoresis
pAsp	Polyaspartic acid
PDH	Pimelic acid dihydrazide
PEEK	Polyetheretherketone
PTH	Parathyroid hormone
qTOF	Quadrupole time-of-flight
SEC	Size exclusion chromatography
siRNA	Small interfering RNA
SPE	Solid-phase extraction
TEM	Transmission electron microscopy

TFA	Trifluoroacetic acid
UV-vis	Ultraviolet-visible
vBMD	Volume bone mineral density
WT	Wild type

CHAPTER 1

INTRODUCTION

Osteocalcin (Oc) is the most abundant non-collagen protein found in the bone, making up 2 – 5 % of the total protein mass in bones [1]. It is found primarily on the outer surface of the bone, the periosteum and endosteum, where it binds to hydroxyapatite matrix in a γ-carboxyglutamic acid (Gla)-dependent fashion. The following sections will focus on the sequence, post-translational modification, structure, and function of Oc.

1.1. Sequence & Post-Translational Modification of Osteocalcin

Osteocalcin (Oc) is encoded by the single *BGLAP* gene in humans, and *Bglap*, *Bglap2*, and *Bglap3* genes in mice. *Bglap* and *Bglap2* gene expression occurs exclusively in osteoblasts within the bone, while the *Bglap3* gene expression occurs in tissues outside of the bone, including the kidneys, lungs, and male gonadal tissues [2, 3]. Oc is expressed by osteoblasts in the bone as a 100 amino acid pre-pro-osteocalcin protein that includes a 23 amino acid residue signal peptide that is cleaved after translocation to the endoplasmic reticulum. The 28 amino acid residues that comprise the pro-peptide include the recognition sequence for the vitamin K-dependent γ-glutamyl carboxylase, an enzyme present in the endoplasmic reticulum of osteoblasts that converts glutamic acid (Glu) residues to γ-carboxyglutamic acid (Gla) residues by replacing one of the hydrogens at the γ-carbon position with a carboxylate. The pro-peptide sequence is cleaved by the proprotein convertase furin in the trans-Golgi network [4], and the mature Oc protein is released into the extracellular matrix. The mature protein is 46 – 50 amino acids in length, depending on the species, and ~ 6 kDa in size. The sequence is highly conserved among vertebrates (**Figure 1.1**), in particular the region of the sequence that includes the three Glu residues that are sites for post-translational modification to Gla residues, positions Glu-17, Glu-21, and Glu-24 in the human Oc sequence. The amino acid sequence of the

mature protein is also identical in humans and Neanderthals, as determined by a combination of Edman sequencing and MALDI-TOF/TOF of Oc samples from Neanderthal specimens dating to approximately 75, 000 years old [5]. Although most mammals almost fully γ-carboxylate their Oc, human Oc circulating in the blood is only ~ 50 % γ-carboxylated, with circulating Oc detected as the uncarboxylated form, Glu-Oc, and with 1 – 3 of the Glu residues that are sites for modification being converted to Gla residues (1-, 2-, and 3-Gla-Oc, or collectively referred to herein as Gla-Oc) [6]. Modification of these Glu residues to Gla residues confers increased binding of Oc to Ca^{2+} and hydroxyapatite bone matrix [7].

In most mammal species, one of the prolines in the mature Oc sequence, Pro-9, is also a site for hydroxylation by prolyl-4-hydroxylase to form 4-hydroxyproline [5]. This reaction is vitamin C-dependent, with the enzyme requiring sufficient vitamin C, O_2, Fe^{2+}, and 2-oxoglutarate for Pro-9 modification to occur [5]. The recognition sequence for prolyl-4-hydroxylase is Leu-Gly-Ala-Pro-9-Ala-Pro-Tyr, and five species have the Pro-9-Val-10 sequence modification that prevents, or reduces in the case of gorillas, hydroxylation. The species that do not hydroxylate their Pro-9 are modern humans, Neanderthals, chimpanzees, and orangutan [5], and this is hypothesized by the authors to have occurred as a result of selective pressure due to a shift from an herbivorous to a more omnivorous diet that resulted in periods of low vitamin C intake.

Mouse Oc is unique because it includes a serine residue, Ser-8, that is a site for *O*-glycosylation, with primary mouse osteoblasts found to express several N-acetylgalactosaminyltransferases, with highest expression of *Galnt1* and *Galnt2* [8]. Glycosylation of Ser-8 was not found to be affected by furin activity or γ-carboxylation of mouse Oc. Additionally, disruption of *O*-glycosylation by GalNAc-bn treatment, which inhibits incorporation of glucosamine into *O*-glycans, or by introducing a Ser-8-Ala

mutation, did not prevent pro-osteocalcin protein processing by furin or the ability of mouse Oc to be γ-carboxylated in primary osteoblasts and HEK293 cells [8]. HCD MS/MS of Oc purified from mouse bone homogenate revealed the Ser-8 on 3-Gla-Oc to be modified with an N-acetylglucosamine-hexose-N-acetylneuraminic acid (HexNAc-Hex-NANA) glycan [8]. Glycosylation of Ser-8 was found to confer an increased half-life of mouse Oc in plasma *ex vivo* by preventing plasmin-mediated proteolytic degradation, with the half-life of non-glycosylated Glu-Oc being ~ 1.8 hrs and the glycosylated Glu-Oc half-life increased to ~ 3 hrs [8]. Although human Oc is not glycosylated, Al Rifai, *et al* [8] created a single Tyr-12-Ser mutation in human Oc and found that this mutation conferred human Oc's ability to be glycosylated, and that this glycosylation of human Glu-Oc increased the protein's stability, with the non-glycosylated human Glu-Oc found to have a half-life of ~ 3 hrs, and the glycosylated human Glu-Oc to remain stable throughout the course of the 5 hr experiment [8]. Thus, the authors concluded that the 5 – 10 times higher levels of circulating Oc found in 1- to 6-month old mice, when serum Oc concentrations are highest in mice, compared to young or adult humans may be explained by the unique ability of mouse Oc to be glycosylated, and thus reduce proteolytic degradation [8].

Although human Oc is not normally glycated, the elevated levels of glucose present in patients with Type-2 diabetes can result in non-enzymatic glycation of extracellular matrix proteins, resulting in the formation of advanced glycation end products (AGEs) [9-11]. Glycated forms of Oc have been observed in human and bovine bones with increased levels of glycation reported with increased age [12]. *In vitro* glycation of Oc identified the N-terminus of Oc as the site for glycation and this modification, being far from the Gla residues, was not thought to affect binding of Gla-Oc to hydroxyapatite matrix but was thought to affect the binding of Oc to type I collagen and osteopontin (Opn) [13]. However,

molecular dynamics simulations have suggested that glycation of the N-terminus of Oc results in reduced α-helical structure and reduced binding affinity of Oc for hydroxyapatite matrix, suggesting that glycation of Oc may play a role in the increased fragility of bones in patients with Type-2 diabetes [14].

1.2. Vitamin K, Bone Mineral Density & Fracture Risk

There are three main types of vitamin K, all sharing the 2-methyl-1,4-naphthoquinone (menadione) structure, and containing a variable side chain located at the 3-position [15, 16]. Based on the differences between the variable side chain, vitamin K compounds can be further subdivided into three groups: vitamin K1, vitamin K2, and vitamin K3, also referred to as phylloquinone, menaquinone, and menadione, respectively. Vitamin K1, primarily found in green leafy vegetables, is the major source of vitamin K in Western diets [17], while vitamin K2, except the MK-4 subtype, is synthesized by bacteria present in the human gut [18] and fermented foods [19]. The role of vitamin K in bone mineral density (BMD) is still controversial, with inconsistent results demonstrating no effect [20, 21], an age-independent association in women only [22], or an association in both men and women, with low dietary intake of vitamin K associated with decreased BMD in the femoral neck of men and in the spine of women without estrogen therapy [23]. In these studies, vitamin K intake was monitored by food questionnaires, which may not accurately reflect dietary levels of vitamin K consumed. Additionally, dietary consumption may not accurately reflect serum concentrations of vitamin K. All studies focused only on vitamin K1, and although this is the primary vitamin K source in our diet, vitamin K2 intake may also have affected the results and therefore should have been accounted for [19]. The studies also did not take into account confounding variables, such as the increase in other nutrients acquired from consumption of vitamin K rich foods, such as calcium and magnesium, that may also have an impact on BMD [19]. In contrast, the link between

vitamin K and fracture risk is well documented, with higher levels of vitamin K intake associated with reduced risk of hip fracture [24, 25]. Although BMD was once used as a measure for bone strength, other factors are now also considered as determinants of bone quality, including bone morphology, composition, and biophysical properties, such as hydroxyapatite crystal size [26]. Thus, other factors may be involved in the improvement in bone mineral quality than just increased mineralization.

1.3. Structure of Osteocalcin

The structure of Oc has been studied using circular dichroism (CD) spectroscopy, nuclear magnetic resonance (NMR), and X-ray crystallography. CD analysis of the secondary structure of chicken Oc found that increased titrations of Ca^{2+} caused a shift in the secondary structure from primarily random coil in the absence of Ca^{2+} (8 % α-helix, 14 % β-sheet, and 78 % random coil), to 38 % α-helical structure with a 5 mM Ca^{2+} titration (8 % β-sheet and 45 % random coil) [27]. Decarboxylated chicken Oc was found to adopt higher α-helical structure compared to the native chicken Oc in the absence of Ca^{2+}, with 18 % α-helical structure in the absence of Ca^{2+} and elicited a similar, yet more modest, shift to 26 % α-helical structure in the presence of 5 mM Ca^{2+} [27]. Similar results were observed with Oc derived from other species, with murine Glu-Oc, and human Gla-Oc, having an increase in α-helical structure with increased titrations of Ca^{2+} [28]. From these titration data, equilibrium dissociation constants were determined for murine Glu-Oc binding to Ca^{2+} (K_d = 3.64 mM) and human Gla-Oc binding to Ca^{2+} (K_d = 115 μM) [28]. Other metals that are present in the bone have been found to bind to Oc, with a similar α-helical structural shift in Oc observed with Pb^{2+} [29] and Mg^{2+} [30] binding. Thermal denaturation studies show an increase in stability of murine Glu-Oc in the presence of the saturating concentration of 30 mM Ca^{2+}, compared to stability of murine Glu-Oc in the absence of added Ca^{2+} [28]. The melting temperature for Glu-Oc in the absence of Ca^{2+} is

T_m = 34 °C, which is below the normal physiological temperature of 37 °C, while the melting temperature increases by approximately 12 °C in the presence of 30 mM Ca^{2+} to T_m = 45.4 °C [28]. For human Gla-Oc, the melting temperature in the absence of Ca^{2+} was determined to be T_m = 48 °C and to increase by 12 °C in the presence of saturating concentrations of Ca^{2+}, 5 mM, to give a melting temperature of T_m = 60 °C [28]. Since the Glu-Oc and Gla-Oc used in these studies were acquired from different source species, comparison between the stability of Glu-Oc and Gla-Oc is not possible from this data. Since physiological levels of free Ca^{2+} in the extracellular fluid (ECF) are between 1.1 – 1.4 mM [31], and based on the binding affinity of Glu-Oc and Gla-Oc for Ca^{2+}, these results suggest that Glu-Oc is largely unbound to Ca^{2+}, while Gla-Oc is primarily bound to Ca^{2+} under these conditions [28].

The X-ray crystal structure of porcine Gla-Oc bound to 5 Ca^{2+} was solved at 2.0 Å resolution [32]. The structure revealed the protein to adopt a tight, globular structure, with the presence of three α-helices, with Gla-17, Gla-21, and Gla-24 present on the first α-helix, and a disulfide bridge between Cys-23 and Cys-29 [32]. The tightly packed core of the protein is highly hydrophobic and includes the conserved residues Leu-16, Leu-32, Phe-38, Ala-41, Tyr-42, Phe-45, and Tyr-46 [32]. Together with the Gla residues, and the Asp-30 residue present on the second α-helix, coordination of up to 5 Ca^{2+} is stabilized with coordination occurring with a spatial arrangement of Ca^{2+} that closely matches that within the structure of crystalline hydroxyapatite matrix [32]. A similar structure was obtained from bovine Gla-Oc using ^1H 2D NMR, depicting three α-helices, with all Gla residues present on the first α-helix, with 3 Ca^{2+} projected onto the structure using the genetic algorithm-molecular dynamics simulation approach [33]. The X-ray crystal structure of bovine Glu-Oc was identified at 1.88 Å in the absence of bound Ca^{2+} and depicts a similarly structured globular protein with three α-helices [34]. Amino acid

positions 1 – 15 were found to be largely unstructured for all the identified structures of Oc.

The structure of porcine Gla-Oc was solved with the 5 Ca^{2+} sandwiched between two monomers of Gla-Oc coordinating with the Ca^{2+} ions [32]. Similar dimeric species were observed for bovine Gla-Oc, with the dimer interface occurring at the location of the Gla residues [33]. Sedimentation equilibrium experiments further confirmed the presence of the dimer species, with increased affinity for porcine Gla-Oc dimer formation in the presence of 10 mM Ca^{2+} (K_d = 200 μM) compared to porcine Gla-Oc dimer formation in the absence of Ca^{2+} (K_d = 800 μM) [32]. Sedimentation equilibrium experiments of bovine Gla-Oc with 6 mM Ca^{2+} were used to determine the equilibrium dissociation constant for dimerization in the presence of Ca^{2+} to be K_d = 160 μM [33]. Although the bovine Glu-Oc crystal structure was solved with two Glu-Oc per asymmetric unit, Glu-Oc was not predicted to support dimer formation based on analysis of the limited subunit interface [34], and complementary sedimentation experiments were not conducted to confirm this prediction.

1.4. Localization in the Bone

As mentioned previously, Oc is found primarily on the outer surface of the bone, on the periosteum and endosteum. Immunohistochemistry studies have shown Oc to be found located in regions of turkey leg tendons that had, or were currently, undergoing mineralization [35] and stronger labelling was observed in rat dentin than in the predentin-dentin junction of rat teeth, with an absence of immunoreactivity in osteoid or predentin [36, 37]. Similarly, only weak immunolabelling was observed in chicken tibiae in the unmineralized osteoid regions [38]. Mineralization on collagen fibrils occurs with an initial amorphous calcium phosphate (ACP) phase that localizes at the gap zones within the a-band region of the collagen fibril [39]. Studies *in vitro* with polyaspartic acid (pAsp)

and fetuin revealed their co-localization with ACP, forming a highly negatively charged complex [39]. Localization of the pAsp-ACP complex at the gap zones of collagen is highly favourable due to the high positive net charge of this region of the collagen fibril, thus offering the site with the lowest electrostatic potential energy for interaction with the negatively charged pAsp-ACP complex [39]. Similar results were observed for Oc in turkey leg tendons, where Oc was found to localize primarily at the interfaces between the overlap and gap zones of collagen fibrils, including the a_4 – a_1 bands [35]. Osteocalcin monomer, being just under 6 kDa in size, was also found in the intrafibrillar collagen spaces [35], where only molecules under 6 kDa in size can fit within the spaces in the collagen fibril [40].

1.5. Circulation in Blood Plasma and Serum and Disease

Oc has also been found in blood circulation and was used as a biomarker for bone turnover due to the observed increase in serum concentrations during periods of increased bone resorption [41] and mineralization [42]. Normal circulating Oc concentrations range from 11.3 – 25.6 ng/mL (~ 2 – 4 nM) [43] and include Glu-Oc, 1-, 2-, and 3-Gla-Oc, and over a dozen forms of Oc fragments [6]. Increased vitamin K intake was shown to result in the increased fractional abundance of 3-Gla-Oc and a reduced fractional abundance of Glu-Oc [6]. Excretion of Oc is primarily accomplished by glomerular filtration in the kidneys, with the half-life of Oc in the blood of rats being only 4 minutes [44], and increased human plasma total Oc concentrations observed in patients with chronic kidney disease (CKD) compared to normal healthy controls [45]. During a 4-week study, participants were either selected to consume a diet rich in vitamin K leafy green vegetables, a diet with low vitamin K vegetables, or a diet without added vegetables with an energy matched control [46]. Participants who ate the diet including vegetables rich in vitamin K had lower total circulating Oc and under carboxylated Oc (Glu-Oc & 1- & 2-Gla-Oc), and higher absolute

3-Gla-Oc concentrations compared to the start of the study and the energy matched control group after the 4-week period [46]. The decrease in total circulating Oc concentrations was hypothesized to be due to an increase in the binding of Oc to the bone due to increased affinity of 3-Gla-Oc for hydroxyapatite matrix [46].

Some results have demonstrated that circulating total Oc levels are higher in women more than two years into menopause [47] and elevated serum total Oc levels were associated with lower BMD in postmenopausal women with osteoporosis compared to non-osteoporotic postmenopausal women [48]. However, conflicting results from a meta-analysis found no difference in the serum total Oc levels in postmenopausal women with osteoporosis compared to postmenopausal women without osteoporosis [49], and thus further work should be conducted before assigning serum total Oc levels as a marker for bone turnover in postmenopausal women with osteoporosis. In patients with metabolic syndrome, lower levels of under carboxylated Oc was shown to be a marker for increased risk of cardiovascular disease and Type-2 diabetes [50].

1.6. Function of Osteocalcin

1.6.1. Function of Osteocalcin in the Bone

Oc was first discovered by two independent groups in the mid 1970's [51, 52] as a γ-carboxyglutamic acid (Gla) containing bone protein and was referred to as Bone Gla Protein (BGP). Although many decades have passed since its first discovery, the function of Oc is still not completely understood. Initial studies with $Oc^{-/-}$ mouse models revealed that $Oc^{-/-}$ mice had higher BMD compared to wild type controls [53]. Additionally, Oc caused dose-dependent chemotaxis, adhesion, and spreading of human osteoclast-like cells, and caused the release of bone sialoprotein, Opn, and fibronectin *in vitro* [54]. Combined with the findings that human and rat osteoclasts caused release of both intact and fragmented Oc from bovine bone slices during bone resorption *in vitro* [55] and that

patients with uremia [56] or induced hypocalcemia [57] had increased serum Oc concentrations, this led to the conclusion that Oc functions in bone resorption.

Although the results mentioned above suggest Oc functions in bone resorption, conflicting results emerged demonstrating Oc may play a role in bone mineralization/remodeling processes. Analysis of cortical and trabecular bones from $Oc^{-/-}$ mice by Fourier transform infrared microspectroscopy (FT-IRM) revealed their cortical bones to have impaired maturation compared to WT control mice, with WT mice having higher carbonate:phosphate ratios and larger, more perfect crystals in their cortical bones compared to the $Oc^{-/-}$ mice [58]. There was no observed difference in the trabecular bones of $Oc^{-/-}$ compared to WT mice [58], which is consistent with the finding that Oc is localized in higher abundance in cortical bone compared to trabecular bone [59, 60]. Differences in the mechanical and chemical properties were found in cortical bones of $Oc^{-/-}$ mice compared to WT mice, with a decrease in the hardness and type-B carbonate substitutions (when CO_3^{2-} substitutes for the PO_4^{3-} ion in the hydroxyapatite crystal) in $Oc^{-/-}$ mice [61], further suggesting Oc plays a role in the maturation of cortical bones. To study the effects of Oc on bone healing, cylindrical hydroxyapatite/collagen implants were inserted into the tibial head of adult rats, with one group of rats also getting 10 μg/g Oc, which was added to the implant during the setting process, and the healing process was monitored [62]. Rats that also had Oc added to their implants had higher numbers of phagocytosing cells and osteoclasts in early stages of bone healing, and earlier and higher expression of Opn, bone sialoprotein, and CD44 at the sites of the implants compared with rats that did not receive Oc in their implants [62]. Newly formed woven bone appeared at the implant surface and was partially replaced by lamellar bone faster in rats with implants with, compared to those without, Oc added, further suggesting that Oc plays a role in bone remodelling [62].

Oc has also been shown to interact with Opn [63, 64], with formation of a dilatational band complex shown to help with energy dissipation during microfractures in the bone [65]. Double knockout $Oc^{-/-}$ $Opn^{-/-}$ mice have bones that are shorter, have larger cross-sectional areas, and larger cortical areas than single knockout $Oc^{-/-}$, $Opn^{-/-}$, or WT mice [66]. Using small angle X-ray scattering and wavelength dispersive spectroscopy, the physical structure and mineral composition of bones was analyzed in more depth, revealing that both Oc and Opn are involved in regulating crystal size and the organization of the crystals, with $Oc^{-/-}$, $Opn^{-/-}$, and $Oc^{-/-}$ $Opn^{-/-}$ mice displaying reduced crystal thickness and a higher degree of misalignment of the crystals compared to WT mice [67]. Studies conducted by Moriishi, *et al* [68] found that $Oc^{-/-}$ mice had normal mineralization hydroxyapatite crystal size compared to WT mice, but that the orientation of hydroxyapatite on the collagen fibrils in the $Oc^{-/-}$ mice was altered, with the orientation of the crystals no longer running parallel to the collagen fibers. This suggested that Oc plays a role in the proper alignment of hydroxyapatite crystals onto collagen fibrils.

The formation of hydroxyapatite matrix requires stabilization of several different mineral species during osteogenic differentiation, some of which can be monitored using Raman spectroscopy [69]. The initial mineral species present in osteoblasts and the extracellular matrix is ACP, as observed by IR analysis of rat bone [70]. Other species that have been observed are dicalcium hydrogen phosphate dehydrate (DCPD), octacalcium phosphate (OCP), β-tricalcium phosphate (β-TCP), all of which can be monitored by Raman spectroscopy [69]. Using siRNA to knockdown Oc gene expression in human mesenchymal stromal cells, the role of Oc in the mineralization process was examined by monitoring the presence of the different mineral species over time and comparing these results to human mesenchymal stromal cells in the sham control group [71]. From these experiments it was determined that Oc knockdown resulted in postponed observation of

OCP, then the β-TCP species was prolonged, and less hydroxyapatite species was formed and appeared later than the control group [71]. In line with these findings, Oc was found to stabilize OCP in the vertebrae of rats and human femur, as well as *in vitro*, using high-resolution TEM, and was found to form "pearl necklace strings" that are used as nucleation sites for OCP on the collagen fibrils [72]. A low molecular weight polyacrylic acid was used to simulate Oc and cross-linked to collagen to study intrafibrillar mineralization in an *in vitro* system [73]. Combined with molecular dynamics simulations, it was found that the cross-linked polyacrylic acid formed chain-like aggregates on the surface of the collagen fibril as a mechanism to increase mineral species that were available for intrafibrillar mineralization [73]. Taken together with the previous knowledge that Oc can localize in the intrafibrillar collagen spaces [34], and form "pearl necklace strings" [72], these results are highly suggestive that Oc may function to collect mineralization components and stabilize precursor mineral species in a site-specific fashion.

1.6.2. Endocrine Function of Osteocalcin

Much of the research focused on Oc, specifically Glu-Oc, functioning as an endocrine hormone was conducted by the Karsenty group using $Oc^{-/-}$ mice derived from the 129Sv;C57BL/6J strain, with key findings of increased blood glucose, impaired glucose and insulin tolerance, reduced β cell mass in the pancreas and increased fat mass compared to wild type mice [74]. Additionally, Glu-Oc was reported to be a ligand for the G protein-coupled receptor family C group 6 member A (GPRC6A) in β cells in the pancreas and Leydig cells in the testes [75, 76]. Mouse models with the $GPRC6A^{-/-}$ phenotype were shown to have impaired glucose metabolism and reduced testosterone levels compared to the wild type mice [77]. However, attempts by other groups to replicate these findings have been unsuccessful, demonstrating no difference in the blood glucose

levels, glucose tolerance, and fat mass in $Oc^{-/-}$ mice [78], $Oc^{-/-}$ rats [79], and glucose metabolism, fertility, and BMD in $GPRC6A^{-/-}$ mice [80-82] compared to wild type controls. Inspection of the mouse strain used by the Karsenty group revealed there to be a mutation in the nicotinamide nucleotide transhydrogenase (Nnt) gene that was absent in the mouse strains used by others [83]. Nnt is known to influence β cell insulin secretion in the pancreas [84, 85], and thus loss of expression of Nnt is more likely to be the explanation for the observed effects on glucose metabolism. As expected, transgenic expression of Nnt resulted in reversal of the observed decrease in insulin secretion and glucose intolerance [84-88], further supporting this notion. A few studies have been conducted in human cell lines *in vitro*, with one demonstrating that in MG-63 human osteoblast-like cells, insulin and low glucose resulted in increased Oc gene expression, while insulin resistance and increased glucose concentrations caused a decrease in Oc gene expression [89]. However, the low reproducibility of the results supporting Oc's role in glucose metabolism conducted using the 129Sv;C57BL/6J mouse strain suggests, at the very least, that further work should be conducted before Oc's role can be confirmed. It is also for this reason that, although studies have not yet, to the best of our knowledge, confirmed or denied a possible role for Oc in brain function, since this work was also conducted by the Karsenty group utilizing this same 129Sv;C57BL/6J mouse strain [90], the relevance of Oc's role in these functions should be taken with caution. Recently, $Oc^{-/-}$ mice were found to have abnormally aligned hydroxyapatite crystals on the collagen fibers and reduced bone strength, but no difference in glucose metabolism, serum testosterone levels, or muscle mass compared to WT controls [68]. For this reason, as well as its minimal relevance to the remainder of this dissertation, discussion of Oc functioning as an endocrine hormone, and the breadth of literature published on this topic, will not be discussed further here.

1.7. Research Scope

To this end, the goal of the work discussed in the proceeding chapters was to assess the metal binding capabilities and structural properties of human Oc, parsing between Gla-Oc and Glu-Oc forms. The work presented improves the knowledge on the vitamin K dependent specificity of Oc binding metals and demonstrated the *in vitro* capability of human Gla-Oc to form filaments in the presence of increased Ca^{2+} concentrations. Finally, a method to quantify the equilibrium dissociation constant for Al^{3+} binding to 3-Gla-Oc was developed. Taken together, the work discussed here improves the knowledge of the structural and binding properties of Oc and, ultimately, will aid in further understanding of the physiological function of this highly abundant protein in humans.

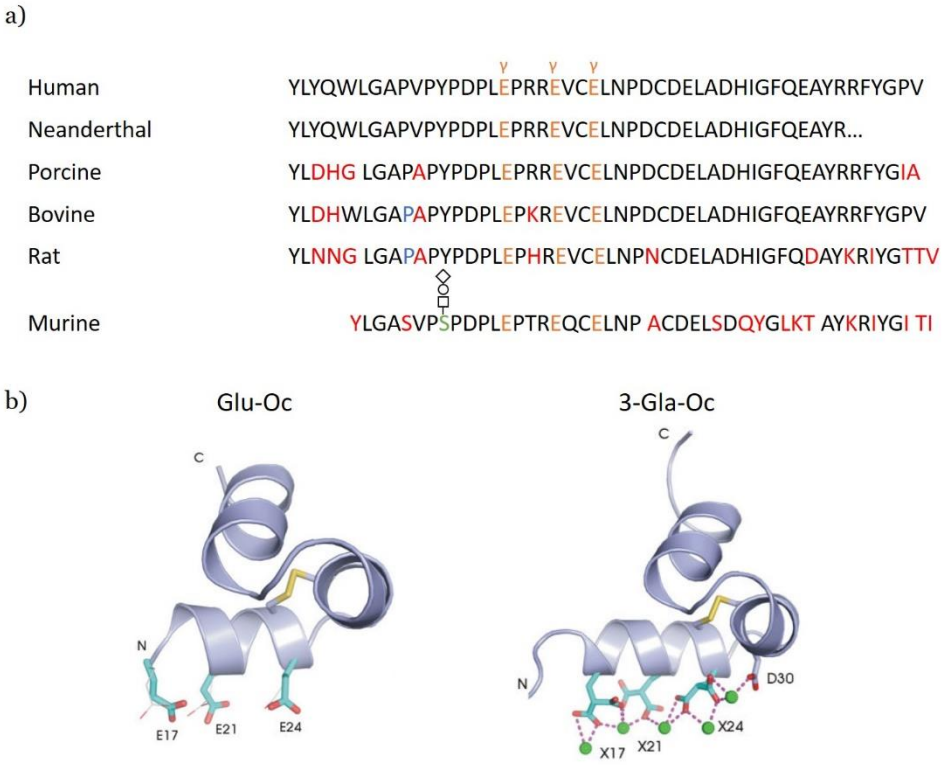


Figure 1.1. Sequence and Structure of Osteocalcin. a) Amino acid sequence of modern human, Neanderthal, porcine, bovine, rat, and murine osteocalcin (Oc). The glutamic acid (Glu, E) residues that are sites for post-translational modification to γ-carboxyglutamic acid (Gla) resides are highlighted in orange in each sequence. Residues that differ from the human sequence are highlighted in red in the other species' sequences. Species that have Pro-9 that can be post-translationally modified to 4-hydroxyproline are highlighted in blue. Ser-8 (highlighted in green) in the murine sequence can be *O*-glycosylated, with the N-acetylglucosamine-hexose-N-acetylneuraminic acid (HexNAc-Hex-NANA) glycan depicted. b) The X-ray crystal structure of bovine Glu-Oc (left, PDB 4MZZ) depicts Glu-17 (E17), Glu-21 (E21), and Glu-24 (E24) on the first α-helix, and a disulfide bridge between Cys-23 and Cys-29 depicted in yellow. The X-ray crystal structure of porcine 3-Gla-Oc (right, PDB 1Q8H) depicts Gla-17 (X17), Gla-21 (X21), and Gla-24 (X24) on the first α-helix together with Asp-30 (D30) coordinating with 5 Ca^{2+} (green spheres), and a disulfide bridge between Cys-23 and Cys-29. Reprinted (adapted) with permission from Malashkevich V N, Almo S C & Dowd T L. (2013). X-ray crystal structure of bovine 3 Glu-Osteocalcin. *Biochemistry*. **52**, 8387-92 [34]. Copyright 2013 American Chemical Society. See **Appendix A** for full copy of permission.

CHAPTER 2

ASSESSMENT OF OSTEOCALCIN'S POST-TRANSLATIONAL MODIFICATION DEPENDENT METAL-BINDING BY NATIVE MASS SPECTROMETRY

2.1 Introduction

Metals play important roles in both biochemical and biophysical processes throughout the body, with the selection of these metals taking place over billions of years of evolution. The driving force behind selection of each metal depends on both the physicochemical properties and the bioavailability of the metals. These metals often coordinate with proteins, with metal-binding proteins performing several structural and functional roles throughout the body. A unique function of some protein-metal complexes is in the formation of mineral crystals, known as biomineralization. Ferritin is one example in this class of metal-binding proteins, where the biomineralization of hydrated ferric oxides is tightly controlled by the nanocage structure of the protein complex. Osteocalcin (Oc) is another protein that may function in the process of biomineralization of hydroxyapatite, as recent literature suggests osteocalcin stabilizes octacalcium phosphate (OCP), a calcium phosphate crystal structure precursor to hydroxyapatite bone matrix [72].

Oc was discovered in the 1970s [51, 52] as a bone protein and was referred to as bone Gla-protein (BGP) due to the presence of γ-carboxyglutamic acid (Gla) residues within its amino acid sequence, a rare post-translational modification that confers binding of up to 5 Ca^{2+} per Oc monomer [32]. Binding of Ca^{2+} by Oc occurs by coordination of the Ca^{2+} ions with the carboxyl groups of three glutamic acid residues, located at positions Glu-17, Glu-21, and Glu-24 in the human Oc sequence, with increased binding affinity when these Glu residues are converted to Gla residues (**Figure 1.1b**). These Glu residues are converted to Gla residues in a post-translational modification dependent fashion via a γ-glutamyl carboxylase that modifies these Glu residues by replacing one of the hydrogens

at the γ-carbon position with a carboxylate. This modification is dependent on the reduced form of vitamin-K, and unlike most mammals who almost fully γ-carboxylate their Oc, only about 50 % of Oc in the blood circulation of humans has been found to be fully γ-carboxylated [6]. Purification of Oc from a cohort of human patient plasma samples has shown that Oc circulates in patients' blood with fractions of the Oc uncarboxylated (Glu-Oc), or containing one (1-Gla-Oc), two (2-Gla-Oc), or all three (3-Gla-Oc) glutamic acid residues γ-carboxylated [6]. The degree of γ-carboxylation of Oc was found to be dependent on the individual's vitamin K status, with vitamin K supplementation shifting about 20 % of the Glu-Oc to 3-Gla-Oc [6], highlighting the potential importance of adequate nutritional intake of this vitamin on bone health.

Post-translational modification of these sites confers increased affinity for Ca^{2+}, with human 3-Gla-Oc binding with a dissociation constant of ~ 0.1 mM, compared to ~ 3.6 mM for mouse Glu-Oc [28]. Although this study did not compare Gla-Oc and Glu-Oc binding within the same species, this, to the best of our knowledge, is the best comparison that has been made to date comparing binding of Ca^{2+} to the different γ-carboxylation states of Oc. There are multiple characteristics of a metal that govern metal-ligand binding including the metal's valence state, ionic radius, charge-accepting ability, and the number of metal-bound atoms present in a ligand [91]. Thus, when Oc is post-translationally modified to contain these rare Gla residues, this increases the number of atoms that can bind Ca^{2+} within each Gla residue compared to Glu residues, known as denticity [91]. Since it is more favourable for Ca^{2+} to bind to one of the oxygen atoms of each carboxyl group in Gla residues than with a pair of Glu residues [91], this modification is critical for effective coordination of Oc with Ca^{2+}, and therefore the protein's function within the bone. This is known as chelation bidentate mode, where both carboxylates from the same Gla residue are monodentately bound [91].

Although Oc's main target binder is thought to be Ca^{2+}, several other metals have also been reported to bind to the protein, including two toxic metals, Pb^{2+} and Al^{3+}. Bone is the primary storage site for these toxic metals, with approximately 95 % and 50 % of Pb^{2+} [92] and Al^{3+} [93] found within the body stored there, respectively. Binding of Pb^{2+} to bovine Oc can displace Ca^{2+} [94], with an equilibrium dissociation constant for Pb^{2+} binding bovine Oc being K_d = 85 nM [29]. Similarly, Al^{3+} binding to bovine Oc has been reported to have an equilibrium dissociation constant of K_d = 83 nM [95]. When Pb^{2+} was titrated with Oc, this resulted in the same increase in α-helical secondary structure as observed with increased titrations of Ca^{2+} with Oc and did not affect the protein's ability to associate with hydroxyapatite crystal [29]. The Ca^{2+} sites within hydroxyapatite crystal can be substituted with Pb^{2+} [96], resulting in decreased bone mineral density (BMD) and fracture resistance [97]. Similar studies assessing the secondary structural effects and hydroxyapatite binding capability of Al^{3+} bound Oc have not been completed. Hydroxyapatite crystal growth has been shown to be altered in low doses and inhibited in high doses of Al^{3+} *in vitro*, with resultant incorporation of Al^{3+} causing structural changes to the growing mineral deposits compared to Al^{3+} free controls [98]. Acute Al^{3+} exposure due to its presence as a contaminant in dialysis solution presented as osteomalacia and dialysis encephalopathy in hemodialysis patients in the 1970's and 1980's [99].

Other metals that have been studied as they relate to bone health and Oc are essential metal cations that are found within the bone, including Mn^{2+}, Mg^{2+}, Fe^{3+}, and Cr^{3+}. Optimal levels of these metals are vital for bone growth, oxygen transport, and carbohydrate and lipid metabolism [100], however toxicity and/or deficiency can result in abnormalities in bone physiology. With about 25 g of magnesium present in the average human body, half of that is stored within the bone and has been shown to be essential for bone health, with Mg^{2+} deficiency being associated with decreased bone mineralization

and bone fracture resistance, resulting in osteopenia and osteoporosis [100]. In contrast to Pb^{2+}, binding of Mg^{2+} to Oc inhibits Oc's ability to bind to hydroxyapatite matrix, although the shift to increased α-helical structure was still observed [30]. Supplementation of Mn^{2+} in the diet of rats was shown to improve BMD and increase their serum levels of Oc compared to controls [101]. Deficiency of Mn^{2+}, although rare, can cause skeletal abnormalities, whereas toxicity leads to a breadth of effects on proper bone development including osteogenesis, bone resorption, and osteoporosis [100]. Although, to the best of our knowledge, the effects of Fe^{3+} on Oc have not been studied in the past, Fe^{3+} toxicity is associated with decreased BMD, decreased bone fracture resistance, osteoporosis and osteopenia, as well as changes in bone microarchitecture [102], and inhibition of hydroxyapatite crystal formation [103]. Present in metal implants used during surgery, Cr^{3+} can leach out over time due to corrosion, leading to toxicity [100]. Rats fed a diet that contained added chromium were found to have increased bone resorption and reduced bone formation [100].

With all of these metals being found at least to some degree within the bone, and evidence to support their negative effect on bone health when toxic and/or deficient levels are present, the goal of this research was two-fold: i) To determine metal binding capability of ***human*** Oc for Mn^{2+}, Fe^{3+} and Cr^{3+}, and validate binding of previously identified binders Ca^{2+} (shown previously to bind to bovine [29] and human Oc [28]) and Pb^{2+} [29], Al^{3+} [95], and Mg^{2+} [30] (shown previously to bind to bovine Oc); and ii) To perform a qualitative assessment of the Gla-dependency of metal binding by comparing binding of each metal to 1-, 2-, and 3-Gla-Oc to its binding capability to Glu-Oc. The novelty of this research lies in the fact that up to this point, most research has focused on metal binding to Oc purified from species other than humans, and the majority of these studies have not parsed Gla-Oc and Glu-Oc binding. This distinction is important when

studying human Oc because the health status (vitamin K intake) of an individual dictates the levels of Gla-Oc and Glu-Oc present at any given time [6]. The results presented here are valuable pilot-study data that may be used to initiate future, more in-depth research on the Gla-dependent binding of Oc to metals.

2.2 Materials and Methods

2.2.1 Materials

Both the Des-Gla synthetic human osteocalcin (Glu-Oc), Cat # AS-65307-025, and [Gla 17, 21, 24] – synthetic human osteocalcin (Gla-Oc), Cat # AS-22830, were purchased from AnaSpec, Co. Ammonium acetate (≥ 98 % purity), ammonium hydroxide, iron(III) chloride, chromium(III) chloride hexahydrate (≥ 98 % purity), human insulin, magnesium chloride hexahydrate (99.995 % trace metals basis), manganese chloride tetrahydrate, lead(II) nitrate (99.999 % trace metals basis), aluminum chloride hexahydrate (99 % purity), and calcium chloride desiccant (96+ %, ACS reagent) were purchased from Sigma-Aldrich, Co. HPLC grade methanol was purchased from Fisher Scientific, Co.

2.2.2 Preparation of Osteocalcin Stocks

For the Gla-Oc vial, the lyophilized powder was reconstituted by first adding 40 µL of 1 % ammonium hydroxide and the vial was swirled gently to ensure all powder was fully reconstituted prior to adding 5 mM ammonium acetate, pH 7.4, to 1 mL final volume. The pH was verified prior to dilution using short range pH paper. For the Glu-Oc vial, the reconstitution technique was identical except the volume of ammonium hydroxide and final volume were half that of the Gla-Oc vial to account for there being half the weight of lyophilized protein. Stock Oc sample concentrations were then measured at A_{280} using a Thermo Scientific Multiskan Go UV-vis spectrophotometer, with Oc's extinction coefficient known, $\varepsilon = 1.33\ (mg/mL)^{-1}cm^{-1}$ [104].

2.2.3 Preparation of Samples

All samples were prepared with 4 μM Gla-Oc, or Glu-Oc, and 4 μM insulin, in 5 mM ammonium acetate, pH 7.4. Metal was titrated into the sample at concentrations of 0, 0.8, 2, 4, 12, 20 and 40 μM and the samples were incubated at room temperature overnight. After initial sample infusion onto the MS, samples were again incubated for an additional day and re-run to verify protein-metal binding had reached equilibrium. If not, the samples were further incubated at room temperature for up to one week. Samples were centrifuged at 17, 000 x g for 5 minutes prior to infusions.

2.2.4 Direct Infusion ESI-MS

Samples described above were infused directly onto a Bruker MaXis 4G qTOF mass spectrometer at 5 μL/min using a syringe and PEEK tubing line. The syringe and line were rinsed several times in between each sample infusion with 100 % acetonitrile, followed by 5 mM ammonium acetate, pH 7.4. Then the same rinsing process was completed for the ESI spray needle by infusing 50/50 acetonitrile/5 mM ammonium acetate, pH 7.4, followed by 5 mM ammonium acetate, pH 7.4, onto the MS at 10 μL/min. The MS was operated in positive ion mode with a capillary voltage set to 4500 V, end plate offset set to 500 V, nebulizer gas at 3.0 bar, dry gas set to 4.0 L/min, and the source temperature set to 225 °C. Sample data were acquired for 2 minutes for each sample, deconvoluted, baseline subtracted, and analyzed using Bruker Compass 1.7 DataAnalysis 4.2.383.1 software.

2.2.5 Non-Specific Dimer Interactions Calculations

The fractional abundance of insulin, 1-, 2-, and 3-Gla-Oc monomers was calculated based on the intensity counts of each protein monomer's 'x + 1' isotope peak divided by the total

intensity counts for all protein monomers in the charge deconvoluted mass spectral data. These monomer fractional abundance values were then used to calculate the fractional abundance of each possible dimer combination using the binomial expansion described by **Equation 2.1**:

$$(I + G_1 + G_2 + G_3)^2 = 1$$

Which can be expanded to the give:

$$I^2 + G_1^2 + G_2^2 + G_3^2 + 2 * (I * G_1 + I * G_2 + I * G_3 + G_1 * G_2 + G_1 * G_3 + G_2 * G_3) = 1$$

Equation 2.1

Where I is the insulin, G_1 is the 1-Gla-Oc, G_2 is the 2-Gla-Oc, and G_3 is the 3-Gla-Oc monomer fractional abundance. The use of binomial expansion to predict the fractional abundance of dimers relative to each other in mass spectra if the dimers are non-specific and driven by probabilities based on the fractional abundance of their monomer counterparts has been used previously [105] to predict the fractional abundance of oligonucleotide dimers. This work demonstrated that for samples where only nonspecific oligonucleotide dimer formation is expected, the fractional abundance of these dimers determined experimentally closely matched the predicted values. As an example, in a system with A and B proteins with equal fractional abundance of the monomers, A:B is 0.5:0.5, the possible dimers that could form would be AA, AB, and BB, with the fractional abundance of these dimers due to non-specific interaction for AA:AB:BB calculated to be 0.25:0.5:0.25. In contrast, the fractional abundance of dimers determined experimentally for samples containing oligonucleotides known to form specific dimer interactions did not closely match the predicted fractional abundance values calculated for nonspecific

interactions using this binomial expansion calculation approach [105]. Therefore, this approach may be used to distinguish between dimer formation observed in mass spectral data that originates completely from non-specific interactions from dimer formation that is not completely non-specific in origin. Note that this approach cannot decipher dimer interactions in mass spectral data that have mixed dimer origins, meaning data with mixed specific and non-specific interactions from those that are purely specific, and thus can only be used to determine whether all dimerization observed is non-specific, or not. To accommodate these more complex sample compositions, a reference protein should be used to quantitatively account for non-specific interactions in more complex situations [106].

Since three of the possible dimer combinations, 2x 1-Gla-Oc, insulin + 1-Gla-Oc, and 1-Gla-Oc + 2-Gla-Oc, were not observed in the actual mass spectral data, the fractional abundance of the remaining observed dimers was normalized to accommodate for this discrepancy. These calculated fractional abundances were then compared to the measured values based on the mass spectral data, and the difference between the calculated and measured fractional abundance values was evaluated for statistical significance using a paired student t-test, with p-values < 0.01 considered statistically significantly different, implying the assumption that all dimers observed in the spectra are non-specific is invalid.

2.2.6 Insulin Control Experiment

Insulin control samples were prepared at 4 μM, 10 μM, 50 μM, and 100 μM in 5 mM ammonium acetate, pH 7.4, or pH 3.3, and centrifuged, infused onto the mass spectrometer, and analyzed following the same protocol as described in **2.2.4**. For the 100 μM insulin sample in 5 mM ammonium acetate, pH 3.3, all buffer washes described

above were completed at this pH. The fraction of insulin as a dimer was determined by summing the peak intensities for all mass spectral peaks associated with Ins-Ins dimer, including Ins-Ins and Ins-Ins bound to metal adducts, and dividing by the total peak counts for all insulin mass spectral peaks found in the spectra. The experimental data were plotted with the fraction bound (fraction of total insulin that was Ins-Ins dimer) against the free concentrations of insulin. Free insulin concentrations were calculated by summing the peak intensities for all mass spectral peaks associated with insulin monomer, dividing by the total insulin peak counts found in the mass spectra to obtain the fraction of insulin as monomer species, and multiplying by the known concentration of total insulin to obtain the concentration of free insulin. To evaluate whether the data fit well to the previously determined dissociation constant for the Ins-Ins dimer, a binding curve was plotted alongside the experimentally determined data, with the equation for the binding curve described by **Equation 2.3** below, where the dissociation constant for Ins-Ins dimer formation is described by:

$$K_d = \frac{\mathrm{Ins} * \mathrm{Ins}}{\mathrm{Ins\ Ins}}$$

Which can be rearranged to:

$$\mathrm{Ins\ Ins} = \frac{\mathrm{Ins}^2}{K_d} \qquad \textbf{Equation 2.2}$$

The fraction of total insulin bound as dimer is described by:

$$\text{Fraction bound} = \frac{2 * \mathrm{Ins\ Ins}}{\mathrm{Ins} + 2 * \mathrm{Ins\ Ins}}$$

Therefore, **Equation 2.2** can be used to substitute for Ins-Ins in **Equation 2.3**:

$$\text{Fraction bound} = \frac{2*\frac{\text{Ins}^2}{K_d}}{\text{Ins}+2*\frac{\text{Ins}^2}{K_d}}$$

Which can be further simplified to give:

$$\text{Fraction bound} = \frac{2*\text{Ins}^2}{2*\text{Ins}^2+K_d*\text{Ins}} \qquad \textbf{Equation 2.3}$$

Where K_d = 100 μM is the dissociation constant for the Ins-Ins dimer [107] and Ins is the concentration of free insulin. The data were plotted in Matlab using the Matlab script found in **Appendix A**.

2.3 Results

2.3.1 Analysis of Metals Binding to Gla-Oc, Glu-Oc, and Insulin

Direct infusion MS of 4 μM Gla-Oc (**Figure 2.1a**), or 4 μM Glu-Oc (**Figure 2.1b**), with 4 μM insulin (used as a reference protein that was not expected to bind the metal) in the absence of added metal, or with 40 μM of the divalent cations Ca^{2+}, Mg^{2+}, Mn^{2+}, or Pb^{2+} was completed to qualitatively evaluate Oc's ability to bind these metals, and to compare metal binding between the post-translationally modified Gla-Oc to the unmodified form of Oc, Glu-Oc. The fraction of Glu-Oc and 3-Gla-Oc bound to each respective metal is summarized in **Figure 2.1c**. The Gla-Oc sample without metal added contains mass spectral peaks associated with Na^+ and K^+ adducts of 3-Gla-Oc (**Figure 2.1a**), which are common protein salt adducts in ESI-MS. Mass spectral peaks associated with 3-Gla-Oc bound singly to either Ni^{2+} or Zn^{2+} are also observed and labeled in the spectra, and are acquired from Oc interacting with nickel and zinc that are electrochemically oxidized and released from the spray needle assembly during the infusion process. The mass spectrum associated with Gla-Oc with Ca^{2+} added was found to include mass spectral peaks

associated with up to five Ca^{2+} bound to 3-Gla-Oc. With Mg^{2+} or Mn^{2+} added to Gla-Oc, mass spectral peaks associated with up to four Mg^{2+}, or Mn^{2+}, bound to 3-Gla-Oc were observed. The mass spectrum associated with Pb^{2+} added to Gla-Oc shows no mass spectral peaks associated with Pb^{2+} bound to 3-Gla-Oc, however, increased titrations of Pb^{2+} with Gla-Oc resulted in a decrease in Gla-Oc peak signals (**Figure 2.2**). The metal-bound 3-Gla-Oc mass spectral peaks are indicated by red asterisks (*), with the number of asterisks corresponding to the number of the indicated divalent metal cation bound. At lower Pb^{2+} concentrations, mass spectral peaks associated with Pb^{2+}-bound Gla-Oc adducts were observed, but did not exceed ~ 0.2 fraction of 3-Gla-Oc bound to Pb^{2+} in the spectrum, with the 4 μM Pb^{2+} sample spectrum depicted in **Figure 2.3**.

Infusion of the Glu-Oc sample with no metal added onto the MS revealed mass spectral peaks associated with unbound Glu-Oc and insulin monomers, as well as Na^{+} and K^{+} adducts for each of the proteins (**Figure 2.1b**). The Glu-Oc sample containing Ca^{2+} was found to include up to three Ca^{2+} bound to Glu-Oc, while the Mg^{2+} sample included mass spectral peaks associated with up to two Mg^{2+} bound to Glu-Oc. The samples with either Mn^{2+} or Pb^{2+} added did not contain mass spectral peaks associated with Glu-Oc bound to either of these divalent cations. The location in the mass spectra where these peaks would be found are indicated by an arrow with a red 'x' (x). Mass spectral peaks associated with insulin bound to up to three Ca^{2+} and two Mg^{2+} were observed and are indicated by the blue asterisks (*). At this 40 μM metal concentration, no peaks associated with insulin bound to either Mn^{2+} or Pb^{2+} were observed, with the location in the mass spectra where these peaks would be found if present indicated by an arrow with a blue 'x' (x). The insulin peaks associated with the Gla-Oc samples, however, demonstrated up to three Mg^{2+} and up to two Mn^{2+} bound to insulin (**Figure 2.4**). Titrations of Glu-Oc with

increased concentrations of either Ca^{2+}, Mg^{2+}, or Mn^{2+} showed metal-bound Glu-Oc mass spectral peaks (**Figure 2.5**).

Identical infusion data were acquired to assess the binding capability of the trivalent metal cations Fe^{3+}, Cr^{3+}, and Al^{3+} to Gla-Oc, Glu-Oc, and insulin (**Figure 2.6**). Consistent with the Gla-Oc control sample for the divalent metal cation data, the Gla-Oc control sample with no metal added for the trivalent metal cation sample was also found to contain mass spectral peaks associated with Na^{+}, K^{+}, Ni^{2+}, and Zn^{2+} bound to Gla-Oc (**Figure 2.6a**). Each of the Gla-Oc sample spectra obtained from infusion MS of Gla-Oc samples incubated with Fe^{3+}, Cr^{3+}, or Al^{3+} show a mass spectral peak associated with the unbound 3-Gla-Oc, and the singly-bound 3-Gla-Oc metal adduct indicated by a red asterisk (*). The Glu-Oc sample with no metal added shows a spectrum that includes mass spectral peaks associated with unbound Glu-Oc and insulin, as well as peaks corresponding to both protein's bound to either Na^{+} or K^{+} (**Figure 2.6b**). The spectra corresponding to Glu-Oc and insulin incubated with either Fe^{3+}, Cr^{3+}, or Al^{3+} indicate no observed mass spectral peaks corresponding to the mass associated with the tested metal binding to either protein. The expected locations of the metal-bound Glu-Oc (x) and insulin (x) peaks are indicated by an arrow and an 'x'. Mass spectra of Gla-Oc, or Glu-Oc, incubated with 0 μM (**Figure 2.7a**), and up to 40 μM Cr^{3+} are shown in **Figure 2.7**. The measured and calculated size (m/z) of each singly-bound protein-metal adduct, and their m/z difference, are shown in **Table 2.1**, with the mass difference not exceeding 0.0705 m/z.

To further assess the effect of γ-carboxylation of Oc on metal binding, the mass spectral data acquired from infusing 4 μM Gla-Oc, or 4 μM Glu-Oc, with 4 μM insulin and 40 μM of either Ca^{2+}, Mg^{2+}, Mn^{2+}, Pb^{2+}, Fe^{3+}, Cr^{3+}, or Al^{3+} were used to calculate the

fraction of insulin, Glu-Oc, 1-, 2-, and 3-Gla-Oc bound to each metal (**Table 2.2**). At this specific metal concentration, 0.33 of insulin was bound to Ca^{2+}, 0.40 was bound to Mg^{2+}, 0.21 was bound to Mn^{2+}, and 0 was bound to Pb^{2+}, Fe^{3+}, Cr^{3+}, and Al^{3+}. The fractions of Glu-Oc bound to each metal were 0.42 Ca^{2+}, 0.24 Mg^{2+}, 0.05 Mn^{2+}, 0 Fe^{3+}, Al^{3+} and Cr^{3+}, while the peak associated with Pb^{2+} was not detected (ND). The Gla-Oc sample containing Ca^{2+} showed 1-Gla-Oc bound to Ca^{2+} ND, and 0.52 of 2-Gla-Oc, and 0.65 of 3-Gla-Oc bound to Ca^{2+}. Similar to the Ca^{2+} sample, no 1-Gla-Oc mass spectral peaks were observed in the Gla-Oc sample containing Mg^{2+}, and thus the fraction bound to Mg^{2+} could not be determined, however 0.55 of 2-Gla-Oc, and 0.86 of 3-Gla-Oc was bound to Mg^{2+}. The Gla-Oc sample containing Mn^{2+} showed 0.4 of 1-Gla-Oc was bound to Mn^{2+}, 0.75 of 2-Gla-Oc, and 0.97 of 3-Gla-Oc. Fractions of Gla-Oc bound to Pb^{2+} could not be determined. For the Gla-Oc sample containing Fe^{3+}, none of the 1-Gla-Oc, 0.35 of the 2-Gla-Oc, and 0.72 of the 3-Gla-Oc were bound to Fe^{3+}. All the 1-, 2-, and 3-Gla-Oc were found to be bound to Cr^{3+}. Finally, no mass spectral peaks associated with 1-Gla-Oc were detected in the Gla-Oc sample containing Al^{3+}, while 0.55 of the 2-Gla-Oc, and all of the 3-Gla-Oc were found to be bound to Al^{3+}.

2.3.2 Qualitative Binding Affinity Comparison of 1-, 2-, and 3-Gla-Oc to Ca^{2+}, Fe^{3+} and Al^{3+}

To further determine the Gla-dependency of metal binding to Oc, 4 μM Gla-Oc, and Glu-Oc, samples were incubated with 0, 40, 120, 200, and 500 μM Ca^{2+} and infused on the MS. The mass spectral peaks associated with unbound 3-Gla-Oc, and up to 5 Ca^{2+} bound to 3-Gla-Oc are shown in **Figure 2.8a**. The fractions of total Gla-Oc that were either 1-, 2-, or 3-Gla-Oc are shown in **Figure 2.8b**, with over 70 % of the Gla-Oc infused in the fully γ-carboxylated 3-Gla-Oc form. Fractions of 1-, 2-, and 3-Gla-Oc bound to up to

5 Ca^{2+} at each metal concentration tested were calculated, with results plotted against total Ca concentrations in **Figure 2.8c**. The fractions of Gla-Oc bound to one to five Ca^{2+} were summed to simplify the data to compare bound to unbound Gla-Oc fractions more clearly. The data depict an increase in the fraction of Oc bound to Ca^{2+} with increased γ-carboxylation of Oc.

Similarly, 4 μM Gla-Oc samples were incubated with 0, 4, 12, 20, and 40 μM Fe^{3+} and infused on the MS. The mass spectral peaks associated with unbound 3-Gla-Oc and singly-bound to Fe^{3+} are shown in **Figure 2.9a**. The fractions of total Gla-Oc that were 1-, 2-, or 3-Gla-Oc were determined (**Figure 2.9b**), with the majority of Gla-Oc being fully γ-carboxylated 3-Gla-Oc. The fraction of 1-, 2-, and 3-Gla-Oc bound to Fe^{3+} at each metal concentration were calculated and results were plotted against total Fe concentrations (**Figure 2.9c**). In the case of 1-Gla-Oc, no Fe^{3+} binding was observed at the tested metal concentrations. Although the binding curves for all forms of Gla-Oc did not reach saturation at the tested Fe concentrations, there is a clear qualitative trend present in the data, with increased γ-carboxylation of Oc associated with higher affinity of Oc for Fe^{3+}.

Complementary experiments to test the binding of Al^{3+} to the various forms of Gla-Oc were also conducted, with Al^{3+} concentrations of 4, 8, 12, 20, and 40 μM. Mass spectral peaks associated with unbound 3-Gla-Oc and 3-Gla-Oc bound singly to Al^{3+} are shown in **Figure 2.10a**. Determination of the fraction of Gla-Oc that was either 1-, 2-, or 3-Gla-Oc demonstrated that the majority of the sample was in the fully γ-carboxylated 3-Gla-Oc form (**Figure 2.10b**). As completed with the Ca^{2+} and Fe^{3+} samples, the fraction of 1-, 2-, and 3-Gla-Oc bound to Al^{3+} was calculated and plotted against each total concentration of Al (**Figure 2.10c**). Similar to the results presented for Ca^{2+} and Fe^{3+} binding Oc, increased γ-carboxylation of Oc increased the affinity of Al^{3+} for Oc.

2.3.3 Insulin and Oc Dimers

When analyzing the direct infusion MS data described above, dimers of insulin and Oc were observed (**Figure 2.11**). The monomer and dimer charge states of the 4 μM Gla-Oc with 4 μM insulin sample (**Figure 2.11a**) and the 4 μM Glu-Oc with 4 μM insulin sample (**Figure 2.11c**) demonstrate that both samples' base peak is the +5 charge state of the monomers. A small proportion of the samples was found to be in the +7 charge state associated with dimers of the proteins. Deconvoluted mass spectra focusing on the mass spectral peaks associated with various insulin and Gla-Oc dimers (**Figure 2.11b**) and insulin and Glu-Oc dimers (**Figure 2.11d**) are also shown. The measured and calculated monoisotopic masses of all possible tested dimer combinations (**Table 2.3**) shows the difference between the two values does not exceed 0.1445 Da. The dimers that were not detected (ND) in the experimental mass spectral data could not be compared to the calculated monoisotopic masses of the species, and thus the difference (Δ) between these values is left blank (-).

To assess whether all the dimer interactions observed in the mass spectra were a non-specific occurrence, the fractional abundance of each observed dimer was compared to their calculated fractional abundance due to probability alone. The non-specific interaction calculations were completed following the method described in **Methods 2.2.5**. using the binomial expansion that describes the sum of the fractional abundance of all possible dimer species in **Equation 2.1**. This method has been used previously by Ding & Anderegg [105] to assess specific and non-specific dimer formation of oligonucleotides in ESI-MS, where use of the binomial expansion to calculate the fractional abundance of each dimer closely matched the experimentally determined fractional abundance of the dimers in mass spectra from samples with dimers expected only to originate from

non-specific binding during the ionization process. Experimentally determined fractional abundances of dimers from samples where oligonucleotides had some degree of specificity in binding did not closely match to the predicted fractional abundance values calculated using the binomial expansion, suggesting that the dimer formation was not completely non-specific [105]. The motivation of these calculations was to assess whether all the observed dimers in the mass spectra were non-specific in origin, or not. Although this approach has a clear limitation, where it can only determine whether all dimers are non-specifically formed or not, this analysis was conducted to evaluate whether to pursue more meticulous studies in the future. The overall workflow is depicted in **Figure 2.12**, where the fractional abundance of the monomeric species is determined based on the deconvoluted mass spectra. Since three of the possible dimer combinations, 2x 1-Gla-Oc, insulin + 1-Gla-Oc, and 1-Gla-Oc + 2-Gla-Oc, were not observed in the actual mass spectral data, the calculated fractional abundance of the remaining observed dimers was adjusted to accommodate for this discrepancy. These calculated fractional abundances were then compared to the measured values based on the mass spectral data (**Table 2.4**) and evaluated whether they were statistically significantly different from one another, based on three replicate measurements. The p-values < 0.01 are highlighted in red and signify the dimers that have measured fractional abundances that are statistically significantly different from the predicted values that assume their occurrence is based on non-specific interaction alone. The dimers that match this criterion are the insulin + 2-Gla-Oc, insulin + 3-Gla-Oc, 2x 2-Gla-Oc/1-Gla-Oc + 3-Gla-Oc, and the 2x 3-Gla-Oc dimers. This implies that the experimentally observed dimers are not all non-specific in origin.

To assess whether the conditions used for native MS were soft enough to maintain protein-protein interactions, a series of insulin concentrations were infused onto the mass spectrometer (**Figure 2.13a-e**), and the fractional abundance of the monomer and dimer

were calculated from the charge deconvoluted mass spectra by summing the peak intensities for mass spectral peaks associated with monomer, or dimer, and dividing by the sum of the total peak intensities for both species observed in the mass spectra. The fractional abundance of the Ins-Ins dimer observed experimentally was plotted against the concentration of insulin monomer, determined based on the known total insulin concentrations and the fractional abundance of the monomer in each mass spectrum. The data were plotted (**Figure 2.13f**) with the line defined by the equation that describes the fraction of insulin as Ins-Ins dimer, based on the known K_d = 100 μM for Ins-Ins dimer formation [107]. The fraction of insulin that was determined experimentally to be the Ins-Ins dimer did not match with the equation defined by the known K_d for Ins-Ins dimer formation. The results obtained in this control experiment suggest the mass spectra obtained here do not accurately reflect true native (i.e., true in-solution) conditions, with the fractional abundance of dimer when 100 μM insulin was infused onto the mass spectrometer not exceeding 0.052 when buffered at pH 7.4, and 0.073 when buffered at pH 3.3, compared to the expected fractional abundance of 0.5.

2.3.4 Monomer and Dimer Metal-Binding Comparison

To compare the metal-binding capability of the insulin, Glu-Oc, 1-, 2-, and 3-Gla-Oc monomers to the various combinations of observed protein dimers, heat maps were created to summarize the data (**Figure 2.14**), with the dimers not tested (dark grey) being the Glu-Oc + Gla-Oc dimers, not detected (light grey) being the 2x 1-Gla-Oc, insulin + 1-Gla-Oc, and 1-Gla-Oc + 2-Gla-Oc, and observed being the insulin-insulin, insulin + Glu-Oc, insulin + 2-Gla-Oc, insulin + 3-Gla-Oc, 2x Glu-Oc, 2-Gla-Oc + 3-Gla-Oc, 2x 3-Gla-Oc, and 2x 2-Gla-Oc/1-Gla-Oc + 3-Gla-Oc dimers (**Figure 2.14a**). The protein monomer and dimer metal-binding capability was then compared in **Figure 2.14b**, with

the Ca^{2+}, Mg^{2+}, Mn^{2+}, Pb^{2+}, Fe^{3+}, and Al^{3+} data separated. Monomers, and dimer combinations, not tested are labeled in dark grey, not observed in light grey, followed by the maximum number of the specific metal bound to the monomer/dimer coloured with an increasing blue coloured gradient for no metal bound (lightest blue) to up to five metals bound (darkest blue).

The deconvoluted monomer mass spectra of insulin with Gla-Oc and Ca^{2+} titrations (**Figure 2.15a,c,e**) depict mass spectral peaks associated with insulin and Gla-Oc monomers, and the monomers bound to Ca^{2+}. Their respective deconvoluted dimer mass spectra (**Figure 2.15b,d,f**) illustrate the presence of insulin-insulin dimer and Ca^{2+} adducts, and insulin + Gla-Oc dimers and Ca^{2+} adducts. The 2x Gla-Oc dimers, or Ca^{2+}-bound 2x Gla-Oc dimers, are not observed in the mass spectra, with increasing Ca^{2+} concentrations causing increased Ca^{2+} adducts of the insulin-insulin dimer, and the insulin + Gla-Oc dimers, but not resulting in observed 2x Gla-Oc dimers.

Similarly, deconvoluted monomer mass spectra of insulin with Glu-Oc and Ca^{2+} titrations (**Figure 2.16a,c,e**) depict mass spectral peaks associated with insulin and Glu-Oc monomers, and the monomers bound to Ca^{2+}. The deconvoluted dimer mass spectra (**Figure 2.16b,d,f**) include mass spectral peaks associated with insulin-insulin dimer and Ca^{2+} adducts, insulin + Glu-Oc dimer and Ca^{2+} adducts, and low abundance of 2x Glu-Oc + Ca^{2+} adducts, with the peaks only confidently identified in **Figure 2.16b** and **d**.

2.4 Discussion

2.4.1 Analysis of Metals Binding to Gla-Oc, Glu-Oc, and Insulin

Based on the results presented in **Figure 2.1**, it is clear all divalent metals tested are capable of coordinating with Gla-Oc. Whether these are the true maximum numbers of each divalent metal cation that Gla-Oc can bind in solution needs to be assessed further, as these maxima could also be explained by the sensitivity limits of this technique at these particular protein and metal concentrations. It has been shown elsewhere [32] that porcine Oc can coordinate with up to five Ca^{2+} in the X-ray crystal structure obtained, which matches the maximum number of Ca^{2+} bound to Gla-Oc in the mass spectral data presented here. It should also be highlighted that the mass shift for Ca^{2+} (Δ 37.947 Da) cannot be confidently distinguished from that for K^{+} (Δ 37.9559 Da), and therefore, although increasing titrations of Ca^{2+} with each protein can confidently determine the protein's ability to bind Ca^{2+} (as seen from increased peak intensities for the Ca^{2+}-protein adducts), one should proceed with caution when making definitive conclusions about the stoichiometry of this interaction. A similar case is also observed with Mg^{2+}, where the mass shift (Δ 21.9694 Da) is close to the mass shift seen for Na^{+} adducts (21.9819 Da). Since increased metal titrations were found to result in protein and metal precipitation (data not shown), this eliminated the possibility of evaluating these maxima further using direct infusion MS.

In contrast, Gla-Oc was found to bind only singly to the trivalent metal cations Fe^{3+}, Cr^{3+}, and Al^{3+} (**Figure 2.6**). A possible explanation for this phenomenon is that trivalent metal coordination with the Gla residues of Oc requires the use of multiple Gla residues, and thus may require a slight conformational shift of the protein to

accommodate binding such metals. This would need to be evaluated further to make such definitive conclusions however, and thus this is purely speculation at this time.

Each metal was also tested for binding capabilities to Glu-Oc (**Figures 2.1 and 2.6**), and all were found to either bind to a lesser degree than to Gla-Oc, or not found to bind in solution to any detectable degree when 4 μM Glu-Oc was incubated with 40 μM of the metal. When comparing the fraction of each protein bound to the metal cation (**Figure 2.1c**, **Figure 2.6c** & **Table 2.2**), it is clear that increasing the number of Gla residues present in Oc also increases the protein's affinity for the metals, with the fraction of Oc bound to each metal increasing with increased Gla residues. These data match what is expected, since Gla residue ligands are known to be important for protein coordination with metal cations, as they have increased net charge and denticity compared to Glu residue ligands [91]. Therefore, it is expected that inclusion of Gla residues in a protein's structure will increase the protein's binding affinity for metal cations. However, whether or not Glu-Oc can bind to Fe^{3+} or Al^{3+} to any detectable degree would need to be evaluated further, since this data only suggests that Gla-Oc has a higher affinity for the metals than Glu-Oc, but does not definitively demonstrate whether or not Glu-Oc is capable of coordinating with these metal cations at higher metal concentrations.

Further studies would need to be completed to successfully analyze binding of Pb^{2+} to Glu-Oc and Gla-Oc. This most likely would need to be completed at a lower pH, since lead hydroxide is not very soluble at pH 7.4 ($K_{sp} = 1.43 \times 10^{-20}$ for $Pb(OH)_2$ [108]). Attempts to infuse samples at pH 4.5, with and without insulin, resulted in complete loss of signal in the MS spectra due to protein precipitation (data not shown). This is suspected to be due to slight acidification of samples, which is known to occur during the infusion process, even in the presence of ammonium acetate [109] and may have resulted in sample pH

decreasing below the isoelectric point of the proteins (pI = 4.6 for Glu-Oc, pI = 4.0 for Gla-Oc [110], and pI = 5.3 for insulin [111]). For this reason, a technique where the pH of the solution can be carefully controlled, such as the circular dichroism (CD) and ICP-MS technique described in **Chapter 5**, should be employed to alleviate these issues while studying Pb^{2+} binding to proteins. It should also be mentioned that although Al^{3+} has a lower solubility limit than that of Pb^{2+} (K_{sp} = 1.3 x 10^{-33} for $Al(OH)_3$), Al^{3+}-bound Gla-Oc adducts were observed in the mass spectra presented here, suggesting the challenge with detecting Pb^{2+}-bound Gla-Oc at higher than ~ 0.2 fraction bound cannot be completely explained by the low solubility limit of $Pb(OH)_2$.

The quaternary structure of proteins is often stabilized by the coordination of metals at the protein-protein interfaces. An example is insulin, which functions in its monomer form, but is stored in the pancreas as a hexamer stabilized by two Zn^{2+} ions located at each of the protein-protein interfaces of the complex. For an in-depth review of the role of metals in protein-protein interfaces, see [112]. Hexameric insulin was not observed in these native MS data. Insulin binding to Ca^{2+} or Pb^{2+} has been shown previously to induce dimerization [113], and insulin has been shown previously to bind to Mg^{2+} [114], however the physiological relevance of insulin binding to the divalent metal cations Ca^{2+}, Mg^{2+}, Mn^{2+}, and Pb^{2+} is unknown. Indirect interactions have been shown, including the observation that Ca^{2+} influx from the extracellular space, or release from cytosolic stores, causes exocytosis of insulin granules from pancreatic β-cells [115, 116]. Intracellular Mg^{2+} in pancreatic β-cells is involved in regulation of insulin secretion, phosphorylation of the insulin receptor and other downstream kinases, with deficiency associated with type 2 diabetes mellitus [117]. The presence of Mn^{2+} has been shown to be important to enhance insulin/Insulin-like growth factor receptor (IGFR) activity,

resulting in increased induction of phosphorylated AKT, while deficiency impairs insulin release and glucose tolerance [118].

Due to the known binding of insulin to many of the tested divalent cations, the use of insulin as a reference protein to account for non-specific protein-metal adducts that occur during the infusion process was not possible. The absence of trivalent metal cation-bound insulin in the mass spectra suggested that non-specific metal binding did not occur for the trivalent metal cations tested here, and that insulin was a successful reference protein for these samples.

2.4.2 Qualitative Binding Affinity Comparison of 1-, 2-, and 3-Gla-Oc to Ca^{2+}, Fe^{3+} and Al^{3+}

Direct infusions of Gla-Oc with titrations of Ca^{2+} (**Figure 2.8**), Fe^{3+} (**Figure 2.9**) and Al^{3+} (**Figure 2.10**) further demonstrate the Gla-dependence of metal binding, with increased binding affinity found with increased Gla residues in Oc. One limitation in the work described here, as mentioned previously, is the inability to distinguish between Ca^{2+} and K^{+} adducts, which will inevitably result in overestimations during calculation of the fraction of Oc bound to Ca^{2+}. Since Gla-Oc is known to bind up to five Ca^{2+}, the previously reported dissociation constants [28, 119] are also inaccurate, as they assume a 1:1 binding mode for Oc to Ca^{2+}, which at least for porcine and human Oc appears to be an invalid assumption.

Completion of the binding curve for Gla-Oc binding to Fe^{3+} was limited by the protein precipitation observed when exceeding 40 μM Fe^{3+} concentrations, likely due to the low solubility of $Fe(OH)_3$ at pH 7.4. An additional challenge when studying Fe^{3+} binding to Gla-Oc was the presence of contaminating Mn^{2+} in the ESI spray needle. The mass shift for Fe^{3+} binding (Δ 52.9115 Da) and Mn^{2+} binding (Δ 52.9224 Da) cannot be

differentiated in the mass spectra, and thus any variation in Mn^{2+} present cannot be confidently identified based on the limitations of this technique.

Although the curve reached binding saturation for the Al^{3+} data, precipitation of Gla-Oc was also a challenge here. With the Al^{3+} data, this precipitation can at least in part be explained by its low solubility limit at pH 7.4 (K_{sp} = 1.3 x 10^{-33} [108]), where Al^{3+} readily forms $Al(OH)_{3(s)}$, with a calculated solubility limit of ~ 82 fM in pure water at pH 7.4. The presence of acetate, chloride, and Gla-Oc, all binders of Al^{3+}, likely enabled the increased solubility of total Al, allowing for observation of in-solution Al^{3+} bound to Gla-Oc in the infusion MS data. As shown in **Chapter 5**, increased solubility of Al^{3+} was accomplished by decreasing solution pH to 4.5. With the inclusion of citrate, an Al^{3+} binder with known affinity constants, it was possible to increase the solubility limit for total Al. The fraction of Gla-Oc bound was determined using CD spectroscopy. Free Al^{3+} concentrations were calculated using a Matlab script to solve the complex system of linear equations, with experimentally determined total Al concentrations, determined by ICP-MS, known total Oc and total citrate concentrations, solution pH, and the calculated concentrations of Al^{3+}-bound Gla-Oc using the CD data input into the Matlab script. These data are shown in **Chapter 5** and provide an example of an optimized method that can both alleviate metal solubility challenges and provide more precise binding affinity calculations for this exceptional case of Gla-Oc binding to Al^{3+}.

Regardless of the qualitative nature of the binding affinity experiments conducted for Ca^{2+}, Fe^{3+} and Al^{3+} binding to 1-, 2-, and 3-Gla-Oc, these data still demonstrate the importance of Gla residues for increased affinity of Oc to metals, and thus the purity of Gla-Oc samples when used for quantitative studies in the future.

2.4.3 Insulin and Oc Dimers

Protein dimerization is a common occurrence in biology and is important for many proteins that function as enzymes, ion channels, receptors, and transcription factors. Insulin readily forms a homomeric dimer, with dimerization acting as a precursor in the self-assembly of insulin into a hexamer [120]. Although the function of insulin lies in its monomeric state, hexamer formation is an important quaternary structure of the protein as it acts as the structure for storage of the protein in Zn^{2+}-rich, and Ca^{2+}-rich, vesicles in pancreatic β cells [121], controlled release of the monomer, Zn^{2+} homeostasis [122], and prevention of protein degradation and fibrillation [123]. Insulin hexamer coordinates with two Zn^{2+}, with Zn^{2+} binding increasing the assembly of Ins-Ins dimers to form hexamer [124] following the monomer-dimer-tetramer-hexamer model [125]. Although less studied, osteocalcin dimerization has also been shown previously [32], with the K_d = 800 μM for Gla-Oc homodimer (2x Gla-Oc) formation in the absence of Ca^{2+}, and 200 μM in the presence of Ca^{2+}, and the X-ray crystal structure of porcine 2x Gla-Oc dimer complexed with five Ca^{2+} [32]. The biological significance of 2x Gla-Oc dimer is unknown, and this interaction is unlikely to be relevant in the blood, where Oc concentrations are between 11.3 – 25.6 ng/mL (~ 2 – 4 nM) [43]. The concentration of Oc in the bone, however, was estimated to be between ~ 430 μM and 9.85 mM, as described in **Chapter 3**, and therefore it is reasonable to hypothesize 2x Gla-Oc dimerization is a biologically significant occurrence in the bone.

When completing the native MS infusions of insulin with Oc and various metals described in this chapter, mass spectral peaks associated with protein-protein dimerization were observed in both the samples without metal added, and the samples including Ca^{2+}, Mg^{2+}, Mn^{2+}, Pb^{2+}, Fe^{3+}, and Al^{3+}. Ins-Ins [107] and 2x Gla-Oc [126] dimers have been seen in mass spectral data in the past, both in metal-free forms, and bound to

Zn^{2+} (Ins-Ins) and Ca^{2+} (2x Gla-Oc). Unexpectedly, mass spectral data presented here demonstrate the presence of both homomeric dimers reported previously, and insulin and Oc heterodimers.

Non-specific protein-metal, and protein-protein, interactions are common in native MS and can alter the gas phase, compared to in-solution, stoichiometry of the complex [106] as solvent evaporation inherently concentrates the protein and ligand during the ESI process. Therefore, it is usually crucial to either prepare samples with a negative control/reference protein known not to bind the protein or ligand specifically [128], employ a mathematical model to deconvolute the binding data [129, 130], or follow up with a secondary technique for validation purposes. Since our initial goal for these direct infusion data was to simply evaluate monomer Oc binding to various metals, a proper reference protein for dimer formation was not included in these samples. Based on the infusion data, it was determined that nearly all possible dimer combinations of insulin, 1-, 2-, and 3-Gla-Oc were present in the mass spectra (**Figure 2.11**). Insulin + 1-Gla-Oc, 2 x 1-Gla-Oc, and 1-Gla-Oc + 2-Gla-Oc dimers were not observed. The reason for the lack of observation of these three dimers is likely because of the low abundance of 1-Gla-Oc in the monomer spectra. This is due to the clear observation of both 2 x Glu-Oc, and insulin + Glu-Oc, suggesting this interaction is not Gla-specific.

Non-specific interactions, by definition, should not depend on a protein's affinity for a given ligand, but rather only depend on the relative abundance of the protein and ligand. In fact, non-specific binding in native MS has also been shown to be independent of protein shape and size [131]. For an initial assessment on the likelihood of these dimers being a non-specific interaction, the predicted fractional abundance of each dimer was calculated based on the fractional abundance of the insulin, 1-, 2-, and 3-Gla-Oc monomers. When all observed dimers were included in the measured and predicted

fractional abundance calculations, all measured and predicted values were statistically significantly different ($p < 0.01$), except for the insulin-insulin, and 2-Gla-Oc + 3-Gla-Oc dimers. However, it has been shown by others [107] that insulin can form Ins-Ins dimers that are observed during native MS, and that these dimers are precursors for concentration-dependent multimeric homomers of insulin that give rise to insulin filamentation. The presence of Gla-Oc dimers has been observed previously, with the porcine Gla-Oc crystal structure showing this interaction occurs with increased affinity with coordination of both Gla-Oc with Ca^{2+} at the metal-binding Gla residue face of the proteins [32]. This suggests the Gla-Oc dimers are a specific interaction and leaves the insulin + Oc dimers as the sole dimer combinations that may be explained by non-specific interactions. Due to the lower fractional abundance of the insulin + Oc heterodimers observed compared to the prediction based on non-specific interaction, this interaction is most likely non-specific in nature.

The control experiment utilizing increased concentrations of insulin is suggestive that the conditions employed here do not reflect true native MS conditions. This is likely a result of the higher ion source voltages, source temperature, and the dimensions of the ESI tip [128]. Although this does not confidently exclude the possibility of non-specific protein-protein interactions, it is at least suggestive that the data presented here does not reflect true in-solution levels of protein dimer, with at least the Ins-Ins dimer formation disrupted in these conditions.

2.4.4 Monomer and Dimer Metal-Binding Comparison

Comparing the binding of metal cations to insulin monomer and insulin dimers, the insulin monomer and Ins-Ins dimer bind to the same number of metals, up to three Ca^{2+}, four Mg^{2+}, two Mn^{2+}, and zero Pb^{2+}, Fe^{3+} and Al^{3+}. These results suggest that the

metal-binding in Ins-Ins dimers is also a specific interaction, and that dimerization does not inhibit metal binding to take place. Whether these results match those found with Zn^{2+} would need to be assessed further, as Zn^{2+} binding was not tested in these experiments.

The decrease in the maximum stoichiometry of metal bound to insulin and Oc homo- and hetero-dimers compared to the monomeric proteins is likely due to the low fractional abundance of these dimers and combined sensitivity limitations of the technique resulting in the higher metal-bound protein dimer adducts being absent from the spectra. For two of the metals tested, Ca^{2+} and Al^{3+}, many of the Gla-Oc dimers were completely absent (all 2x Gla-Oc dimers with Ca^{2+} and all 2x Gla-Oc dimers except 2-Gla-Oc + 3-Gla-Oc with Al^{3+} data) from the mass spectra. In contrast to the 2x Gla-Oc dimers, the 2x Glu-Oc dimer was observed in the Ca^{2+} samples, with up to one Ca^{2+} bound visible in the 40 μM Ca^{2+} sample. Increased concentrations of Ca^{2+} increased the number of Ca^{2+} bound to Ins-Ins and Insulin + Glu-Oc dimers, with the 2x Glu-Oc dimers not confidently assigned in the 500 μM Ca^{2+} sample due to low signal intensity (**Figure 2.16f**). When combining this information for Ca^{2+} with the precipitation data discussed in **Chapter 4**, a possible explanation emerges: The metal-induced Oc precipitation is Gla-dependent and Gla-Oc dimers complexed with increasing numbers of metal are less soluble than their non-metal bound counterparts. Whether the Gla-Oc dimers are intermediates for filamentation would need to be assessed further. The reason for the Gla-Oc dimers absent in the Al^{3+} data is more likely to be explained by the low solubility of Al^{3+} at pH 7.4, resulting in $Al(OH)_3$ and Gla-Oc precipitation, and is discussed in **Chapter 5**.

2.5 Conclusions

Based on the data presented here, Oc binds metals in a Gla-dependent fashion and an increased number of Gla residues further increases its metal-binding affinity and stoichiometry. The ability of these metals to bind Gla-Oc can in part be explained by the metal-binding Gla residues being located on the protein's surface. Often proteins evolve a metal-binding pocket buried deep within the interior of the protein, enabling high specificity of metal binding due to the pocket's intolerance for coordination with metals that do not closely match the valence state and ionic radius of the target metal. The reason for this lack of specificity of metal binding to Oc is not fully understood, but could be attributed to lack of an evolutionary strain on Oc to do so, due to high Ca^{2+} concentrations and the low relative abundance of other competing metal cations in the bone. In addition, some metals were not bioavailable until the last few thousand years, as is the case with the toxic metal Al^{3+}, providing insufficient time for evolutionary adaptation. Binding of Gla-Oc to the hydroxyapatite surface has also been shown [7, 29], and may be involved in the protein's function, thus eliminating Oc's ability to evolve a metal-binding pocket buried deep within the protein's tertiary structure. Alternative hydroxyapatite structures, for example where Ca^{2+} is replaced by Mg^{2+} [132], are also found in low levels in human bone, and thus Gla-Oc's lack of metal-binding specificity may be important to accommodate for differences in crystal lattice structure of hydroxyapatite when these other metals are incorporated. Whatever the case, systematic studies on the mechanism by which Gla-Oc functions in the bone, and the effects of competitive metal binding, are necessary before conclusions can be drawn.

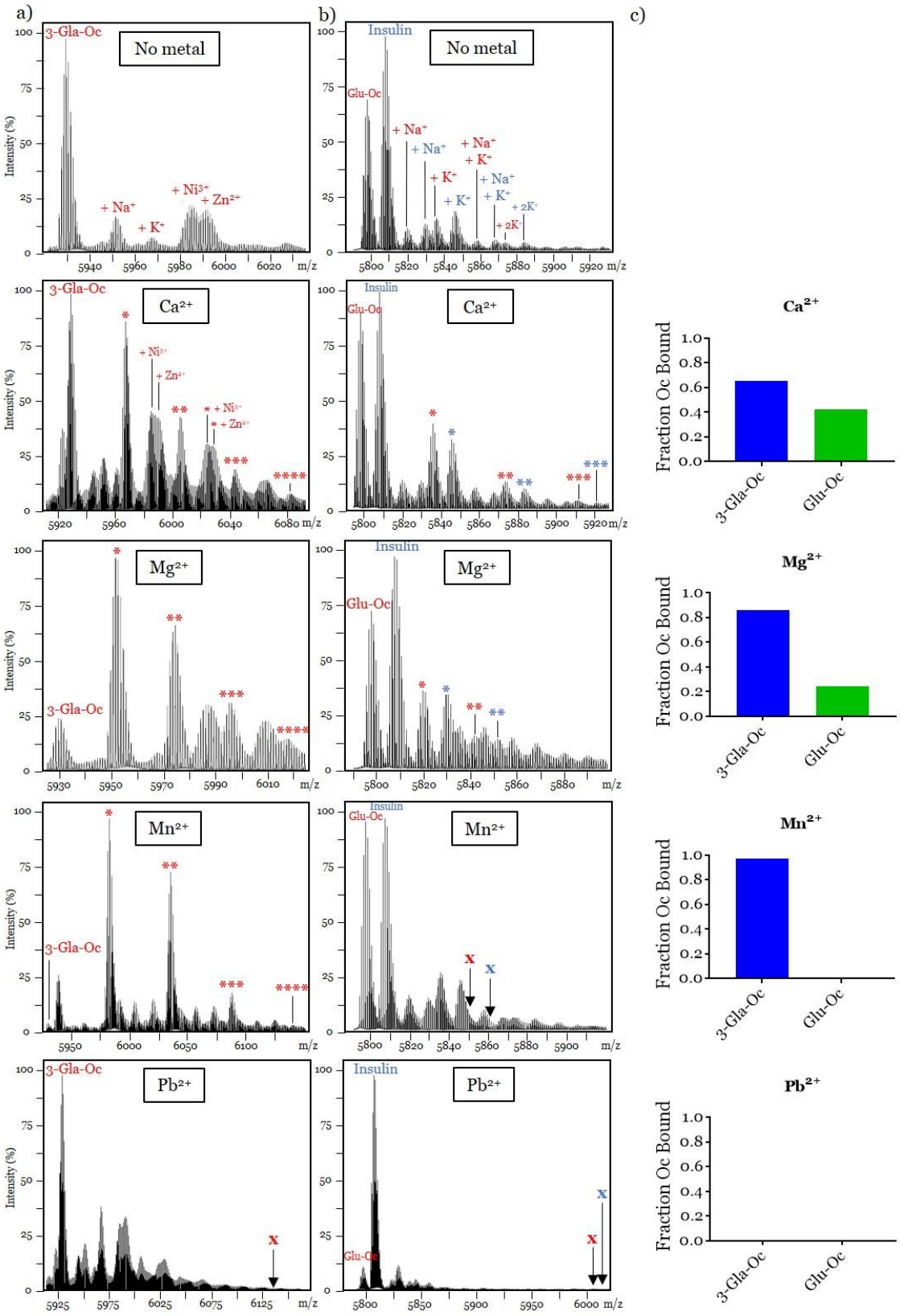

a)
b)
c)
3-Gla-Oc
No metal
Insulin
Glu-Oc
Ca²⁺
Mg²⁺
Mn²⁺
Pb²⁺
Intensity (%)
m/z
+ Na⁺
+ K⁺
+ Ni³⁺
+ Zn²⁺
+ 2K⁺
Fraction Oc Bound
1.0
0.8
0.6
0.4
0.2
0.0

Figure 2.1. Qualitative Assessment of Divalent Metal Cations Binding to Gla-Oc and Glu-Oc. Deconvoluted mass spectra of direct infusions of 4 μM insulin (blue) with 4 μM of either a) Gla-Oc or b) Glu-Oc (red) that was incubated with 0 μM (no metal), or 40 μM of Ca^{2+}, Mg^{2+}, Mn^{2+}, or Pb^{2+}. The mass spectral peaks associated with the metal-bound Oc (*) and insulin (*) are denoted with asterisks, where the number of asterisks corresponds to the number of the specified metal that is bound to the protein. An "x" with an arrow demonstrates that the metal-bound Oc (x) or insulin (x) mass spectral peak is not present in the spectrum, and points to the location where it would be found if present. c) Summary graphs displaying the fraction of Glu-Oc and 3-Gla-Oc bound by the specified metal.

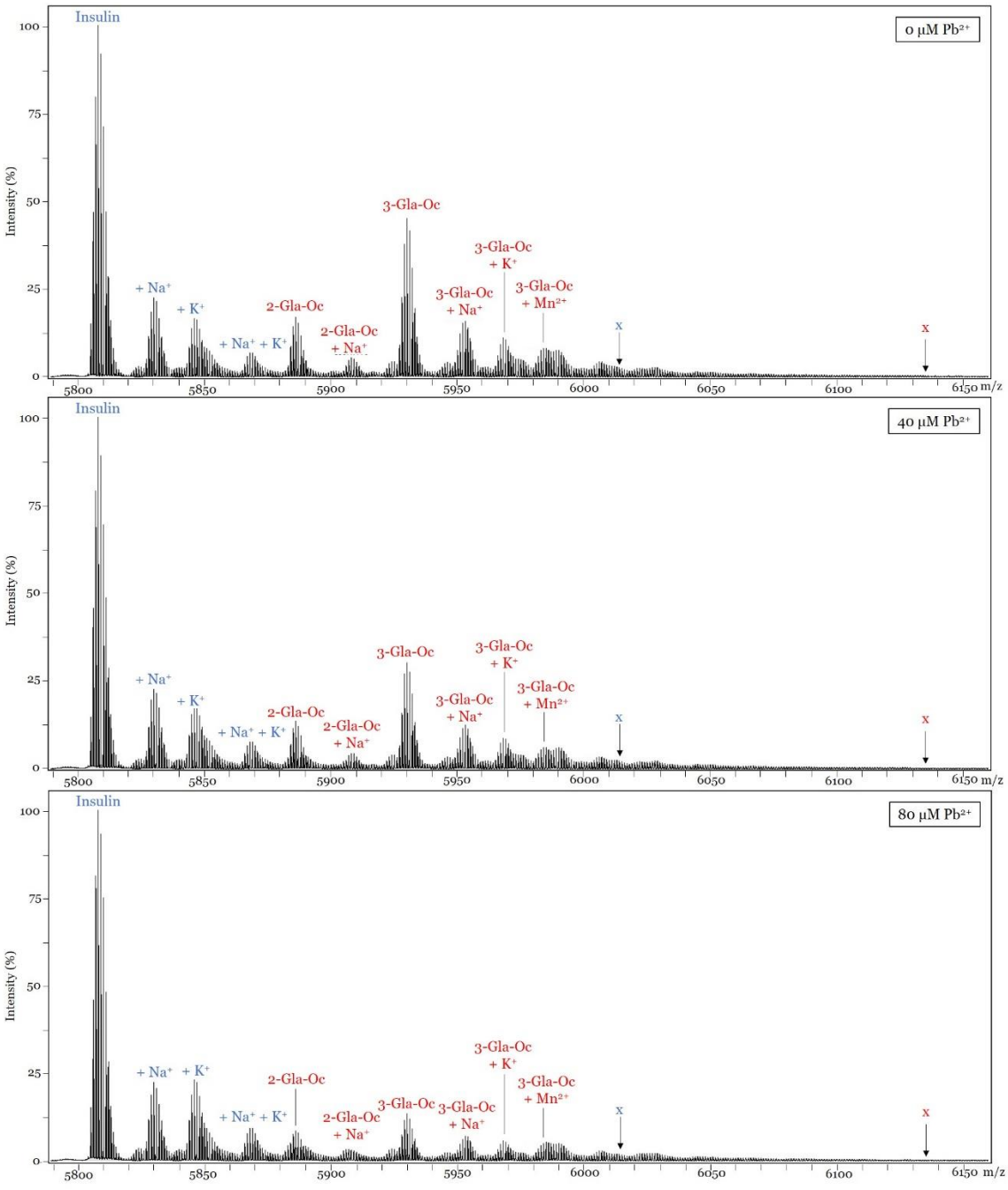


Figure 2.2. Analysis of Samples of Gla-Oc with Titrations of Pb^{2+} by Direct Infusion MS. Deconvoluted mass spectra of direct infusions of 4 μM Gla-Oc (red) + 4 μM insulin (blue) and 0, 40, or 80 μM Pb^{2+}. Insulin adducts (blue) and Gla-Oc adducts (red) are also specified, with the expected location of Pb^{2+}-bound insulin (x) and 3-Gla-Oc (x) marked with an arrow and an "x".

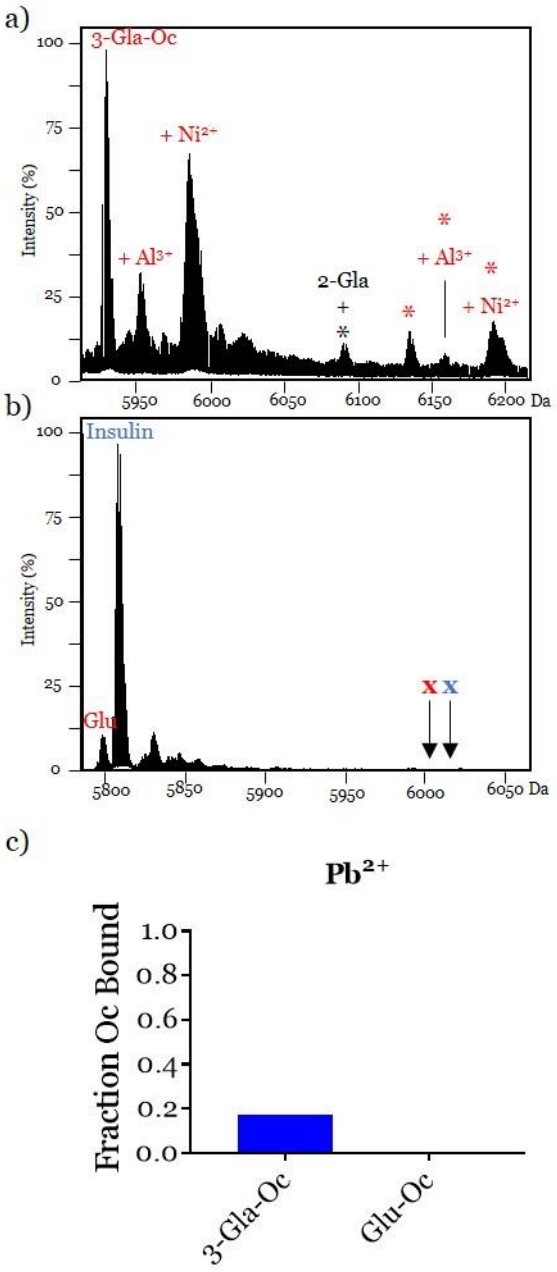


Figure 2.3. Gla-Oc Binding of Pb^{2+} at Lower Pb^{2+} Concentrations. a) 4 μM Gla-Oc and 4 μM insulin incubated with 4 μM Pb^{2+}, with 2-Gla-Oc bound singly to Pb^{2+} (2-Gla + *) and 3-Gla-Oc bound singly to Pb^{2+} (*), as well as Al^{3+} and Ni^{2+} adducts labeled in the spectrum. b) 4 μM Glu-Oc and 4 μM insulin incubated with 40 μM Pb^{2+}, with no observable Pb^{2+}-bound adducts observed (x for Glu-Oc and x for insulin in spectrum where Pb^{2+} adducts would be observed if present). c) The fraction of 3-Gla-Oc (0.17) and Glu-Oc (0) bound to Pb^{2+}.

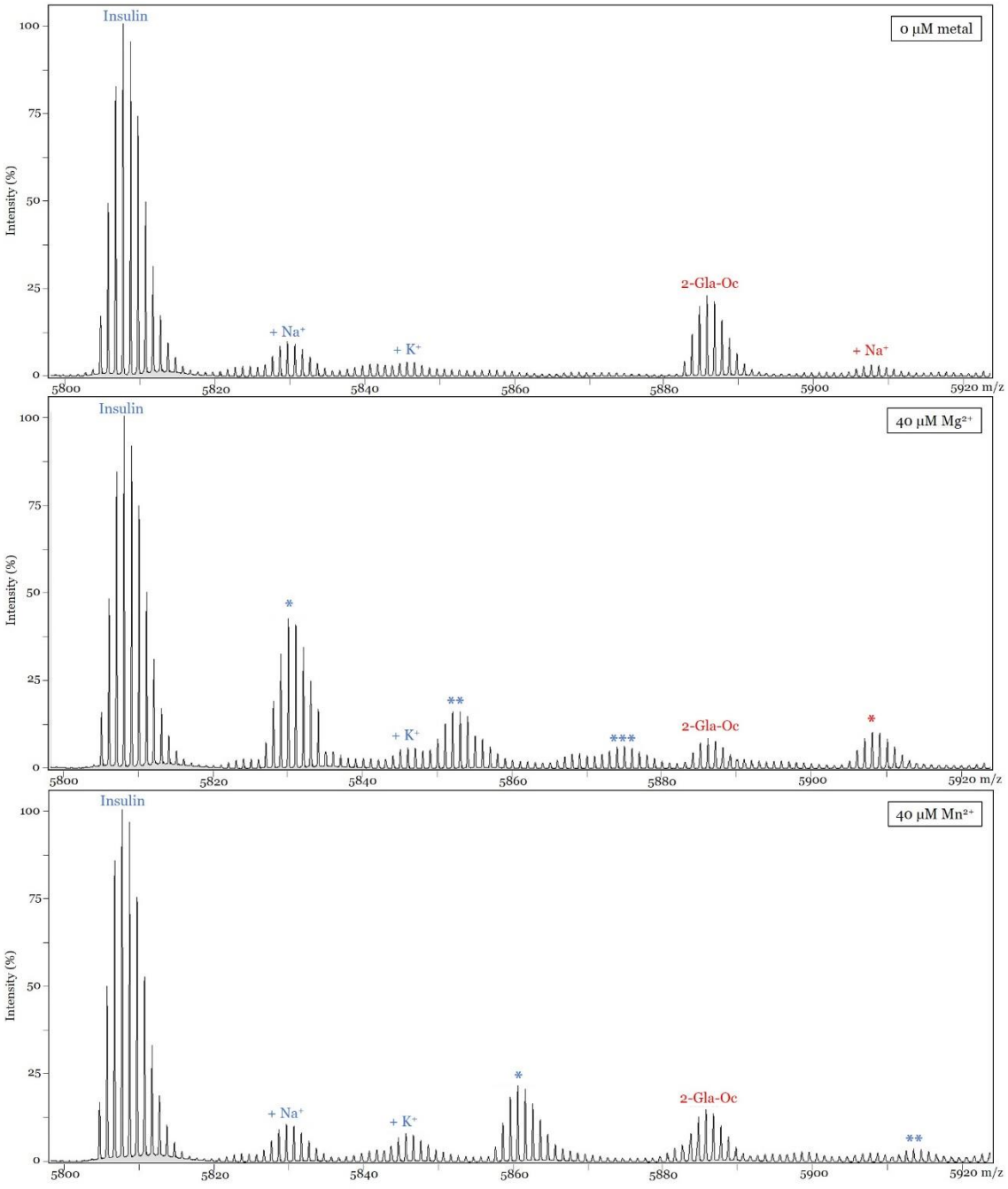


Figure 2.4. Analysis of Samples of Gla-Oc with Titrations of Mg^{2+} or Mn^{2+} by Direct Infusion MS. Deconvoluted mass spectra of direct infusions of 4 µM Gla-Oc (red) + 4 µM insulin (blue) with no metal (top), or 40 µM Mg^{2+} (middle) or Mn^{2+} (bottom). Mass spectral peaks associated with insulin bound to one (*) or two (**) Mg^{2+} or Mn^{2+}, and three (***) Mg^{2+} are found. 2-Gla-Oc bound to one Mg^{2+} (*) is also visible in the spectrum.

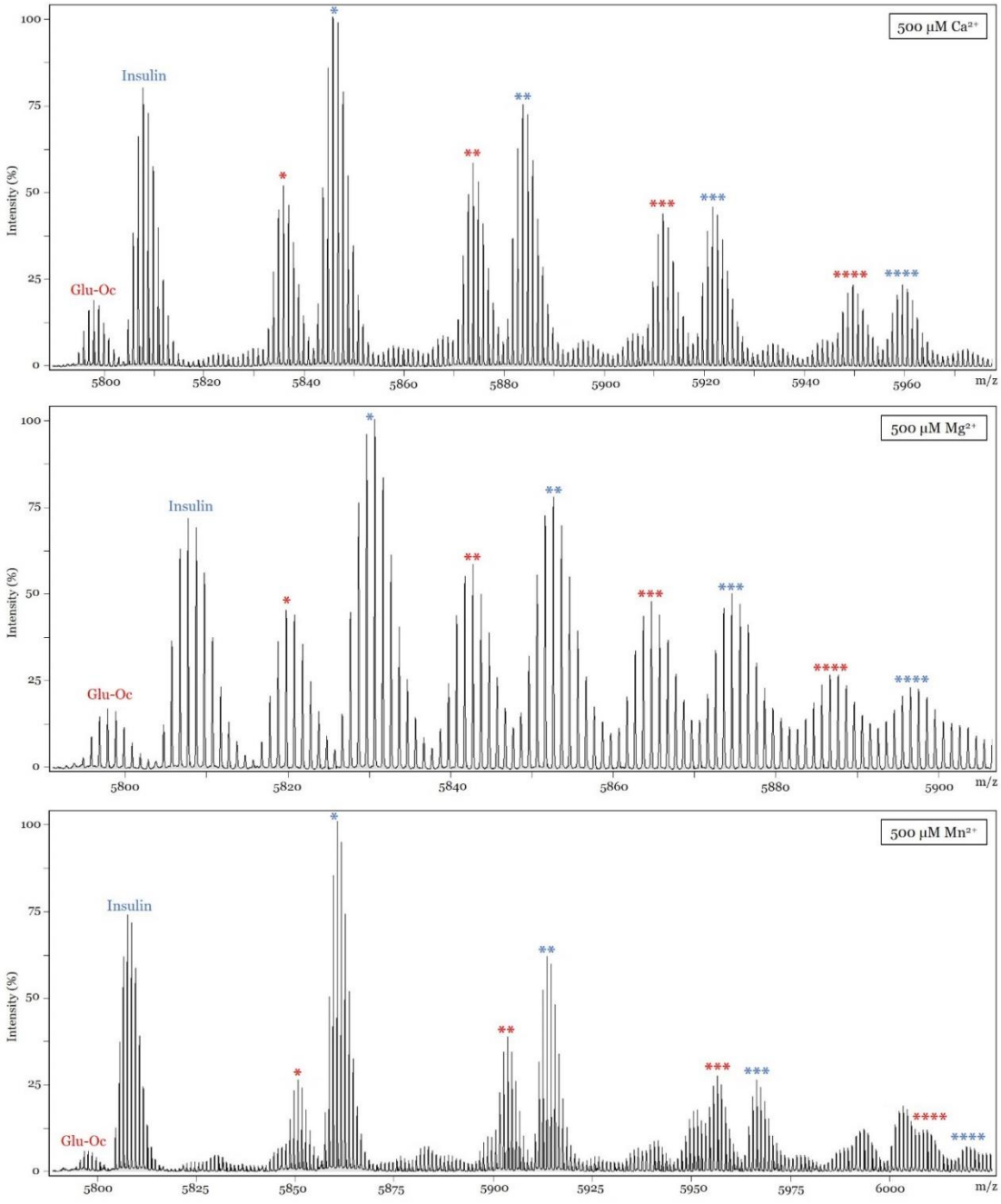


Figure 2.5. Analysis of Samples of Glu-Oc with Titrations of Ca^{2+}, Mg^{2+} or Mn^{2+} by Direct Infusion MS. Deconvoluted mass spectra of direct infusions of 4 µM Glu-Oc (red) + 4 µM insulin (blue) with 500 µM Ca^{2+} (top), Mg^{2+} (middle), or Mn^{2+} (bottom). Mass spectral peaks associated with insulin binding the respective metal are indicated by the blue asterisks ('*'), whereas Glu-Oc metal-bound adducts are indicated by the red asterisks ('*'), with the number of asterisks corresponding to the number of the specified metal bound.

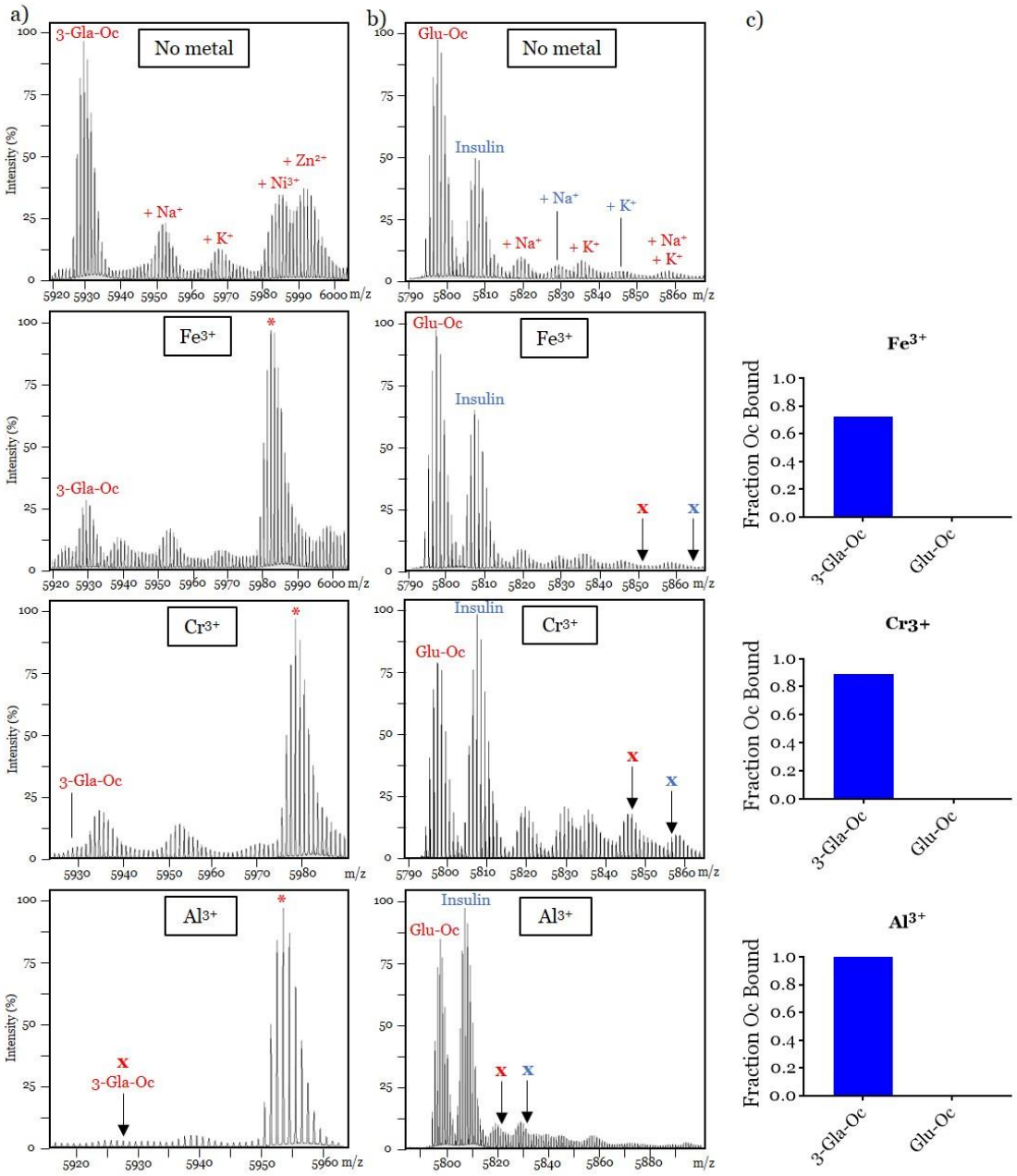


Figure 2.6. Qualitative Assessment of Trivalent Metal Cations Binding to Gla-Oc and Glu-Oc. Deconvoluted mass spectra of direct infusions of 4 μM insulin (blue) with 4 μM of either a) Gla-Oc or b) Glu-Oc (red) that was incubated with 0 μM (no metal), or 40 μM of Fe^{3+}, Al^{3+}, or 12 μM Cr^{3+}. The mass spectral peaks associated with the metal-bound Oc are denoted with asterisks "*", where the number of asterisks corresponds to the number of the specified metal that is bound. An "x" with an arrow demonstrates that the metal-bound protein peak is not present in the spectrum, and points to the location where it would be found if present for Oc (x) and insulin (x).

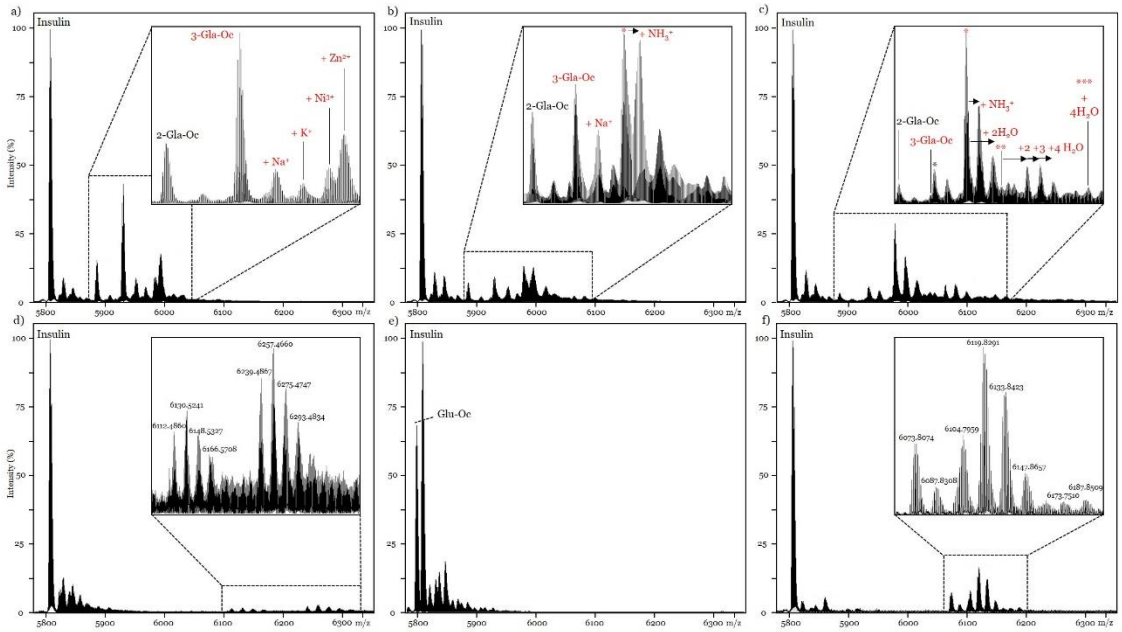


Figure 2.7. Analysis of Samples of Gla-Oc, or Glu-Oc, with Titrations of Cr^{3+} by Direct Infusion MS. Deconvoluted mass spectra of 4 μM Gla-Oc incubated with a) 0 μM, b) 4 μM, c) 12 μM, d) 40 μM, or 4 μM Glu-Oc incubated with e) 0 μM, or f) 40 μM Cr^{3+}. Full spectra include insulin labeled as a reference. Zoom-in regions of the spectra that are boxed off with dashes include 2-Gla-Oc (black), 2-Gla-Oc bound singly to Cr^{3+} (*) in c), 3-Gla-Oc (red) and bound to Cr^{3+} (*), with the number of asterisks corresponding to the number of Cr^{3+} bound. All other adducts that are classified include arrows from the mass spectral peak corresponding to the species of interest, with "+" indicating the additional adduct involved in the mass spectral shift. The species were not identified in d) and f) and thus the monoisotopic mass (m/z) is indicated for the x + 4 isotope peak.

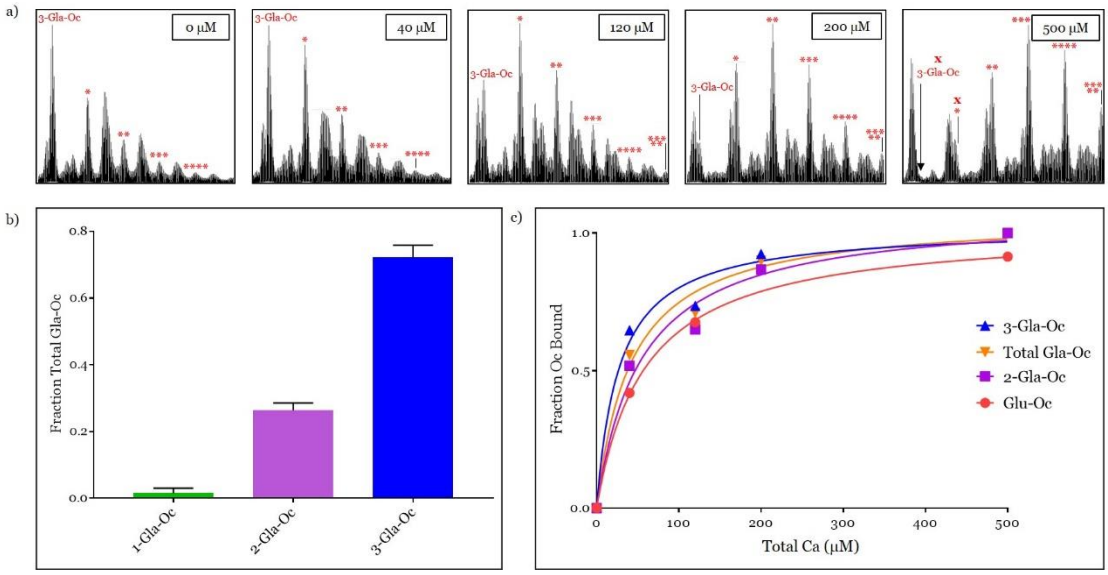


Figure 2.8. Qualitative Effect of Under Carboxylation of Gla-Oc on Binding to Ca^{2+}. Direct infusions of 4 µM Gla-Oc with 4 µM insulin and 0, 40, 120, 200, and 500 µM total Ca titrations with the deconvoluted mass spectra in a) focusing on the peaks associated with 3-Gla-Oc (red and labeled) and 3-Gla-Oc coordinating with one (*), two (**), three (***), four (****), and five (*****) Ca^{2+}. b) Fraction of total Gla-Oc that is either 1-Gla-Oc (green), 2-Gla-Oc (purple), or 3-Gla-Oc (blue), n = 3 with error bars depicted based on SD. Binding curves are plotted in c) that distinguish between the fractions of Glu-Oc (red), 2-Gla-Oc (purple), and 3-Gla-Oc (blue) bound to Ca^{2+}, plotted against the total Ca at the concentrations discussed above. The effect of combined peak intensities from all forms of Gla-Oc on the fraction bound by Ca^{2+} is shown in the "Total Gla-Oc" data set (orange). Insufficient levels of 1-Gla-Oc were present in the mass spectra to obtain a binding curve for this sample set. n = 1.

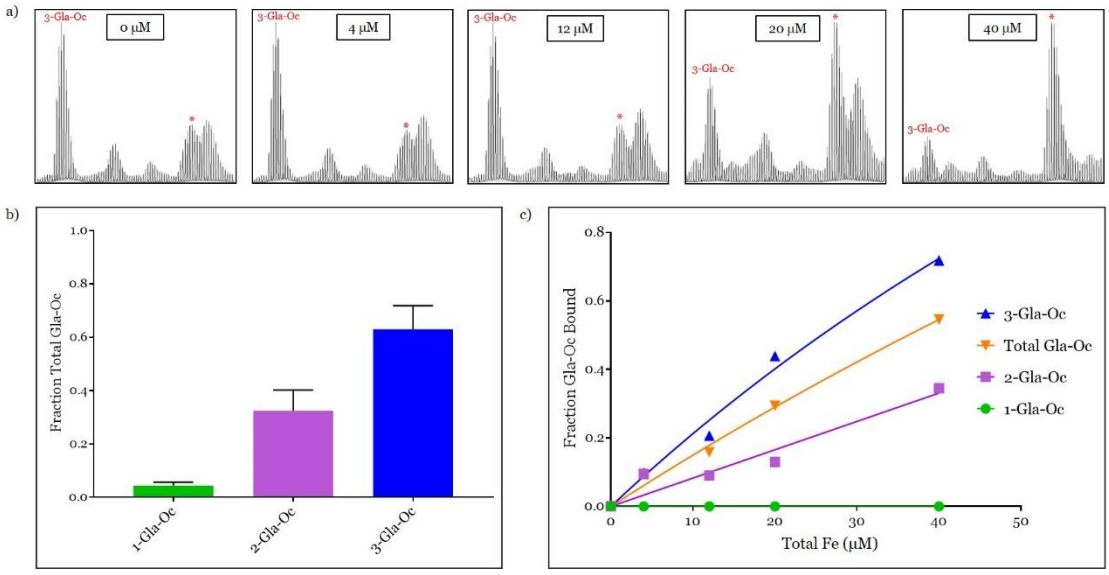


Figure 2.9. Qualitative Effect of Under Carboxylation of Gla-Oc on Binding to Fe^{3+}. Direct infusions of 4 µM Gla-Oc with 4 µM insulin and 0, 4, 12, 20, and 40 µM total Fe titrations with the deconvoluted mass spectra in a) focusing on the peaks associated with 3-Gla-Oc (red and labeled) and 3-Gla-Oc coordinating with one Fe^{3+} (*). b) Fraction of total Gla-Oc that is either 1-Gla-Oc (green), 2-Gla-Oc (purple), or 3-Gla-Oc (blue), n = 3 with error bars depicted based on SD. Binding curves are plotted in c) that distinguish between the fractions of 1-Gla-Oc (green), 2-Gla-Oc (purple), and 3-Gla-Oc (blue) bound to Fe^{3+} at the total Fe concentrations discussed above. The effect of combined peak intensities from all forms of Gla-Oc on the fraction bound by Fe^{3+} is shown in the "Total Gla-Oc" data set (orange). n = 1.

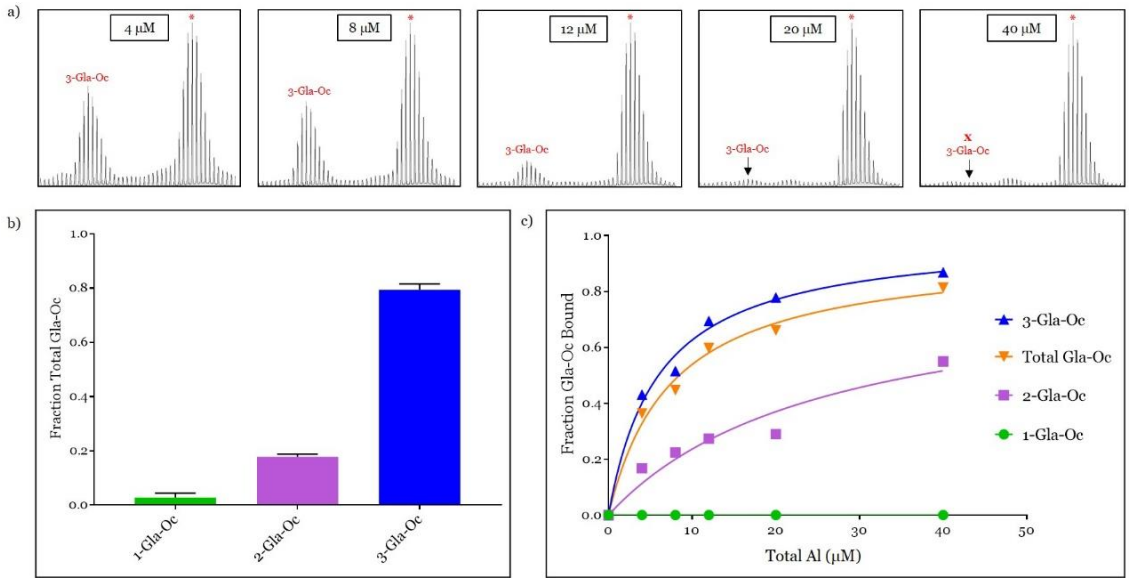


Figure 2.10. Qualitative Effect of Under Carboxylation of Gla-Oc on Binding to Al^{3+}. Direct infusions of 4 µM Gla-Oc with 4 µM insulin and 4, 8, 12, 20, and 40 µM total Al titrations with the deconvoluted mass spectra in a) focusing on the peaks associated with 3-Gla-Oc (red and labeled) and 3-Gla-Oc coordinating with one Al^{3+} (*). b) Fraction of total Gla-Oc that is either 1-Gla-Oc (green), 2-Gla-Oc (purple), or 3-Gla-Oc (blue), n = 3 with error bars depicted based on SD. Binding curves are plotted in c) that distinguish between the fractions of 1-Gla-Oc (green), 2-Gla-Oc (purple), and 3-Gla-Oc (blue) bound to Al^{3+} at the total Al concentrations discussed above. The effect of combined peak intensities from all forms of Gla-Oc on the fraction bound by Al^{3+} is shown in the "Total Gla-Oc" data set (orange). n = 1.

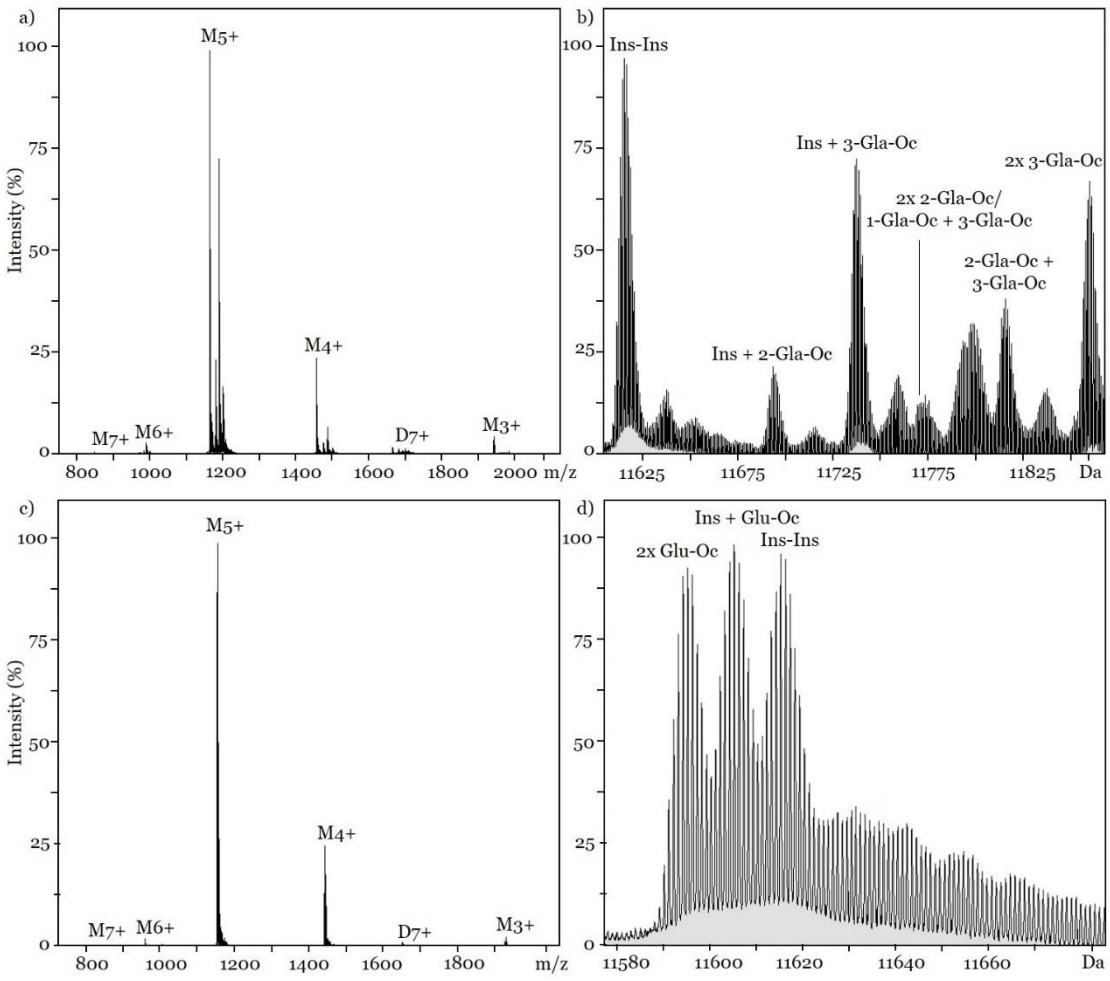


Figure 2.11. Insulin and Oc Monomer and Homo- and Hetero-Dimers Present in Direct Infusion MS. Monomer and dimer charge states for a) 4 μM Gla-Oc, and b) 4 μM Glu-Oc, incubated with 4 μM insulin. The mass spectral peaks associated with monomers (M) and dimers (D) are indicated in the spectra, with the charge state of each peak cluster also specified. The deconvoluted mass spectrum for c) the Gla-Oc sample, and d) the Glu-Oc sample, depict the combinations of dimers observed in each spectrum.

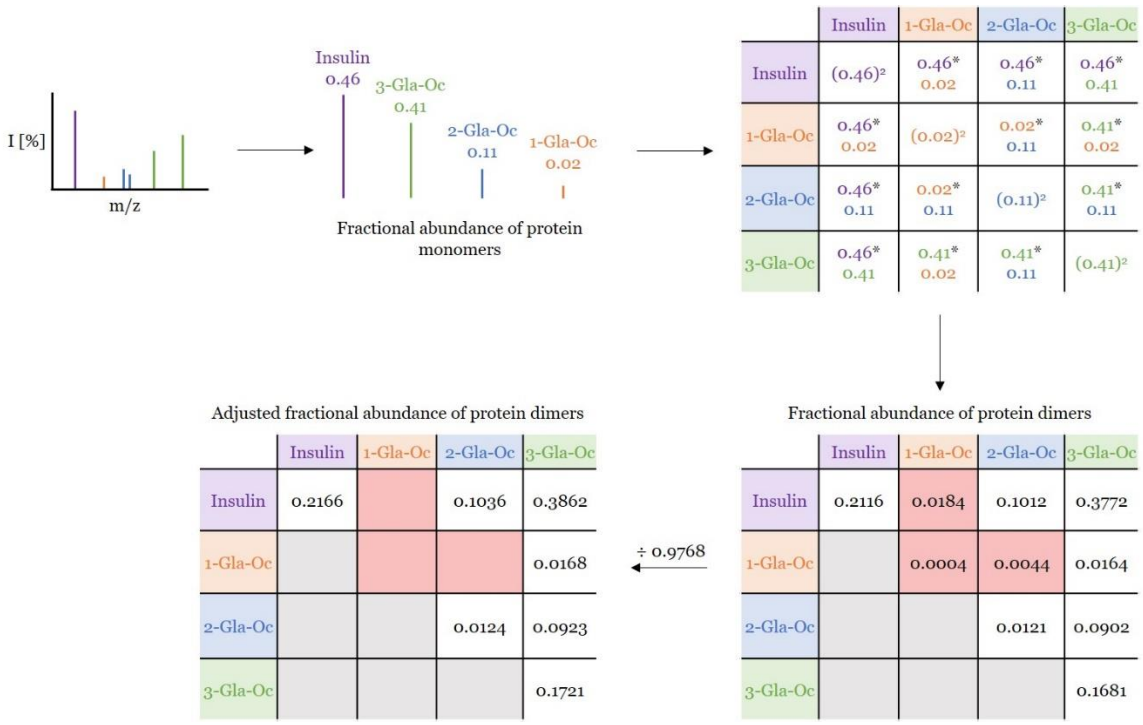


Figure 2.12. Binomial Expansion to Predict Fractional Abundance of Protein Dimers Based on Non-Specific Interaction. Prediction of the fractional abundance of each dimer due to non-specific interactions is dependent on the fractional abundance of each protein's monomer, determined using the deconvoluted mass spectra, and the probability of each protein to interact based on this abundance alone. First, each mass spectral peak for the monomers is classified as either a peak associated with insulin, or 1-, 2-, or 3-Gla-Oc. The intensity of each peak (counts) is recorded, the peaks associated with each protein are summed, followed by dividing by the total counts recorded for all proteins present. These fractions are then used to calculate the fraction of each dimer that should be present if the interactions are solely based on non-specific binding, as shown in the table on the far right. The dimers absent from the spectra, Insulin + 1-Gla-Oc, 2x 1-Gla-Oc, and 1-Gla-Oc + 2-Gla-Oc, are excluded and the fractional abundance of the remaining dimers is normalized so the sum of the final fractional abundance values is adjusted to 1. In this example, this is done by dividing by 0.9768, which is the sum of the fractional abundance values once the unobserved dimers are excluded.

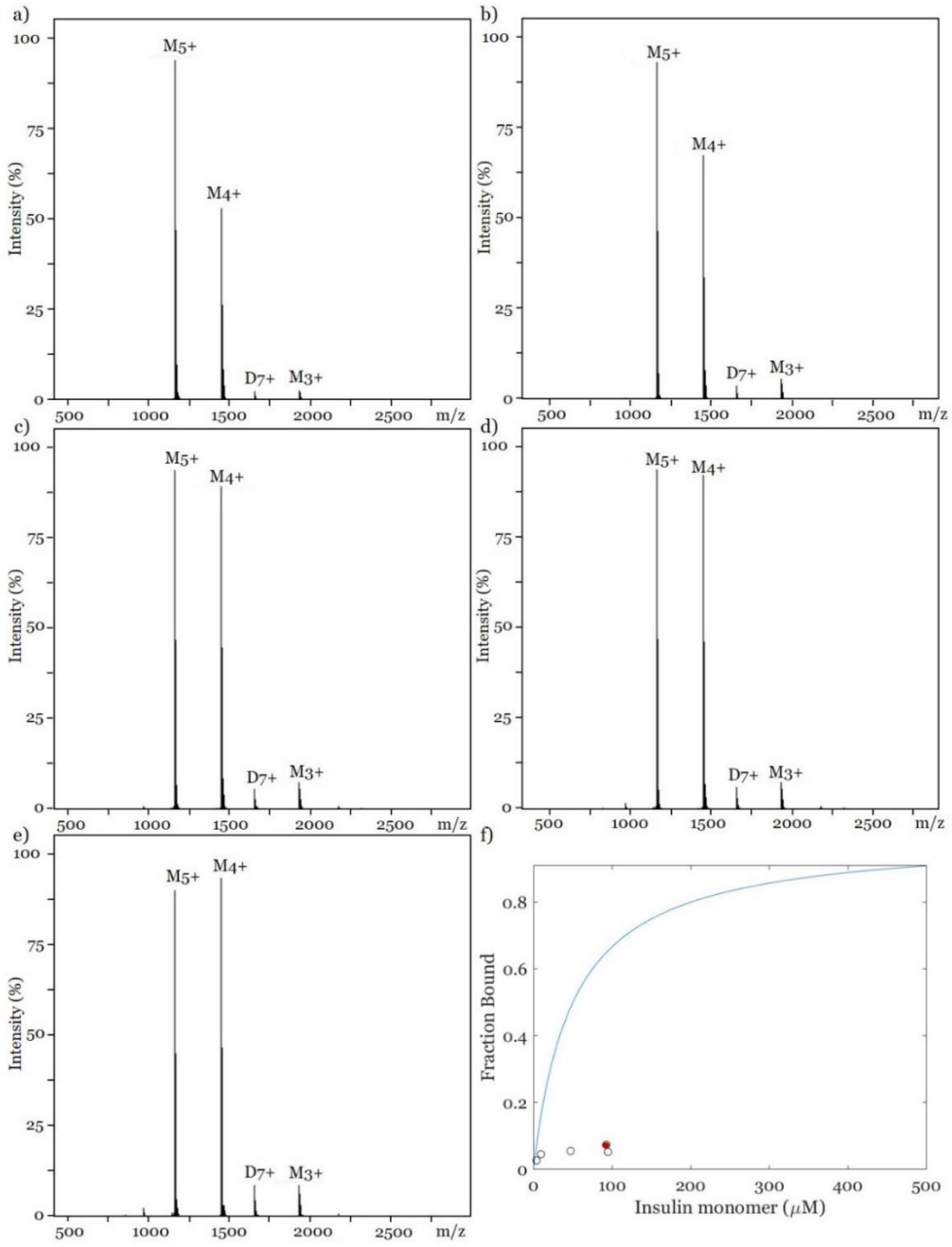


Figure 2.13. Insulin Control Experiment to Evaluate Dimer Formation in Tested Native MS Conditions. Monomer and dimer charge states for a) 4 μM, b) 10 μM, c) 50 μM, and d) 100 μM insulin in 5 mM ammonium acetate, pH 7.4. d) 100 μM insulin in 5 mM ammonium acetate, pH 3.3. All mass spectra include monomer 5+ (M5+), 4+ (M4+), 3+ (M3+), and dimer 7+ (D7+) charge states clearly labelled. f) Fraction of infused insulin in dimer form is plotted against the concentration of insulin monomer in each infusion sample at pH 7.4 (black circle data points) and the sample infused at pH 3.3 (red circle data point). The data were plotted alongside the binding curve based on the known K_d = 100 μM for monomer to dimer formation and is represented by the blue trace.

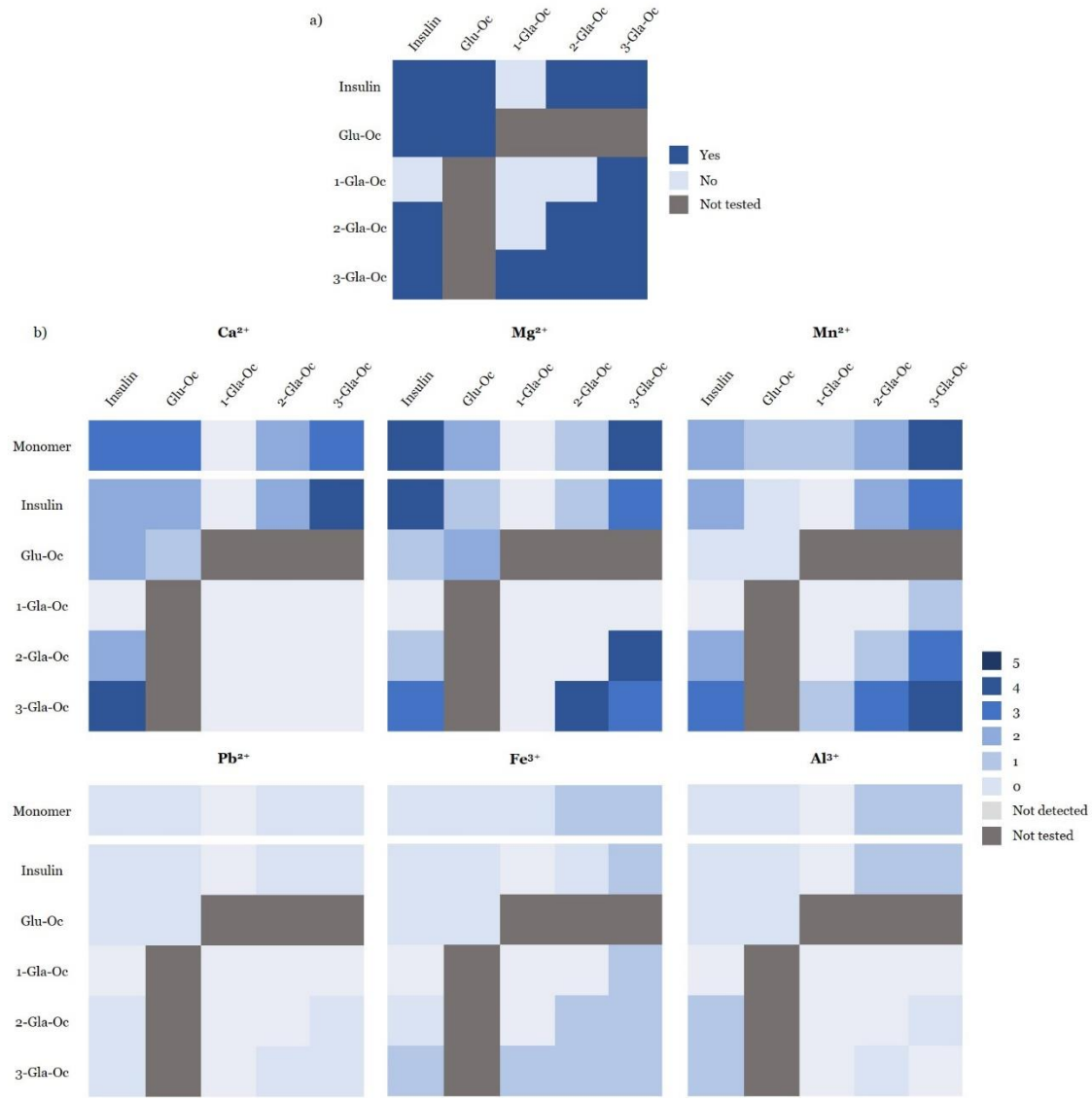


Figure 2.14. Stoichiometry of Metal Binding to Monomer and Homo- and Hetero-Dimers of Oc and Insulin. a) The possible dimer combinations are found in the histogram, with observed, not observed, and not tested clearly shown. A comparison of the monomer and dimer metal binding capacity is depicted in b), with the maximum number of each metal bound denoted by the colour coding shown, with dimers not tested in dark grey, monomers or dimers not detected in the mass spectra in light grey, and observed but no metal binding detected, or maximum of one to five of the specified metal bound depicted by the blue coloured gradient shown in the legend. The results depicted here are derived from 4 μM insulin with 4 μM Oc and 40 μM of the specified metal.

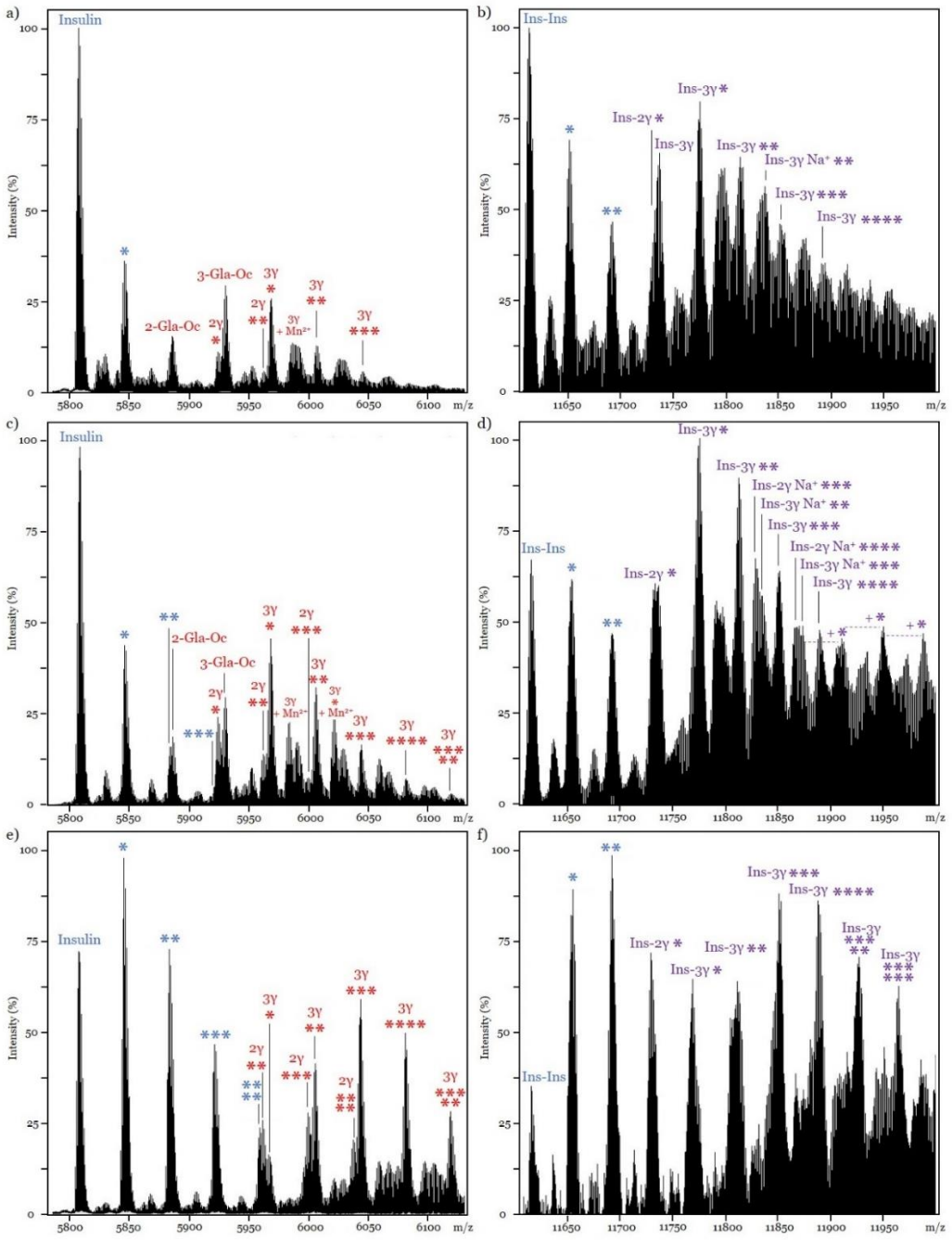


Figure 2.15. Ca^{2+} Binding to Insulin and Gla-Oc Monomer and Dimers. Monomer (a, c, e) and dimer (b, d, f) mass spectra of insulin with Gla-Oc and 40 μM (a & b), 120 μM (c & d), or 500 μM (e & f) Ca^{2+}. Monomer mass spectra include mass spectral peaks for insulin labelled in blue, 2-Gla-Oc (2γ) and 3-Gla-Oc (3γ) in red, and up to five Ca^{2+} bound, indicated by the asterisks ('*') in the appropriate colour, with the number of asterisks corresponding to the number of Ca^{2+} bound. Dimer mass spectra include mass spectral peaks for the insulin-insulin dimer (Ins-Ins) and Ca^{2+} bound dimers ('*'), and insulin + 2-Gla-Oc (Ins-2γ) and insulin + 3-Gla-Oc (Ins-3γ), and Ca^{2+} bound dimers ('*').

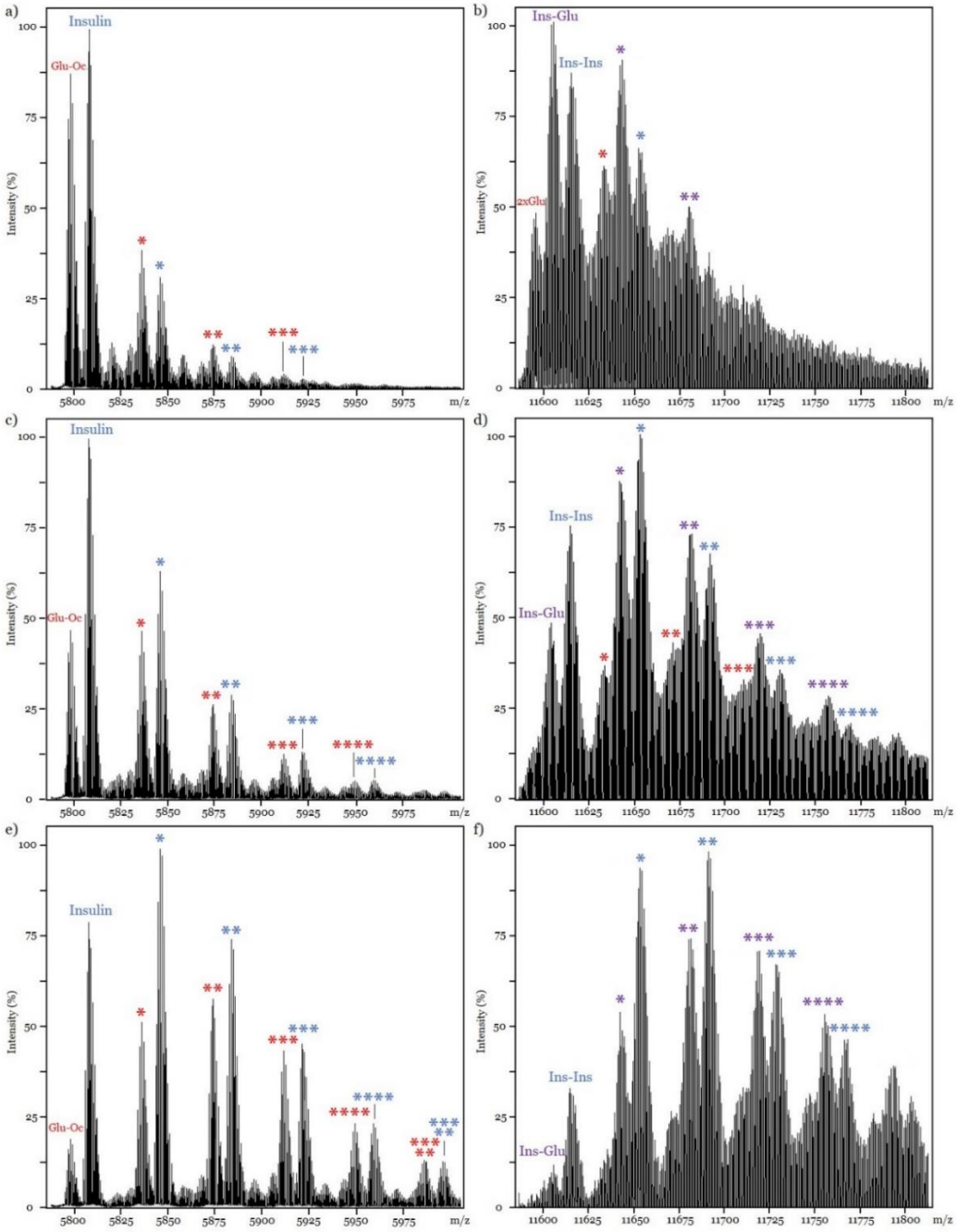


Figure 2.16. Ca^{2+} Binding to Insulin and Glu-Oc Monomer and Dimers. Monomer (a, c, e) and dimer (b, d, f) mass spectra of insulin with Glu-Oc and 40 μM (a & b), 120 μM (c & d), or 500 μM (e & f) Ca^{2+}. Monomer mass spectra include mass spectral peaks for insulin labelled in blue, Glu-Oc in red, and up to five Ca^{2+} bound, indicated by the asterisks ('*') in the appropriate colour, with the number of asterisks corresponding to the number of Ca^{2+} bound. Dimer mass spectra include mass spectral peaks for the Glu-Oc + Glu-Oc (2xGlu), insulin-insulin dimer (Ins-Ins), and insulin + Glu-Oc (Ins-Glu) and Ca^{2+} bound dimers ('*') in the appropriate colour.

Table 2.1. Comparison of Measured and Calculated Monoisotopic Masses of Metal-Free and Metal-Bound Insulin and Oc Monomers. Measured and calculated monoisotopic mass (Da) of insulin, Glu-Oc, 1-, 2-, and 3-Gla-Oc monomers and each singly-bound protein-metal adduct, with the difference between these values (Δ) calculated. ND, not detected.

	INSULIN			GLU-OC			1-GLA-OC			2-GLA-OC			3-GLA-OC		
Metal	Measured	Calculated	Δ	Measured	Calculated	Δ	Measured	Calculated	Δ	Measured	Calculated	Δ	Measured	Calculated	Δ
No metal	5806.6837	5806.6601	0.0236	5796.8226	5796.7679	0.0547	5840.7091	5840.7577	0.0486	5884.7683	5884.7475	0.0208	5928.7621	5928.7374	0.0247
Ca^{2+}	5844.6763	5844.6071	0.0692	5834.7790	5834.7149	0.0641	ND	5878.7047	-	5922.7564	5922.6945	0.0619	5966.7488	5966.6844	0.0644
Mg^{2+}	5828.6511	5828.6295	0.0216	5818.8066	5818.7373	0.0693	ND	5862.7271	-	5906.7299	5906.7169	0.013	5950.7218	5950.7068	0.015
Mn^{2+}	5859.5981	5859.5825	0.0156	5849.6829	5849.6903	0.0074	ND	5893.6801	-	5937.6876	5937.6699	0.0177	5981.682	5981.6598	0.0222
Pb^{2+}	ND	6012.6212	-	ND	6002.729	-	ND	6046.7188	-	ND	6090.7086	-	ND	6134.6985	-
Fe^{3+}	ND	5859.5716	-	ND	5849.6794	-	ND	5893.6692	-	5937.6524	5937.659	0.0066	5981.6393	5981.6489	0.0096
Cr^{3+}	ND	5855.5772	-	ND	5845.685	-	ND	5889.6748	-	5933.7276	5933.6646	0.063	5977.725	5977.6545	0.0705
Al^{3+}	ND	5830.6182	-	ND	5820.726	-	ND	5864.7158	-	5908.7052	5908.7056	0.0004	5952.7017	5952.6955	0.0062

Table 2.2. Fraction of Protein Monomers Bound to Metal. Fraction of protein (4 μM) monomers that were bound to the specified metal (40 μM). Calculations completed using peak intensities (counts) for each protein peak and metal-bound protein adduct. Maximum number of each metal bound to the specified protein observed at the given metal concentration is shown in brackets. ND, not detected.

	FRACTION BOUND				
Metal	Insulin	Glu-Oc	1-Gla-Oc	2-Gla-Oc	3-Gla-Oc
Ca^{2+}	0.33 (3)	0.42 (4)	ND	0.52 (2)	0.65 (5)
Mg^{2+}	0.40 (4)	0.24 (2)	ND	0.55 (1)	0.86 (4)
Mn^{2+}	0.21 (2)	0.05 (1)	0.40 (1)	0.75 (2)	0.97 (4)
Pb^{2+}	0	ND	ND	ND	ND
Fe^{3+}	0	0	0	0.35 (1)	0.72 (1)
Cr^{3+}	0	0	1	1	1
Al^{3+}	0	0	ND	0.55 (1)	1 (1)

Table 2.3. Measured and Calculated Monoisotopic Masses of Insulin and Oc Dimers. Measured and calculated monoisotopic mass of all possible insulin, Glu-Oc, 1-, 2-, and 3-Gla-Oc dimers, with the difference between these values (Δ) calculated.

	MH^+ (Da)		
Dimer	Measured	Calculated	Δ
Insulin-Insulin	11613.4081	11613.3202	0.0879
Insulin + Glu-Oc	11603.5725	11603.428	0.1445
Insulin + 1-Gla-Oc	ND	11647.4178	-
Insulin + 2-Gla-Oc	11691.4876	11691.4076	0.08
Insulin + 3-Gla-Oc	11735.4813	11735.3975	0.0838
2x Glu-Oc	11593.678	11593.5358	0.1422
2x 1-Gla-Oc	ND	11681.5154	-
1-Gla-Oc + 2-Gla-Oc	ND	11725.5052	-
2x 2-Gla-Oc/ 1-Gla-Oc + 3-Gla-Oc	11769.5173	11769.495	0.0223
2-Gla-Oc + 3-Gla-Oc	11813.545	11813.4849	0.0601
2x 3-Gla-Oc	11857.5485	11857.4748	0.0737

Table 2.4. Comparison of the Measured and Predicted Fractional Abundance of Each Insulin and Oc Dimer Observed. P-value < 0.05 indicated by an asterisk (*).

	FRACTIONAL ABUNDANCE OF PROTEIN DIMERS		
Protein	Measured	Predicted	p-value
Insulin-Insulin	0.2296 ± 0.0166	0.2185 ± 0.0460	0.713351
Insulin + 2-Gla-Oc	0.0496 ± 0.0071	0.0994 ± 0.0023	0.000327*
Insulin + 3-Gla-Oc	0.2143 ± 0.0237	0.3875 ± 0.0037	0.000235*
2x 2-Gla-Oc/1-Gla-Oc + 3-Gla-Oc	0.0701 ± 0.0030	0.0279 ± 0.0029	0.000063*
2-Gla-Oc + 3-Gla-Oc	0.1216 ± 0.0070	0.0903 ± 0.0157	0.034385
2x 3-Gla-Oc	0.3149 ± 0.0236	0.1764 ± 0.0330	0.004106*

CHAPTER 3

PENTAMERIC HUMAN OSTEOCALCIN DESTABILIZATION IN THE PRESENCE OF SODIUM CHLORIDE

3.1 Introduction

The quaternary structure of proteins is described as how different protein chains, or subunits, assemble with one another. These protein complexes can be homomeric – made up of multiple copies of a single protein subunit – or heteromeric – made up of protein subunits that are distinct from one another. Typically, the subunits that make up heteromeric complexes are encoded on separate genes, although they can also be formed by proteolytic cleavage of single chains [133]. Approximately 45 % of eukaryotic proteins can form homomeric complexes [134], with the majority of these complexes displaying structural symmetry. There are six different classes of homomeric protein complexes that are categorized based on their symmetry, extensively reviewed by Marsh & Teichmann [133] and Bergendahl & Marsh [134]. Each symmetry class has distinct protein functions and demonstrates the interconnectedness of structure and function in protein complexation. The formation of homomeric complexes is beneficial because it allows for regulation of complex assembly by monomer concentration and reduces error during DNA replication, transcription, and translation due to decreased genetic space encoding the assembly. Although many proteins adopt highly uniform homomeric complex states, some proteins assemble into two or more quaternary structural states that exist in equilibrium with one another. For example, the small heat-shock protein αB-crystallin self-assembles into multiple homomeric complex states that involve between approximately 10 to 40 monomer subunits [135-137]. Insulin, as mentioned previously, is another example protein that can function in monomer, dimer, and hexamer states [120].

As mentioned in previous chapters, the function of Oc is still not completely understood. Since the structure of a protein is often linked to its function, understanding the quaternary structure of Oc may assist to uncover the mechanisms by which it functions. Presented here, it was found that Oc can form multiple different homomeric complexes with the number of monomers involved in the assemblies governed by NaCl concentration. At 50 μM Oc concentrations and in the absence of NaCl, Gla-Oc was found to form a pentamer, while Glu-Oc formed both pentamer and tetramer structures, with both complexes formed by Oc destabilized with increasing concentrations of NaCl, with Oc in the monomeric state at roughly the normal concentration of sodium in extracellular fluid (ECF), ~ 136 mM [138]. Estimation of the concentration of Oc within the human bone revealed Oc to be within the hundreds of micromolar to even millimolar concentration range, suggesting the self-assemblies observed here in an *in vitro* setting may be relevant biologically. Although additional experiments are required for confident complex size characterizations and to better understand the significance of complexation of Oc, these preliminary experiments set the framework for potential future research on Oc homomeric complexation.

3.2 Materials and Methods

3.2.1 Materials

Both the Des-Gla synthetic human osteocalcin (Glu-Oc), Cat # AS-65307-025, and [Gla 17, 21, 24] – synthetic human osteocalcin (Gla-Oc), Cat # AS-22830, were purchased from AnaSpec, Co. Ammonium hydroxide, bovine serum albumin (BSA) heat shock fraction, protease free, pH 7 (≥ 98 %), Coomassie brilliant blue R-250 dye, formaldehyde, glacial acetic acid, recombinant human insulin expressed in yeast (≥ 98 %), silver (I) nitrate (99.9999 % trace metals basis), sodium carbonate (≥ 90.0 % purity), sodium chloride (trace metal grade, 99.999 % purity), sodium hydroxide (minimum 98 % purity),

sodium thiosulphate, Tris base (≥ 99.9 %), Tris hydrochloride (≥ 99.0 % purity), and uridine (> 99 % purity) were purchased from Sigma-Aldrich, Co. The small-volume 0.2 µm filters were purchased from Pall, Corp. Mini-PROTEAN TGX precast 4 – 20 % native PAGE gels, precision plus protein standards, and gel filtration standard were purchased from Bio-Rad Laboratories, Co. The 0.2 µm filters used to filter SEC running buffers were purchased from Corning, Inc.

3.2.2 Preparation of Osteocalcin Stocks

When preparing Oc stock solutions in 20 mM Tris + 150 mM NaCl, pH 7.4, the Gla-Oc lyophilized powder was reconstituted initially with 40 µL of 1 % ammonium hydroxide in dIH_2O, swirled gently to ensure full reconstitution of the protein powder, and then the final volume was brought to 500 µL with 20 mM Tris, pH 7.4. The pH was verified using short range pH paper and adjusted with small volumes of HCl if necessary. For the Glu-Oc vials, the reconstitution technique was identical as described for Gla-Oc vials above, except the volume of 1 % ammonium hydroxide and final volume were half that of the Gla-Oc vials (20 µL, 250 µL) to account for there being half the weight of lyophilized protein (~ 0.25 mg vs. ~ 0.5 mg). Stock Oc sample concentrations were then determined by measuring absorbance at 280 nm using a Thermo Scientific Multiskan Go UV-vis spectrophotometer, with Oc's extinction coefficient known, $\varepsilon = 1.33\ (mg/mL)^{-1}cm^{-1}$ [104].

3.2.3 Size Exclusion Chromatography

Size exclusion chromatography (SEC) was completed by injecting 100 µL of 35 – 50 µM Oc using a syringe onto either a Superdex 200 Increase, 10 x 300 mm, or Superdex 75 Increase, 10 x 300 mm, connected to a Bio-Rad NGC Chromatography System FPLC. Samples were centrifuged for 5 minutes at 17, 000 x g immediately prior to injection. A 100 µL injection loop was used for sample injections and was rinsed with 2 – 4 mL of

18.2 Ω-cm dIH_2O, followed by 2 – 4 mL of the running buffer prior to each sample injection. Running buffer was 100 % of the specified buffer also used to prepare each individual sample, 20 mM Tris, pH 7.4, with either 0, 50, or 150 mM trace metal grade NaCl in 18.2 Ω-cm dIH_2O and 0.2 μm filtered prior to use. Columns were pre-equilibrated with > 1 column volume of the running buffer. Flow rate was set to 0.75 mL/min during sample runs, with total run volumes of 26 mL. Prior to storage, columns were cleaned overnight using a rinsing method with a 4 hr rinse with 18.2 Ω-cm dIH_2O, followed by a 4 hr rinse with 20 % ethanol in 18.2 Ω-cm dIH_2O. Prior to use of columns after long-term storage, running buffer was used to pre-equilibrate the columns using a flow rate of 0.5 mL/min for > 2 hrs.

3.2.4 Native PAGE

Precast native PAGE gels were used, with 20 μL of 50 μM Gla-Oc in 20 mM Tris, pH 7.4, with 12.5 % glycerol (or percentage specified in **Figure 3.3**) and 0.005 % bromophenol blue. Each gel included a lane that was loaded with 10 μL of protein standard. Gels were run with 120 V for 1 hr in a 4 °C cold room with 25 mM Tris base in 200 mM glycine running buffer, stained overnight using 0.05 % (w/v) Coomassie dye in 30 % methanol and 10 % acetic acid, and de-stained with 18.2 Ω-cm dIH_2O overnight, or until sufficient destaining had taken place, prior to imaging using a Gel Doc EZ Imager, Bio-Rad.

To better observe the low-abundance protein bands, gels were silver stained following the protocol described below. First, the gels were fixed by incubating the gels on a plate rocker using 50 % (v/v) ethanol in 7.5 % (v/v) glacial acetic acid in 18.2 Ω-cm dIH_2O for 30 min. Gels were washed 3 times for 2 min each using 18.2 Ω-cm dIH_2O, and then pre-treated using 50 mL of freshly prepared, room temperature 15 % (v/v) formaldehyde in 18.2 Ω-cm dIH_2O for 30 min. Gels were impregnated with 50 mL of

freshly prepared 1 mg/mL silver (I) nitrate in 18.2 Ω-cm dIH_2O for 20 min. Then, gels were rinsed with 100 mL of 18.2 Ω-cm dIH_2O twice for 5-10 seconds each. Gels were developed for less than 30 seconds using 100 mL of freshly prepared 30 mg/mL sodium carbonate with 0.002 mg/mL sodium thiosulphate in 18.2 Ω-cm dIH_2O that had been pre-chilled in an ice bath. Gels were quickly rinsed with 50 mL of pre-chilled 18.2 Ω-cm dIH_2O, followed by a 10 min incubation with pre-chilled 7.5 % (v/v) glacial acetic acid in 18.2 Ω-cm dIH_2O to stop the reaction. Finally, gels were quickly rinsed with 50 mL of 18.2 Ω-cm dIH_2O prior to imaging.

For analysis, the protein standard lanes were used to prepare a standard curve, plotting the migration distance (pixels) against the log(kDa) for each molecular weight marker. The y = mx + b trendline generated was then used to estimate the size of the protein bands in the Gla-Oc sample lanes on the gel.

3.2.5 Estimation of Concentration of Oc in Bone

The concentration of Oc in human bones will be approximated for a 65 kg individual below. First, the mass of bone is ~ 15 % of the total body mass [139], which means that the individual's bone mass is ~ 9.75 kg. Approximately 1/3 of total bone mass is made up of proteins [140], which means the total mass of protein in the bone is ~ 3.25 kg. Since Oc makes up 2 – 5 % of protein mass in the bones [1], this means that the mass of Oc is 0.065 – 0.1625 kg for this 65 kg individual. From these numbers, we first need to calculate the range of bone volumes for the 65 kg individual by solving for v in the density equation described by **Equation 3.1**.

$$d = \frac{m}{v} \qquad \textbf{Equation 3.1}$$

Where d is the density, m is the mass, and v is the bone volume. The volume bone mineral density (vBMD) ranges from 135 – 467 g/L [141] and 2 – 5.34 % of total body weight is bone mineral [142]. Thus, the bone volume ranges from 2.78 – 25.7 L.

Next, the density of Oc in the bone will be calculated, assuming a homogeneous distribution of Oc throughout the bone. Note that this is known not to be the case, with Oc primarily found on the outer surface of bone (periosteum and endosteum) but this assumption is made for simplicity purposes. Using **Equation 3.1**, along with the range of masses calculated for Oc, and the range of bone volumes calculated above, the density of Oc is calculated to be 2.53 – 58.45 g/L. Using the molecular weight of 5929.5 g/mol for human Gla-Oc, the densities of Oc can be converted to molarity, with the concentration of Oc calculated to be 427 – 9857 µM.

3.3 Results

3.3.1 Oc Complexation and Quaternary Structure Destabilization by NaCl

To evaluate the heterogeneity of the Gla-Oc quaternary structures, and purify Gla-Oc complex states if necessary, 35 µM Gla-Oc in 20 mM Tris, pH 7.4, was injected onto an FPLC with a Superdex 200 column used to separate Gla-Oc complexes based on size. As shown in **Figure 3.1a**, the Gla-Oc sample was highly homogenous in complex size, with the size calculated to be approximately 30 kDa, corresponding to a pentamer of Gla-Oc. The reproducibility of these findings is shown by the overlay of the chromatograms from two separate Gla-Oc injections (dark blue chromatographic traces). Unexpectedly, when Gla-Oc samples were prepared in the 20 mM Tris, pH 7.4, buffer containing 150 mM NaCl the chromatographic peak shifted to elute at a volume corresponding to a size of approximately 2 kDa, which again was reproducible (light blue chromatographic traces). The size was determined using the standard calibrant protein mix SEC data (**Figure 3.1b**)

to prepare a standard curve (**Figure 3.1c**), with the trendline used to estimate the Gla-Oc complex size. Note that the size of Gla-Oc monomer is outside of the range of the standard curve, with the lowest molecular weight species used to generate the trendline being myoglobulin at 17 kDa, and outside of the linear range of molecular weights that can be separated by this column.

The larger complex of Gla-Oc was visualized by native PAGE with a faint band observed at about 30 kDa (Band 1), and a more prominent band observed at about 7 kDa (Band 2) (**Figure 3.2a**). The 30 kDa band was more clearly visible after silver staining of the native gel (**Figure 3.2b**). Molecular weights of each band were determined using the standard curve shown in **Figure 3.2c**, generated using the ladder/molecular weight marker loaded alongside the sample lanes (L). Altering sample conditions with either 0 – 500 mM NaCl, or 0 – 0.5 mM $CaCl_2$, did not appear qualitatively to result in any substantial difference in the fraction of Gla-Oc found in Band 1 or Band 2. Stabilization of the ~ 30 kDa complex was attempted by altering the percentage of glycerol in solution. Glycerol is known to stabilize native protein structures into more compact states and prevent protein unfolding and aggregation [143] via a glycerol exclusion model that reduces the solvent volume available for water, forcing increased hydration at the protein surface [144]. Changes in glycerol percentage between 7.5 – 25 % (v/v) in the samples did not qualitatively affect the stability of the complex (**Figure 3.3**), although they demonstrated the reproducibility of the native gel results, with Band 1 corresponding to a 29 kDa complex of Gla-Oc, and Band 2 a 10 kDa species of Gla-Oc. Note that since the standard curve was generated with a lowest molecular weight species of 20 kDa, the accuracy of the size estimation for the species at Band 2 is low. The species present at Band 2 in both native gels is assumed to be monomeric Gla-Oc.

Since the Superdex 200 column separates molecular weight species between 10 and 600 kDa linearly, the size estimate for the lower molecular weight Gla-Oc species described above is unlikely to be accurate because it is outside of the linear range of species separated by the column. To improve the accuracy of Oc size estimates, a Superdex 75 column was used. Molecular weight estimates of Oc species described below were completed by preparing a standard curve including the data for the ovalbumin (44 kDa) and myoglobulin (17 kDa) chromatographic peaks (**Figure 3.4a**). Since the Superdex 75 column separates globular proteins and protein complexes between 3 and 70 kDa linearly, all remaining species in the standard calibrant mix could not be used to prepare a standard curve. For this reason, an in-house calibrant mixture was prepared including bovine serum albumin (BSA, 66 kDa), human insulin (5.8 kDa), and uridine (0.244 kDa) (**Figure 3.4b**). Each species was also run with individual injections to confirm proper chromatographic peak assignment (**Figure 3.4c-e**). Human insulin eluted at a volume that is suggestive of dimer formation, based on the elution volumes for ovalbumin and myoglobulin, and has been shown previously to form dimer species [107]. Therefore, the human insulin chromatographic peak was set to have a molecular weight twice that of the monomer species (11.6 kDa). Using the BSA, ovalbumin, myoglobulin, and human insulin dimer data, a standard curve was generated (**Figure 3.4f**) and used to estimate the size of the Gla-Oc species with different NaCl concentrations.

To further evaluate the effects of NaCl on Oc complexation, 50 μM Gla-Oc (**Figure 3.5a**), or 50 μM Glu-Oc (**Figure 3.5b**) samples in 20 mM Tris, pH 7.4, with 0, 50, or 150 mM NaCl, and complementary running buffer conditions, were injected onto the FPLC with a Superdex 75 column. The chromatographic traces for Gla-Oc sample runs with different NaCl conditions were overlaid in the figures to help with ease of interpretation, with the chromatographic peak for the 0 mM NaCl samples corresponding

to 28 kDa molecular weight, the 50 mM NaCl samples corresponding to 10 kDa molecular weight, and the 150 mM NaCl sample corresponding to 6 kDa molecular weight. Similar results were observed with Glu-Oc, with the 0 mM NaCl samples having two major peaks, one at 27 kDa, and the second at 21 kDa. The 50 mM NaCl samples included a chromatographic peak that corresponds to a 7 kDa molecular weight species, and the 150 mM NaCl sample includes a chromatographic peak that corresponds to a 4 kDa molecular weight species.

3.4 Discussion

3.4.1 Oc Complexation and Quaternary Structure Destabilization by NaCl

The Gla-Oc pentamer shown here is reproducibly observed with SEC and with native PAGE. This is highly suggestive that the chromatographic peak eluting at a volume associated with ~ 30 kDa molecular weight species is in fact the pentamer structure, and not an artifact caused by Gla-Oc interacting with the solid phase of the column. The two bands observed in the native gels, with a weak band corresponding to the pentamer, suggests that this pentamer structure is destabilized during native PAGE. The lower molecular weight species is likely the Gla-Oc monomer. Since the size of Gla-Oc monomer is outside the range of molecular weights used to generate the standard curve, this likely explains the inaccuracy of this size estimation. Similarly, the Gla-Oc samples run on the Superdex 200 column with 150 mM NaCl had inaccurate size estimates most likely for this same reason, as well as the fact that Gla-Oc monomer is outside the linear range of molecular weights that can be separated by this column. The use of the Superdex 75 column should more accurately estimate both the pentamer and monomer Gla-Oc, and Glu-Oc molecular weights since both species lie within the linear range of separation for this column. However, since the lowest molecular weight species used to generate the

standard curve was the insulin dimer at 11.6 kDa, the size of Oc monomers were still extrapolated outside the range used to generate the standard curve, and thus error is still expected. This is a limitation of the techniques used here, and thus further research should be conducted for conclusive data on the complex sizes to be achieved.

Based on the SEC runs using the Superdex 75 column, Gla-Oc appears to be a pentamer in the absence of NaCl, and destabilizes first to approximately dimer species, and finally to monomer with increasing concentrations of NaCl. Glu-Oc is found in both the pentameric and tetrameric states in the absence of NaCl, suggesting there may be a γ-carboxylation dependency for improved pentamer stabilization. Similar to Gla-Oc, with increasing concentrations of NaCl, Glu-Oc is destabilized to dimer and monomer species. This suggests that higher order complexation of Oc is dependent on NaCl concentration.

3.5 Conclusions

Although the precise size and structure of the Oc self-assemblies were not well characterized here, the reproducible observation of larger Oc complexes is shown. The concentrations of Oc in the bone were estimated here to range from roughly 400 μM to nearly 10 mM. This is highly suggestive that Oc will adopt higher-order structures *in vivo*, although since experiments were conducted in an *in vitro* setting here, these structures have yet to be confirmed in bone samples. The destabilization of Gla-Oc pentamer, and Glu-Oc pentamer and tetramer, with increasing NaCl concentrations is an unusual occurrence. Much of the sodium content of bone is inaccessible at neutral pH, with decreased pH causing resolubilization of Na^+, suggesting Na^+ present in the bone is incorporated in the hydroxyapatite matrix rather than free in the ECF [145]. However, ECF normally contains on average 136 mM Na^+ with normal physiological conditions resulting in only 1 – 2 % variation in concentration [138], which is thought to be tightly

regulated using Na^+ stores in the bone [138]. Thus, although the pentameric structure of Gla-Oc is observed here in an *in vitro* setting, it is unlikely to be biologically relevant in the ECF within the bone. It does highlight the importance of use of near physiological Na^+ concentrations on maintenance of proper monomeric Oc structure, and thus these findings were of importance when designing experiments for **Chapter 5**.

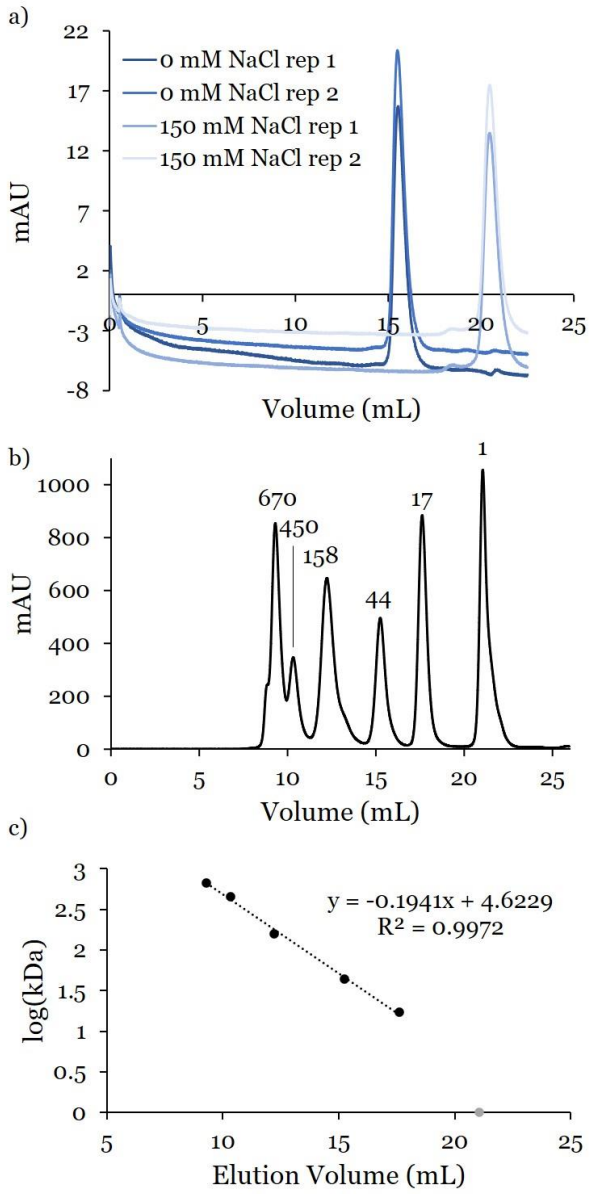


Figure 3.1. Initial SEC Runs Using a Superdex 200 Column. a) Overlaid individual 35 µM Gla-Oc sample injections with either 0 mM or 150 mM NaCl, pH 7.4, in both the samples and the running buffers. Injections of samples with either condition completed in duplicate. b) Injection of commercial standard calibrant mix containing thyroglobulin (670 kDa), immunoglobulin G (158 kDa), ovalbumin (44 kDa), myoglobulin (17 kDa), and vitamin B12 (1.35 kDa). Chromatographic peaks are labelled by the molecular weight (kDa) of the species that corresponds to each peak separation. c) Standard curve generated by plotting the elution volume versus the log(kDa) for each species found in b). The trendline was made based on the species that separate in the linear range limits of the column and include the 17 – 670 kDa species (black data points), while the 1 kDa species, vitamin B12, was omitted from the trendline for this reason (grey point).

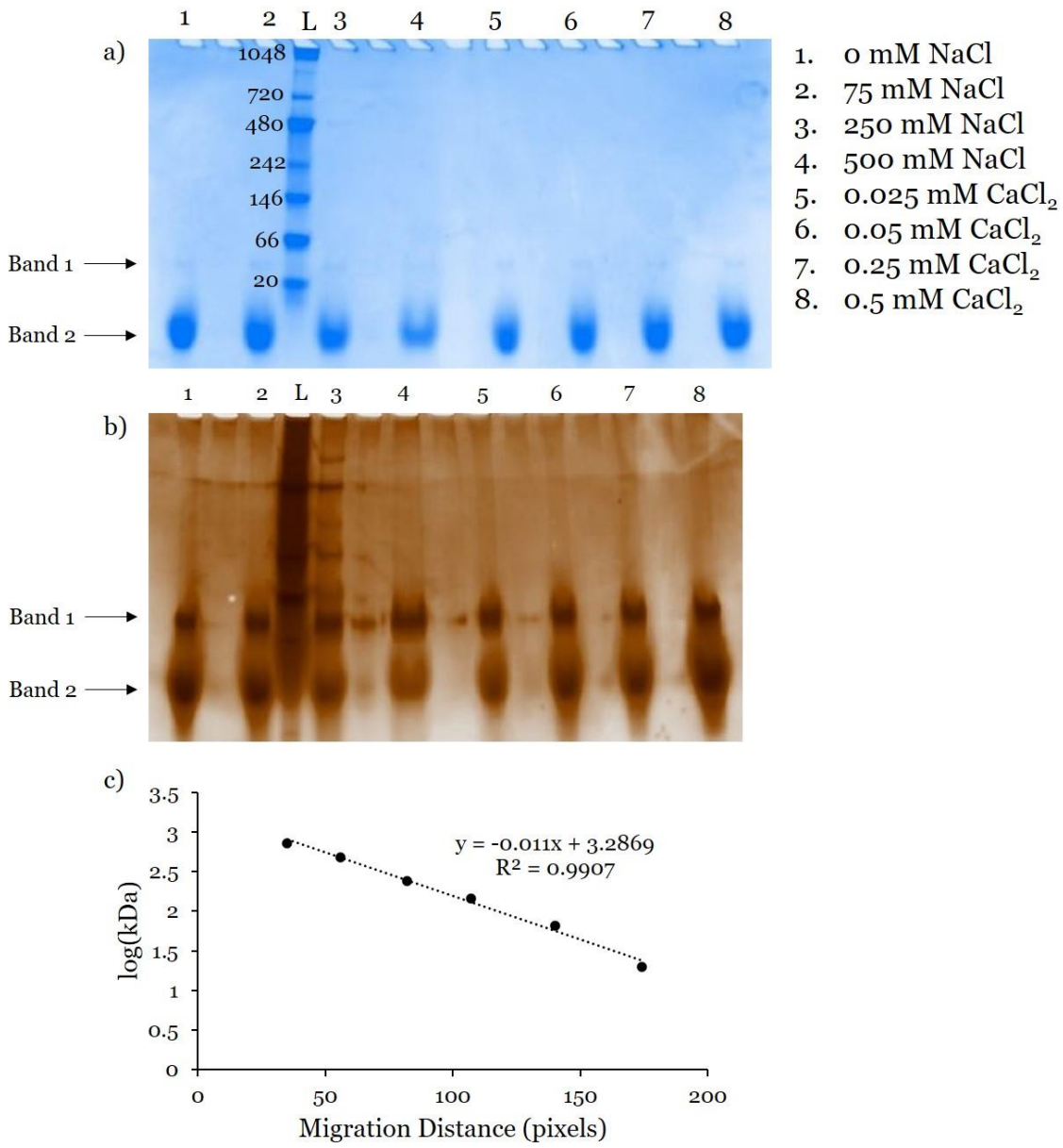


Figure 3.2. Native PAGE of Gla-Oc with Changes in NaCl and $CaCl_2$ Concentrations in the Samples. a) Native PAGE of 20 µL of 50 µM Gla-Oc samples stained with Coomassie dye for initial visualization of the protein bands reveals a faint band (Band 1), and a primary band (Band 2). Samples were prepared in 20 mM Tris, pH 7.4, with the specified concentrations of NaCl or $CaCl_2$. L = protein ladder, with the molecular weights specified in kDa. The silver-stained gel shown in b) confirms the presence of Band 1 in all samples loaded onto the gel. c) The standard curve generated from the protein ladder, with migration distance measured in pixels plotted against the log(kDa), and a trendline generated as shown in the figure. The trendline was used to estimate the size of Band 1 = 30 kDa, and Band 2 = 7 kDa.

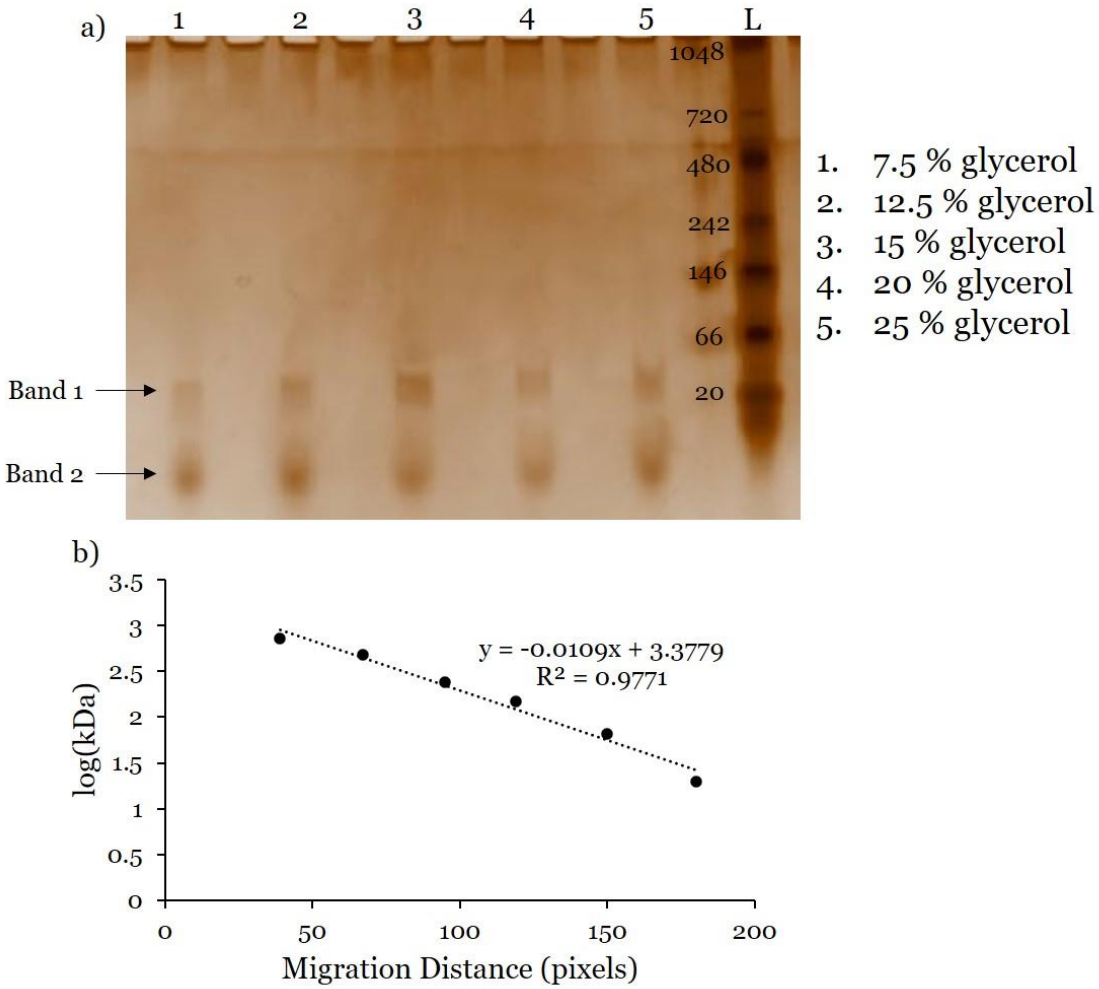


Figure 3.3. Silver-Stained Native PAGE of Gla-Oc Samples with Different Percentages of Glycerol. a) Samples prepared in 7.5 – 25 % glycerol did not significantly improve the stability of the pentameric Gla-Oc species. L = protein ladder, with the molecular weights specified in kDa. Based on the standard curve shown in b), Band 1 = 29 kDa, and Band 2 = 10 kDa.

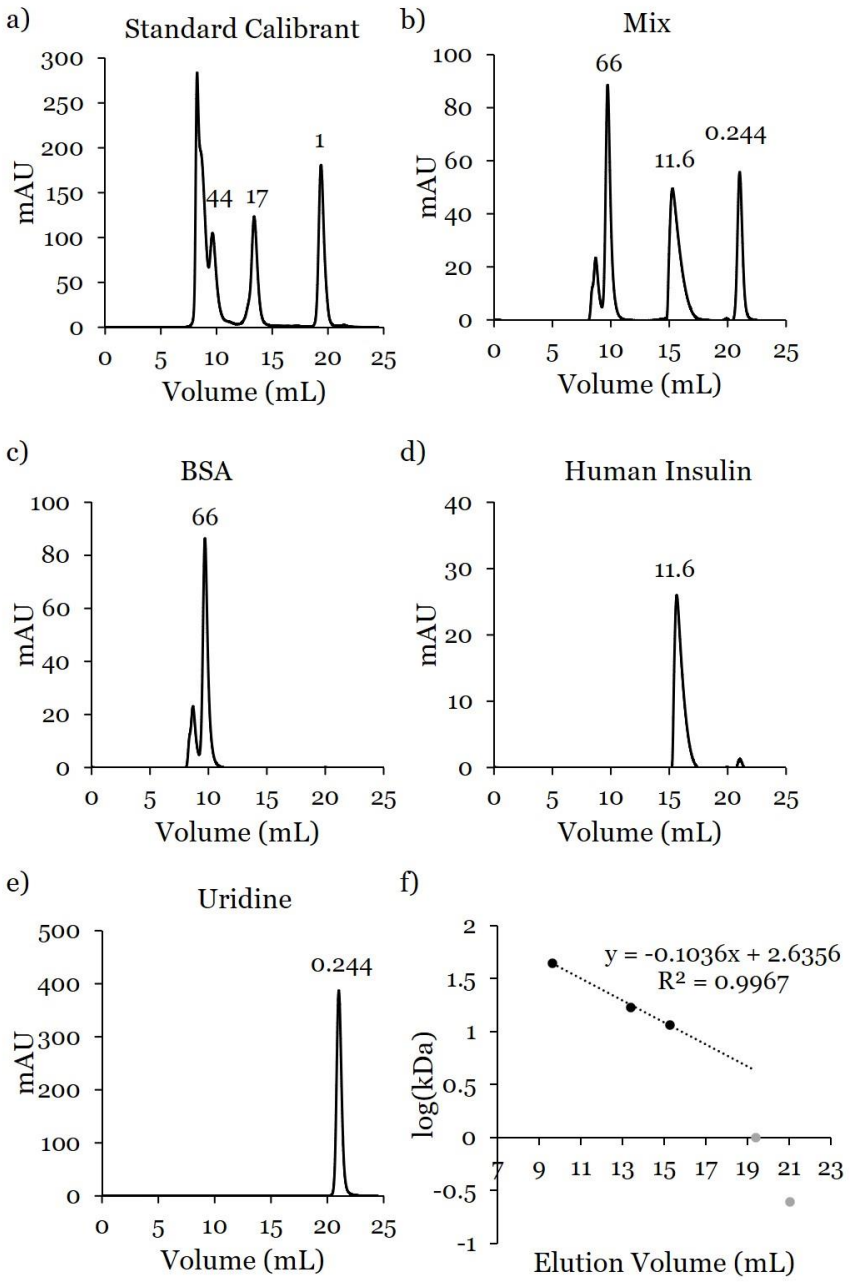


Figure 3.4. SEC Runs Using a Superdex 75 Column. a) Chromatogram of the standard calibrant mix including thyroglobulin (670 kDa), immunoglobulin G (158 kDa), ovalbumin (44 kDa), myoglobulin (17 kDa), and vitamin B12 (1.35 kDa), with the molecular weight species within the linear separation range of the column labelled with their molecular weights (kDa). b) In-house calibrant mix containing BSA (66 kDa), human insulin dimer (11.6 kDa), and uridine (0.224 kDa). Elution volumes for each species were confirmed by individual injections, as shown in c)-e). f) Standard curve generated using ovalbumin, myoglobulin, and insulin dimer data points determined from a) and b) (black data points), with the data points for vitamin B12 and uridine excluding due to the known linear separation limits of the column (grey data points).

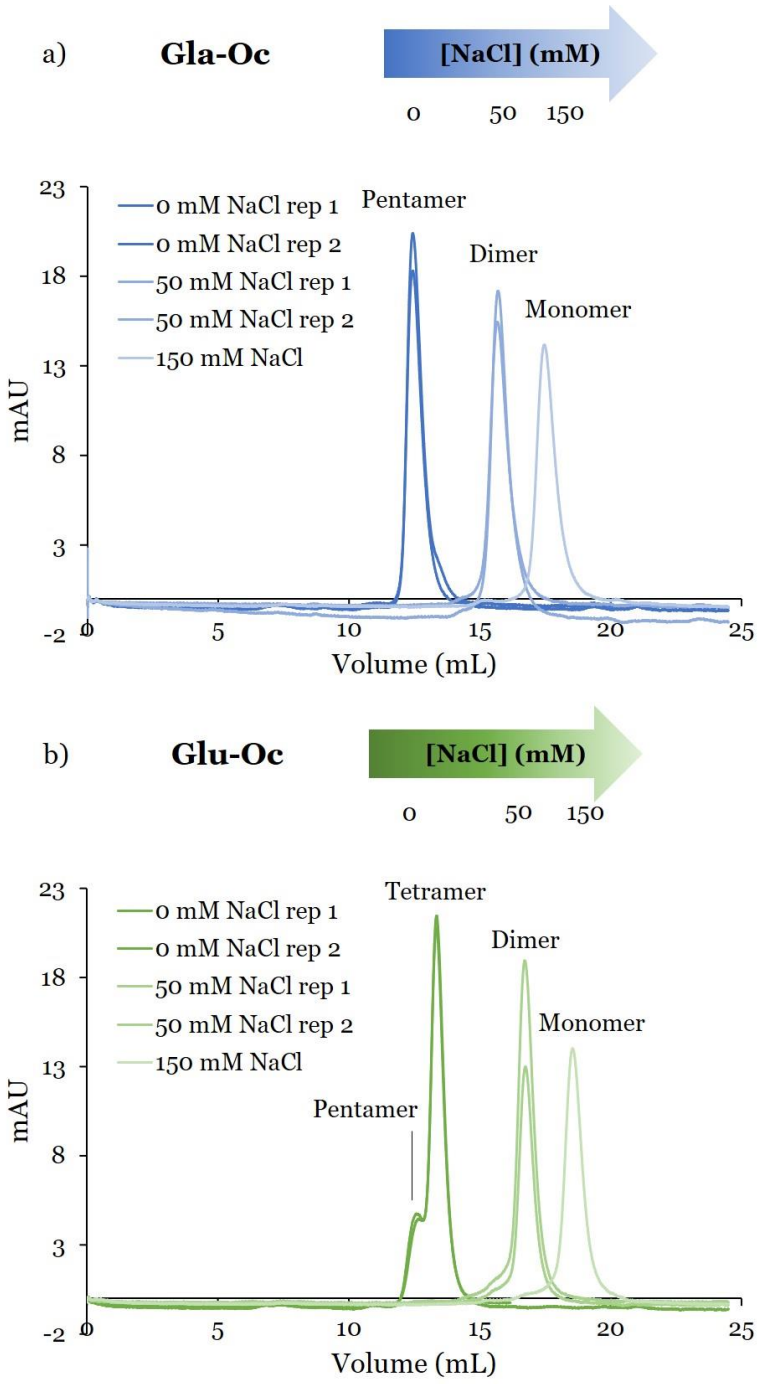


Figure 3.5. Oc Complexation States with Varying NaCl Concentrations. a) Overlaid individual SEC injections of 50 µM Gla-Oc in 20 mM Tris, pH 7.4, with 0, 50, or 150 mM NaCl in samples and in the running buffers. Complex size determined based on the standard curve in **Figure 3.4f** and specified above each corresponding chromatographic peak. Complementary experiments completed with 50 µM Glu-Oc SEC injections shown in b).

CHAPTER 4

CALCIUM-INDUCED LOSS OF OSTEOCALCIN FROM SOLUTION AND FILAMENTATION IS BOTH POST-TRANSLATIONAL MODIFACTION- AND pH-DEPENDENT

4.1 Introduction

Protein filamentation is a valuable biological mechanism for regulating both structural and enzymatic functions of proteins. In addition to their ability to form structural frameworks within cells, as is the case with actin and tubulin filaments, non-cytoskeletal proteins also form filaments both during normal physiological conditions, and in response to cellular stress. Cytoskeletal actin and tubulin filaments are involved in cell structure and cell movement and have been extensively studied [146], with "treadmilling" [147] and "dynamic instability" [148] controlled movement mechanisms dependent on the binding and hydrolysis of nucleotides by individual subunits. Although less studied, enzymatic filaments have been observed for over the past 50 years *in vitro* using techniques such as electron microscopy (EM) [149-164], with over 20 different enzymes now confirmed to form filaments with structures resolved, and another 91 enzymes with suspected filamentation capability [165]. Due to the surge in interest in X-ray crystallography, much of the work on protein filaments came to a halt due to the vast majority being incapable of forming well-ordered protein crystals. Recently, however, many have shown enzyme filamentation *in vivo* using fluorescently-labelled proteins or detection antibodies [166-173], resulting in the revival of interest in the topic.

Filament forming proteins span a wide array of cellular and extracellular processes, including metabolic pathways (glycolysis, and H_2O_2, nitrogen, glutamate, and nucleotide metabolism), biosynthetic pathways (fatty acid, pyrimidine, nucleotide, amino acid, and protein synthesis), host defense, cellular detoxification, fermentation, cellular stress response and unfolded protein response, homologous recombination, DNA

replication, cell cycle control, differentiation and proliferation, apoptosis, gene expression, and the innate immune response [165]. These are just some of the processes filament-forming proteins take part in, and exemplify the commonality of this phenomenon in biology.

Formation of protein filaments has been found to be regulated by several different factors, including substrate binding, product formation, binding to allosteric effectors, ligands, post-translational modification, and buffer conditions. Glucokinase is a kinase that initiates glycolysis in fungi by an ATP-dependent reaction that results in phosphorylation of glucose to form glucose-6-phosphate. In the presence of substrates ATP and glucose, glucokinase filaments form when the protein concentration is above the critical concentration of 2 μM, resulting in inactivation of the enzyme, and self-regulation of rate of catalysis [174]. Casein kinase 2 is a serine/threonine kinase that can form oligomeric species including heterotetramers, ring structures, and linear polymers. The oligomeric states of the enzyme are affected by ionic strength, Mg^{2+} concentration, pH, and temperature [175-177], with the ring structures associated with increased enzyme activity [176]. In the case of acetyl-CoA carboxylase, filamentation is induced by high citrate concentrations, resulting in increased activity of the enzyme [161, 164]. However, another filament structure can occur when the BRCT domains of BRCA1 bind to phosphorylated acetyl-CoA carboxylase, which is thought to result in enzyme inactivation [178]. Of the filament-forming enzymes that are confirmed to modulate enzyme activity, 13 have been demonstrated to be activated in the filament state, while 6 are catalytically inactive when in the filament state [165].

Filamentation is a diverse tool used by biology with structural, storage, and regulatory purposes. In the case of the hydrogen-dependent CO_2 reductase, filamentation occurs in the presence of $MgSO_4$ and appears to be used to protect against enzyme

degradation and oxidation [179]. In homologous recombination, recombinase A filamentation occurs when bound to single-stranded DNA and ATP and increases activity, while ATP hydrolysis causes dissociation of the filaments and reduced enzyme activity [180, 181]. In the bone, the self-assembly of type I collagen forms extended protein fibers used as a scaffold for bone mineralization [182]. With diversity in the processes proteins involved in filament formation take part in, the method used to regulate filamentation, and the ultimate function of protein filamentation, this field of research offers an exciting opportunity to better understand the interplay between protein structure and function.

Although osteocalcin (Oc) is one of the ten most abundant proteins in the human body, and the second most abundant protein in the bone, its precise function is still not well characterized. Presented here, we found that Ca^{2+}-dependent loss of in-solution Oc concentrations is associated with protein filamentation. The aims of this chapter were to i) assess the effect of increased Ca^{2+} titrations on the loss of Oc from solution, ii) evaluate the effect of changes in solution pH and post-translational modification of Oc on Ca^{2+}-dependent Oc loss from solution, iii) visualize Oc filaments using transmission electron microscopy (TEM), and iv) evaluate the reversibility of Oc filamentation. To the best of our knowledge, this is the first time Oc has been shown to form filaments and may offer a starting point to better understanding the function and/or possible mechanism of regulation of function of this protein in bone.

4.2 Materials and Methods

4.2.1 Materials

Both the Des-Gla synthetic human osteocalcin (Glu-Oc), Cat # AS-65307-025, and [Gla 17, 21, 24] – synthetic human osteocalcin (Gla-Oc), Cat # AS-22830, were purchased from AnaSpec, Co. Ammonium acetate (≥ 98 % purity), ammonium hydroxide, calcium chloride desiccant (96+ %, ACS reagent), guanidine hydrochloride, sodium chloride

(≥ 99.5 % purity), sodium hydroxide (minimum 98 % purity), Tris base (≥ 99.9 %), and Tris hydrochloride (≥ 99.0 % purity) were purchased from Sigma-Aldrich, Co. A 1 mm pathlength quartz cuvette with a 10 mm thick base, to reduce sample volume, and concentrated cuvette cleaner solution were purchased from Starna Cells, Inc. The small-volume 0.2 μm filters were purchased from Pall, Corp. Formvar, carbon-coated copper, 400 mesh grids and uranyl acetate dihydrate were purchased from Ted Pella, Inc. LC-MS grade acetonitrile was purchased from Fisher Scientific. Uranyl formate, hydrate, was purchased from Electron Microscopy Sciences, Inc.

4.2.2 Preparation of Osteocalcin Stocks

When preparing Oc stock solutions in 20 mM Tris + 150 mM NaCl, pH 7.4, the Gla-Oc lyophilized powder was reconstituted initially with 40 μL of 100 mM NaOH, swirled gently to ensure full reconstitution of the protein powder, and then the final volume was brought to 500 μL with 20 mM Tris + 150 mM NaCl, pH 7.4. The pH was verified using short range pH paper and adjusted with small volumes of HCl if necessary. For the Oc stock solutions in 5 mM ammonium acetate + 150 mM NaCl, pH 4.5, Gla-Oc vials were reconstituted by first adding 40 μL of 1 % ammonium hydroxide, the vial was swirled gently to ensure all powder was fully reconstituted, and the final volume was brought to 500 μL with 5 mM ammonium acetate + 150 mM NaCl, pH 7.4. The pH was verified prior to dilution using short range pH paper, ensuring pH did not exceed 7.4. Note that the Oc stock solutions prepared in ammonium acetate buffer were stored at higher pH (~ 6 – 7.4) due to observed decreased shelf-life of Oc stocks prepared at pH 4.5 (data not shown). For the Glu-Oc vials, the reconstitution technique was identical as described for Gla-Oc vials above, except the volume of NaOH/ammonium hydroxide and final volume were half that of the Gla-Oc vials (20 μL, 250 μL) to account for there being half the weight of lyophilized protein. Stock Oc

sample concentrations were then determined by measuring absorbance at 280 nm using a Thermo Scientific Multiskan Go UV-vis spectrophotometer, with Oc's extinction coefficient known, $\varepsilon = 1.33\ (mg/mL)^{-1}cm^{-1}$ [104].

4.2.3 Absorbance Spectrophotometry

For the Ca^{2+} titration experiments, all Oc samples were prepared at roughly 20 μM starting protein concentration in 300 μL sample volume in 20 mM Tris + 150 mM NaCl, pH 7.4, or 5 mM ammonium acetate + 150 mM NaCl, pH 4.5, in triplicate. Corresponding triplicate buffer blank samples were also aliquoted with starting volumes of 300 μL. Samples were centrifuged for 5 minutes at 17, 000 x g in a tabletop microcentrifuge, and the top 200 μL of sample volume was aliquoted into a UV 96-well microplate. Absorbance at 280 nm was measured for each sample using a Thermo Scientific Multiskan Go UV-vis spectrophotometer, with Oc's extinction coefficient known, $\varepsilon = 1.33\ (mg/mL)^{-1}cm^{-1}$ [104], to determine total starting concentrations of Oc in each sample prior to $CaCl_2$ titrations. Samples were all carefully transferred back to their respective microcentrifuge tubes using gel loading pipette tips. $CaCl_2$ stocks prepared at 30.5 – 2500 mM were used to titrate 5 – 40 μL of $CaCl_2$ stock into the Oc sample to arrive at final $CaCl_2$ concentrations ranging from 0.5 – 500 mM in the samples. After each $CaCl_2$ titration, samples were centrifuged for 5 minutes at 17, 000 x g to separate in-solution Oc from precipitated Oc, and the top 200 μL of sample volume was aliquoted into the UV 96-well microplate. No incubation between $CaCl_2$ titrations was required, as increased incubation times did not affect the amount of protein precipitated (data not shown). To reduce protein loss due to adsorption to surfaces, the same gel loading pipette tips were used throughout the entire experiment, labelled for each specific sample. Samples were also added to the same well in the microplate, specific for each sample, during each $CaCl_2$ titration. To account for reduced

in-solution Oc concentrations due to dilution alone, side-by-side sample titrations were completed with buffer without $CaCl_2$ added, with corresponding buffer blanks. These samples were used to determine the total concentration of Oc during each sample titration, and to calculate the fraction of Oc precipitated when compared to concentrations of Oc in the $CaCl_2$ titrated samples.

To further test the effect of pH on the loss of Oc from solution, 20 μM Gla-Oc, and Glu-Oc, samples with 500 mM $CaCl_2$ were prepared in i) 5 mM ammonium acetate + 150 mM NaCl at pH 4.5, 5.0, 5.5, 6.0, and 6.5, or ii) 20 mM Tris + 150 mM NaCl at pH 6.0, 6.5, 7.0, and 7.4. Control samples without $CaCl_2$ added were used to calculate the total concentration of in-solution Oc at each pH condition. All samples were incubated overnight at room temperature, followed by centrifugation for 5 minutes at 17, 000 x g, and measurement of absorbance at 280 nm. To ensure samples reached equilibrium they were incubated for an additional 24 hours, centrifuged, and measurements were re-taken. The fraction of Oc precipitated was calculated by subtracting the concentration of in-solution Oc in the samples containing excess $CaCl_2$ from the total concentration of Oc, determined using the control samples without $CaCl_2$ added, and then dividing this value by the total concentration of Oc.

4.2.4 Circular Dichroism

In-solution Oc secondary structure in the presence and absence of Ca^{2+} was evaluated by preparing 50 μM Gla-Oc, or Glu-Oc, in 20 mM Tris + 150 mM NaCl, pH 7.4. Gla-Oc was titrated with 0, 0.05, 0.1, 1, 5, and 10 mM $CaCl_2$, and centrifuged for 5 minutes at 17, 000 x g prior to analysis by circular dichroism (CD). Glu-Oc was titrated with 0, 0.05, 0.1, 1, 3, 5, 10, 25, and 50 mM $CaCl_2$ and followed identical preparation and analysis procedures.

For the reconstituted Oc precipitate samples, initial in-solution Oc CD analyses were completed using i) 50 μM Gla-Oc in 20 mM Tris + 150 mM NaCl, pH 7.4, ii) 75 μM Gla-Oc in 5 mM ammonium acetate + 150 mM NaCl, pH 4.5, and iii) 125 μM Glu-Oc in 20 mM Tris + 150 mM NaCl, pH 7.4. Samples were then titrated with $CaCl_2$ to a final concentration of 500 mM, centrifuged for 5 minutes at 17, 000 x g to remove the supernatant, and the precipitates were reconstituted in the appropriate buffer without $CaCl_2$ added.

All sample runs were completed on a Jasco J-815 circular dichroism spectrometer after a 15 minute N_2 purge, followed by 30 minutes lamp warm-up. Instrument parameters were optimized to be a 1 nm bandwidth, 5 nm/min scanning speed, and 8 s time constant, with a single acquisition per sample replicate. Sample data were acquired every 0.5 nm from 190 – 260 nm wavelengths in a 1 mm pathlength quartz cuvette. Buffer blank runs were completed in between each sample run and included the appropriate $CaCl_2$ concentrations. After completion of each sample run, the cuvette was washed 3 times with deionized H_2O, followed by 2 drops of cuvette cleaning solution in 400 μL of water, 3 times with deionized H_2O, 3 times with 6 M guanidine-HCl, 3 times with deionized H_2O, and 3 times with 100 % acetonitrile and air dried prior to running the next buffer blank. In between each sample run, the Oc samples were centrifuged for 5 minutes at 17, 000 x g and absorbance at 280 nm was measured to determine the in-solution Oc concentrations to convert instrument mdeg units to mean residue ellipticity (MRE = $degcm^2dmol^{-1}res^{-1}$). Data were smoothed using an 11-point moving average smoothing technique and plotted in GraphPad Prism 7.

4.2.5 Transmission Electron Microscopy

Samples of precipitated Gla-Oc were prepared as described in **4.2.4**, using 500 mM $CaCl_2$, in 20 mM Tris, pH 7.4, centrifuged for 5 minutes at 17, 000 x g, the supernatant was removed, and precipitated Gla-Oc filaments were mixed with 20 µL of buffer containing 500 mM $CaCl_2$. Then, 3.8 µL aliquots of the sample were spotted on the carbon-coated formvar grid, incubated for 1 minute, followed by 4 wash steps with 18.2 Ω-cm dIH_2O, once with the staining solution, followed by a 1 minute incubation with the negative stain. The negative stain used was either 0.75 % uranyl formate (w/v), or 2 % uranyl acetate (w/v), prepared fresh daily and filtered using a 0.2 µm syringe filter prior to use. Grids were analyzed using a Philips CM-12 TEM, with a tungsten filament and high-tension voltage set to 80 kV.

4.3 Results

4.3.1 Effect of Ca^{2+} Titrations, pH, and γ-Carboxylation Status on Oc Solubility

The effect of increased Ca^{2+} concentrations on Oc solubility was evaluated using absorbance spectrophotometry at 280 nm wavelength, with the results depicted in **Figure 4.1**. Gla-Oc at pH 7.4 (**Figure 4.1a**, p-value < 0.0001), Gla-Oc at pH 4.5 (**Figure 4.1b**, p-value = 0.0001), and Glu-Oc at pH 7.4 (**Figure 4.1c**, p-value = 0.0002) were all found to have a statistically significant difference in in-solution Oc concentrations for samples titrated with Ca^{2+}, compared to those titrated with buffer only when comparing the average difference between the series using a two-tailed Wilcoxon matched-pairs signed rank test**.** In contrast, Glu-Oc at pH 4.5 (**Figure 4.1d**) titrated with Ca^{2+} did not have a statistically significantly different in-solution concentration compared to Glu-Oc titrated with buffer in the absence of Ca^{2+} (p-value = 0.1324). All sample data were also compared to the

expected Oc concentrations due to dilution alone, as determined using the average concentration of Oc prior to Ca^{2+} titrations in each condition.

To further evaluate the effect of pH on loss of Oc from solution, Gla-Oc, and Glu-Oc, samples were prepared in buffers at pH values between 4.5 – 7.4 with excess Ca^{2+}. In-solution protein concentrations were evaluated by measuring absorbance at 280 nm and demonstrate the increased Ca^{2+}-dependent loss of in-solution Oc with increased pH (**Figure 4.2**). The fraction of Gla-Oc compared to Glu-Oc precipitated was statistically significantly different at pH 6.5, 7.0, and 7.4 for samples in 20 mM Tris with 150 mM NaCl (paired t test, $p < 0.05$, **Table 4.1**).

The secondary structure of in-solution Gla-Oc (**Figure 4.3a**) and Glu-Oc (**Figure 4.3b**) was evaluated by circular dichroism (CD) spectroscopy and demonstrates a clear increase in α-helical content with increased titrations of Ca^{2+}. This finding is in good accordance with previously published literature [28, 29].

4.3.2 Visualization of Precipitated Oc Filamentation by TEM

To visualize the Ca^{2+}-dependent precipitated Oc, a sample was prepared by precipitating Gla-Oc with excess Ca^{2+} at pH 7.4, spotted on a carbon-coated formvar grid, and imaged using negative stain TEM. The Gla-Oc sample contained higher-order filamentous structures that appeared like pearls on a string (**Figure 4.4a-c**), with the individual pearls ranging in diameter from approximately 15 to 30 nm. A grid spotted with a sample prepared with excess Ca^{2+} in the absence of Gla-Oc did not display any observable filamentous structures (**Figure 4.4d**). Replicates of the Ca^{2+}-dependent Gla-Oc filaments prepared at pH 7.4 demonstrated the reproducibility of filamentation, but the heterogeneity of the number of connected strings of pearls that may be formed (**Figure 4.5a**), with one, two, or four adjacent strings observed in the replicate samples. At lower

magnification, Gla-Oc filaments can be seen to extend for several μm, involving an intricate meshwork of strands of filaments (**Figure 4.5b**).

4.3.3 Resolubilization of Oc Filaments

To assess the reversibility of Oc precipitates, Gla-Oc at pH 7.4 and 4.5, and Glu-Oc at pH 7.4 precipitates were prepared by titrating samples with excess Ca^{2+}. Buffer with or without excess Ca^{2+} was then added to the precipitate and mixed thoroughly by pipetting. After separation of any remaining precipitate from resolubilized Oc by centrifugation, the nanomoles of in-solution Oc was determined and compared to the nanomoles prior to addition of excess Ca^{2+}, and after generation of precipitated Oc with excess Ca^{2+}. The methods described here are also depicted in **Figure 4.6**. The amount of Oc at each step is depicted in **Figure 4.7a**, with the initial amount, precipitated amount, and reconstituted amount of Oc for each sample condition shown. The data are then re-expressed in **Figure 4.7b** to depict the fraction of precipitated Oc that was reconstituted in buffer containing excess Ca^{2+} compared to buffer without Ca^{2+} added. The difference between fractions of precipitated Oc that were reconstituted in buffer with or without excess Ca^{2+} added was evaluated for statistical significance using the Welch's unpaired t test and it was determined that Gla-Oc at both pH 7.4 ($p = 0.0007$) and 4.5 ($p = 0.014$) had statistically significantly higher reconstitution of precipitate using the buffers without Ca^{2+} added when compared to the samples reconstituted using buffers that included excess Ca^{2+}.

Samples reconstituted in buffer without Ca^{2+} added were analyzed by CD to determine the secondary structure of reconstituted Oc. All the reconstituted Oc samples were found to have increased α-helical content compared to Oc prior to precipitation with excess Ca^{2+} (**Figure 4.7c**).

4.4 Discussion

4.4.1 Effect of Ca^{2+} Titrations, pH, and γ-Carboxylation Status on Oc Solubility

From the data presented here, it is clear that loss of in-solution Oc occurs in a Ca^{2+}-dependent fashion, and that this event is further affected by solution pH and the γ-carboxylation status of Oc, with increased Ca^{2+}-dependent precipitation observed with increased pH and with Gla-Oc compared to Glu-Oc. The control samples where Oc is titrated with buffer without Ca^{2+} added accounted for any protein loss due to adsorption, as well as precipitation of Oc due to other factors, such as any Oc denaturation during the experimental procedure. Since the isoelectric point (pI) for Glu-Oc is 4.6 [110], the observation of significant precipitation of Glu-Oc at pH 4.5 with buffer titrations is suspected to be due to disruption of proper secondary and/or tertiary Glu-Oc structure at this pH due to loss of surface charge of the protein. In contrast, the pI for Gla-Oc is 4.0 [110], which may be the reason for the maintenance of in-solution Gla-Oc concentrations at pH 4.5, although this would need to be evaluated further before any conclusions may be drawn.

The maximum fraction of Oc precipitated in the presence of excess Ca^{2+} is both pH and Oc γ-carboxylation status dependent, with increased precipitation at higher pH and with Gla-Oc compared to Glu-Oc. The reason for this plateau in loss of in-solution Oc with excess Ca^{2+} is likely due to an equilibrium state between in-solution and precipitated Oc under each experimental condition. The precipitation maxima presented here may not be a true reflection of biological maxima, although the experimental set-up enables comparison between experimental conditions tested. As was discussed extensively in **Chapter 2**, the binding affinity of Ca^{2+} by Oc is improved with increased γ-carboxylation of Oc, as this increases the ability of Oc to coordinate with metal ions in a bidentate

fashion. In addition, the charge state of the Gla residues is highly dependent on the pH, and thus the ability of these residues to coordinate with metal ions is also affected by pH. Taken together, this data is highly suggestive of the importance of Ca^{2+} coordination with Oc on Oc's propensity for precipitation and lends to the hypothesis that precipitation of Oc with Ca^{2+} is biologically relevant. The observation that in-solution Gla-Oc and Glu-Oc consists of increased α-helical secondary structure with increased Ca^{2+} titrations, which matches previously published literature [28, 29], suggests that the results presented here are not due to improperly folded protein and thus supports the biological relevance of these findings.

4.4.2 Visualization of Precipitated Oc Filamentation by TEM

Based on the negative stain TEM images obtained of the Gla-Oc filaments, there are multiple conclusions that can be drawn. First, these filamentous structures are Oc specific, since they are only visualized in the presence of Gla-Oc, but not in the sample prepared with $CaCl_2$ alone. Next, the formation of these filaments involves hundreds of Gla-Oc monomers. With the diameter of the individual pearl-like structures ranging from 15 to 30 nm, this corresponds to roughly 240 to 2000 homomeric complex structures of Gla-Oc. This implies that Ca^{2+} binding to Gla-Oc may induce higher-order complexation to form these filamentous structures. In fact, Gla-Oc dimerization has already been shown to occur with increased affinity in the presence of Ca^{2+} [32] and taken together with the observation that in the presence of Ca^{2+}, Gla-Oc dimers are absent from the mass spectra shown in **Chapter 2** this is suggestive that these Gla-Oc dimers may be building blocks for the filaments observed here. The mechanism of monomer to filament transition was not addressed here, although it would be an interesting potential future direction of the work. Finally, the heterogenous nature of the filaments observed is suspected to be due to sample

preparation on the grid, with the uranyl formate and uranyl acetate stains having low pH (~ 4) to prevent precipitation of the stain, but this would have to be validated through the use of stains compatible with neutral pH prior to making any definitive conclusions.

4.4.3 Resolubilization of Oc Filaments

The observation that proteins can form filaments, and that these structures are important for regulation of protein and enzymatic function is becoming a rapidly evolving field in part due to improvements in TEM and cryo-EM techniques, which, unlike X-ray crystallography and NMR, enables one to solve less ordered protein structures. To demonstrate the biological relevance of these protein filament structures, it is important to either show their formation and dissolution in an *in vivo* setting, or at the very least, demonstrate their ability to disassemble *in vitro*. Similar to Narayanaswamy, *et al* [166], we test here the ability of the Oc filaments, which precipitate out of solution, to disassemble and resolubilize when excess Ca^{2+} is removed from solution *in vitro*. Since Oc is an extracellular protein, testing filament formation in a cell-based *in vivo* system was not evaluated. Based on the results presented here, all Oc precipitates were able to be resolubilized in buffer without excess Ca^{2+}. The fractions of precipitated Oc resolubilized with buffer without Ca^{2+} added were statistically significantly higher compared to the fractions resolubilized using buffer with excess Ca^{2+} for the Gla-Oc samples at pH 7.4 and 4.5. This further exemplifies the importance of solution Ca^{2+} levels on the equilibrium between in-solution Oc and Oc precipitation and filamentation. The re-solubilization of Glu-Oc precipitated samples at pH 7.4 with buffer without excess Ca^{2+} was not statistically significantly different from the fraction of precipitated Glu-Oc resolubilized with buffer with excess Ca^{2+}. This is likely due to the low abundance of the precipitate in these samples, resulting in increased overall error in the experimental measurements.

To further test the biological relevance of precipitated Oc, the secondary structure of Oc resolubilized from precipitates using buffer without excess Ca^{2+} was evaluated to determine whether these Oc samples maintained similar secondary structure as observed with in-solution Oc. All resolubilized Oc precipitate samples were found to contain increased α-helical structure, suggestive of Ca^{2+} binding, and maintenance of normal secondary structural behaviour. It should be noted that comparison of the degree of secondary structural components between different conditions tested cannot be made here, as these variances likely arise from differences in quantities of Oc precipitate, Oc concentration during resolubilization, and in-solution Ca^{2+} concentrations. It will, however, be noted that all conditions tested resulted in consistent maintenance of some degree of increased α-helical structure compared to Oc's degree of α-helical structure prior to addition of excess Ca^{2+} and Oc precipitation. This suggests that with all conditions tested here, the resulting resolubilized Oc precipitates maintain their binding capability with Ca^{2+} and proper secondary structure.

4.5 Conclusions

Osteocalcin filamentation is ligand (Ca^{2+}), post-translational (γ-carboxylation), and buffer (pH) dependent. The observation that these filaments are reversible suggests they may have a biological role, although the function of these filaments was not addressed here. Since the function of osteocalcin is still not completely understood, the full significance of this finding is unfortunately not realized, although a hypothesis can be developed, as discussed below. Free Ca^{2+} in the extracellular fluid (ECF) is tightly maintained between 1.1 – 1.4 mM to maintain normal physiological functions [31]. This is regulated by parathyroid hormone (PTH), where hydroxyapatite matrix is used as a storage system for Ca^{2+} that can be accessed during times of need. The challenge with this system is the fact

that hydroxyapatite has low solubility at neutral pH. Based on the reported solubility product constant for hydroxyapatite ($Ca_5OH(PO_4)_3$, $K_{sp} = 3.73 \times 10^{-58}$ [183]), at pH 7.4, the solubility limit for free Ca^{2+} is only ~ 0.54 μM. This suggests that hydroxyapatite alone cannot regulate free Ca^{2+} in the ECF at normal physiological levels and that other factors must be present in the bone that shift the free Ca^{2+} solubility limit to 1.1 – 1.4 mM. The rate of calcium influx into the bone has been quantified previously as 15 $pMcm^{-2}s^{-1}$ [184] and is suggestive that the use of osteoclast-based bone resorption as a method to access Ca^{2+} reserves only accounts for a small percentage of the total Ca^{2+} resorption that takes place, since the maximum amount of Ca^{2+} resorption by osteoclasts is ~ 500 mg/day in humans [31]. The rapid, almost instantaneous, correction of plasma Ca^{2+} concentration is primarily adjusted using bone Ca^{2+} stores without the use of bone remodeling [184, 185]. Additionally, parathyroid hormone (PTH) was not found to be involved in the rapid correction of Ca^{2+} concentrations in the ECF by Dedic, *et al* [186], and the authors suggested that for this rapid correction of free Ca^{2+} concentrations in the ECF to take place, there may be a protein present in the osteocyte lacunae forming "coastal crystals" of Ca^{2+} linked to a protein template. In line with this hypothesis, Talmage & Mobley [31], suggested Oc may play a role in the process of maintaining physiological Ca^{2+} in the ECF by improving the solubility limits of free Ca^{2+} through its binding to the surface of hydroxyapatite. Due to the Ca^{2+}-dependency of filamentation observed here, and the reversibility of this process, it can be hypothesized that Oc filamentation is used as an easily accessible storage method of bone Ca^{2+} that can be rapidly tapped into when ECF free Ca^{2+} concentrations are depleted. Of course, this is purely speculation at this time, and would require further studies for conclusions to be drawn, but this offers a potential biological role for this filamentation process. Whatever the case, the finding that Oc can form protein filaments adds yet another protein to the list of hundreds of proteins already

found to form filaments that span structural and enzymatic functions throughout the body.

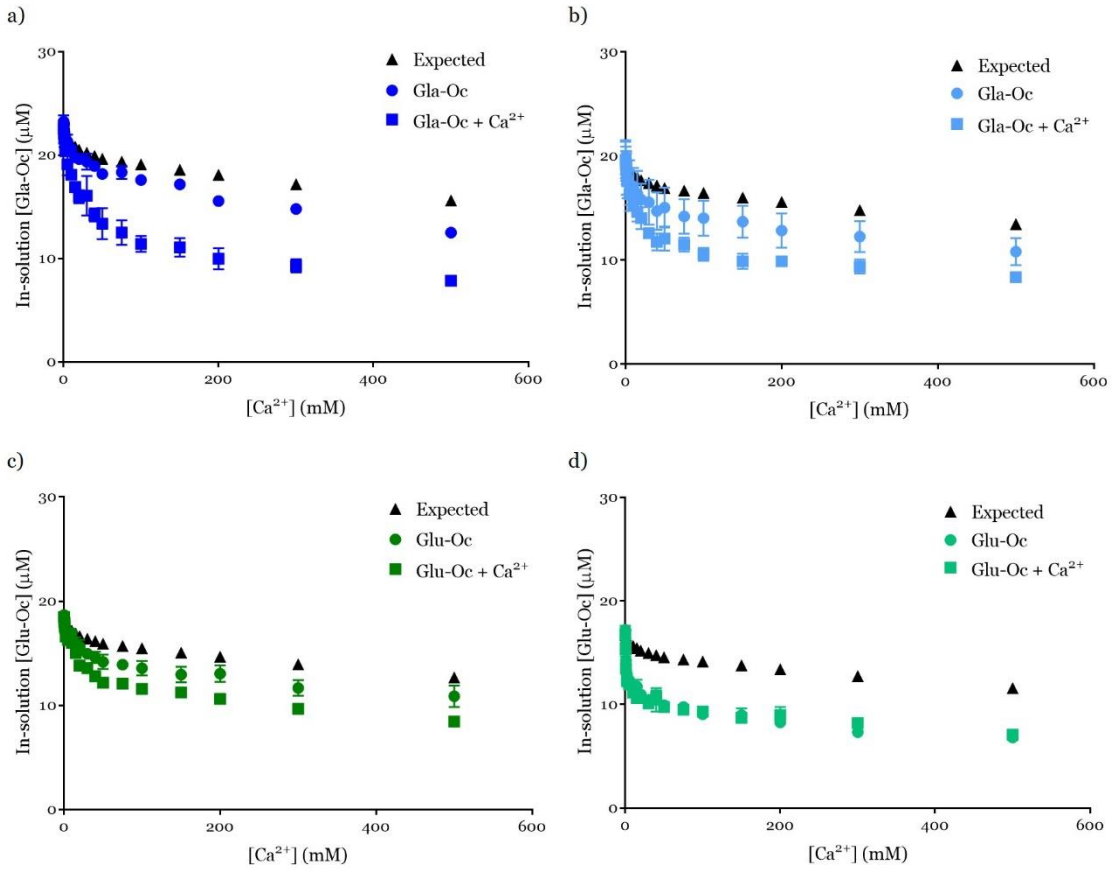


Figure 4.1. In-Solution Oc Concentrations with Increased Calcium Titrations. In-solution concentrations of Gla-Oc in a) 20 mM Tris + 150 mM NaCl, pH 7.4, or b) 5 mM ammonium acetate + 150 mM NaCl, pH 4.5, and Glu-Oc in c) 20 mM Tris + 150 mM NaCl, pH 7.4, or d) 5 mM ammonium acetate + 150 mM NaCl, pH 4.5. Each sample set includes a data set depicting the expected concentration (µM) of Oc due to dilution alone (black triangles), the experimental concentration of Oc when titrated with buffer (circle datasets), and the experimental concentration of Oc when titrated with Ca^{2+} in the specified buffer (square datasets). Total Ca^{2+} concentrations were between 0 – 500 mM. n = 3. Error bars represent SD. Significant differences between the average difference between the the buffer titrated and Ca^{2+} titrated data series were found for Gla-Oc at pH 7.4 (p-value < 0.0001) and 4.5 (p-value = 0.0001) and Glu-Oc at pH 7.4 (p-value = 0.0002) but not for Glu-Oc at pH 4.5 (p-value = 0.1324) using a two-tailed Wilcoxon matched-pairs signed rank test.

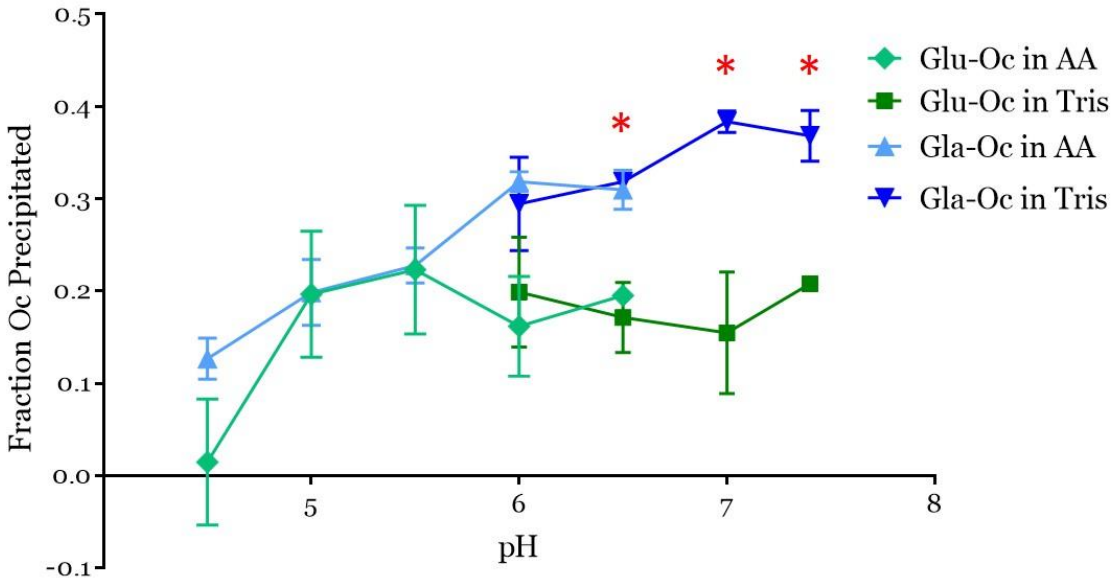


Figure 4.2. Fraction of Oc Precipitated at Various pH Values Between pH 4.5 – 7.4. Glu-Oc (green traces), or Gla-Oc (blue traces) tested between pH 4.5 – 6.5 were prepared in 5 mM ammonium acetate + 150 mM NaCl (AA), while samples tested between pH 6.5 – 7.4 were prepared in 20 mM Tris + 150 mM NaCl (Tris). Control samples were prepared in the absence of Ca^{2+} to determine the total concentration of Oc at each pH value. All samples with added Ca^{2+} were completed with excess Ca^{2+} (500 mM). The Gla-Oc fraction precipitated in Tris at pH 6.5, 7.0, and 7.4 were statistically significantly different from the fractions of Glu-Oc precipitated from complementary pH samples (student paired t-test; p-value < 0.05), as indicated in **Table 4.1**. n = 3 and error bars represent SD.

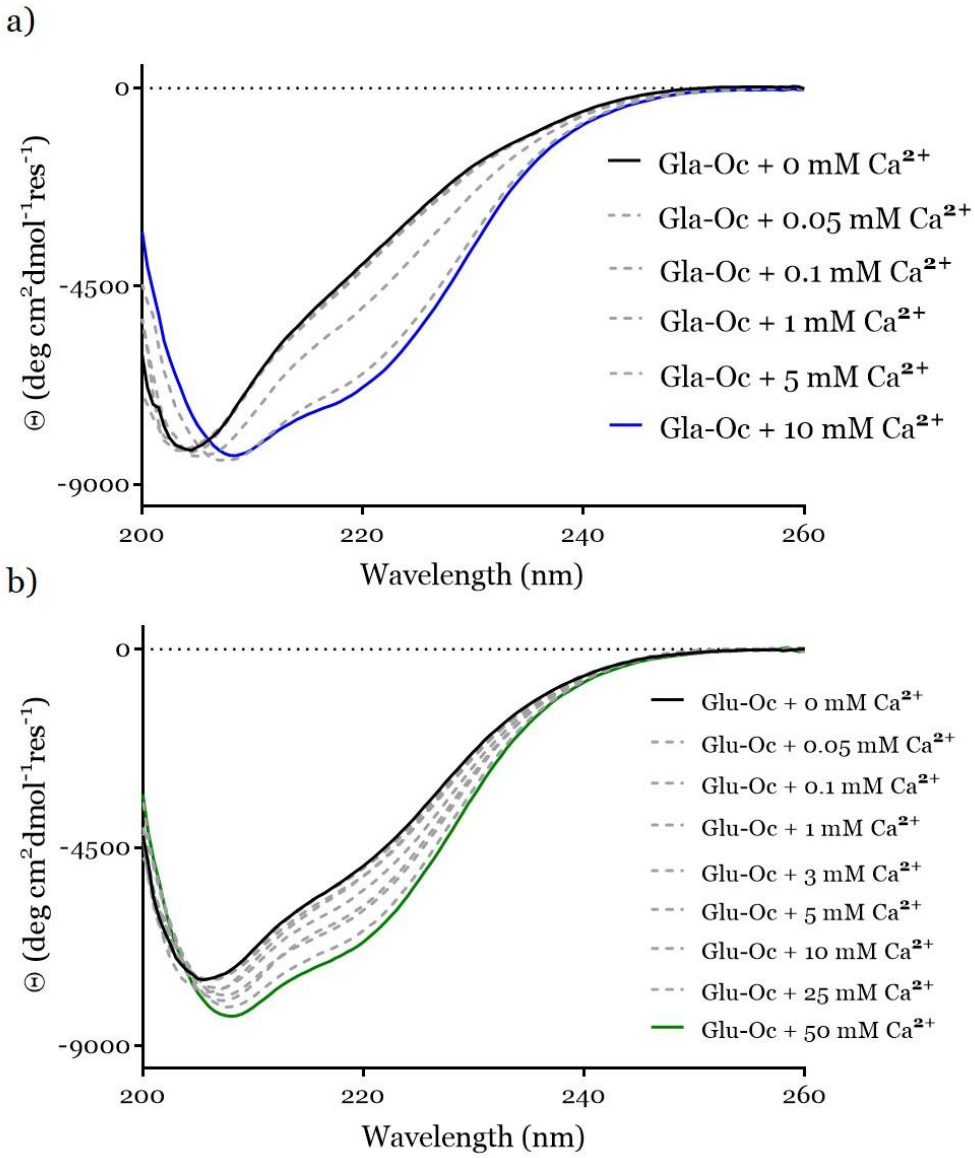


Figure 4.3. CD Spectroscopy of In-Solution Oc with Calcium Titrations. Ellipticity (Θ, degcm^{2}dmol^{-1}res^{-1}) of in-solution a) Gla-Oc, or b) Glu-Oc, with 0 – 50 mM Ca^{2+} between 200 – 260 nm. n = 1.

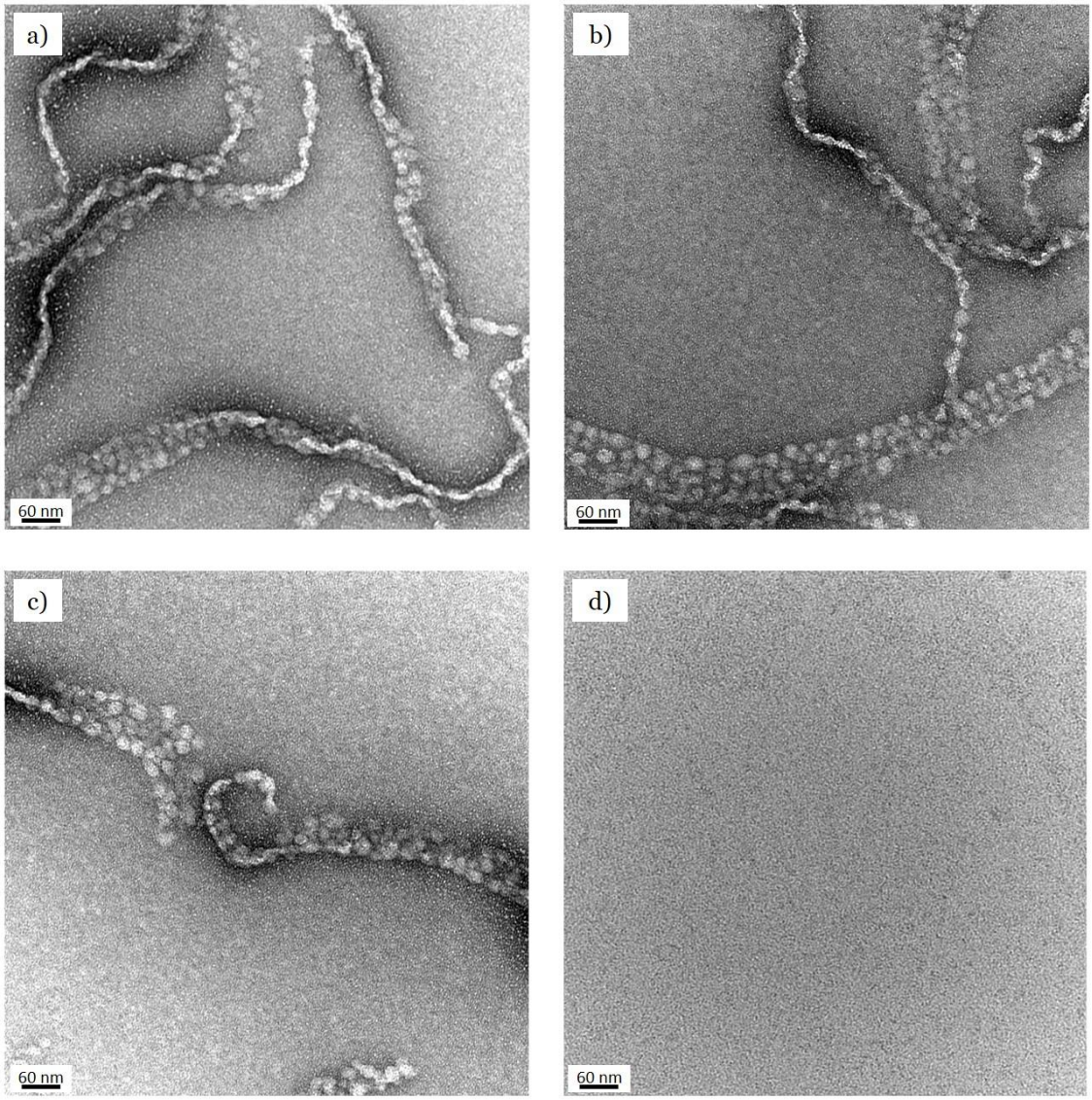


Figure 4.4. Negative Stain TEM Images of Ca^{2+}-Dependent Gla-Oc Filaments. Gla-Oc precipitated in the presence of 500 mM $CaCl_2$ at pH 7.4 was spotted on a carbon-coated formvar grid and stained using 0.75 % uranyl formate. Filaments are imaged in a)-c) at 110,000 x magnification. Image of grid shown in d) was obtained from the negative control sample containing 500 mM $CaCl_2$ only. Scale bars are 60 nm.

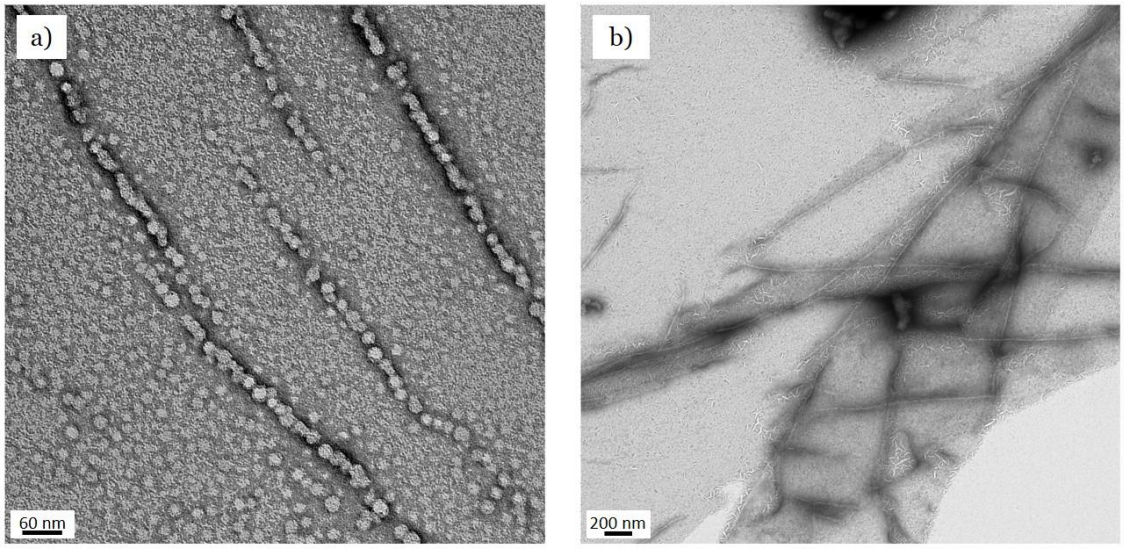


Figure 4.5. Replicates of Ca^{2+}-Dependent Gla-Oc Filamentation. Separate replicates of Gla-Oc precipitated with 500 mM $CaCl_2$ at pH 7.4 were imaged by negative stain TEM, with a) at 110,000 x magnification and a 60 nm scale bar, and b) at 25,000 x magnification and a 200 nm scale bar shown.

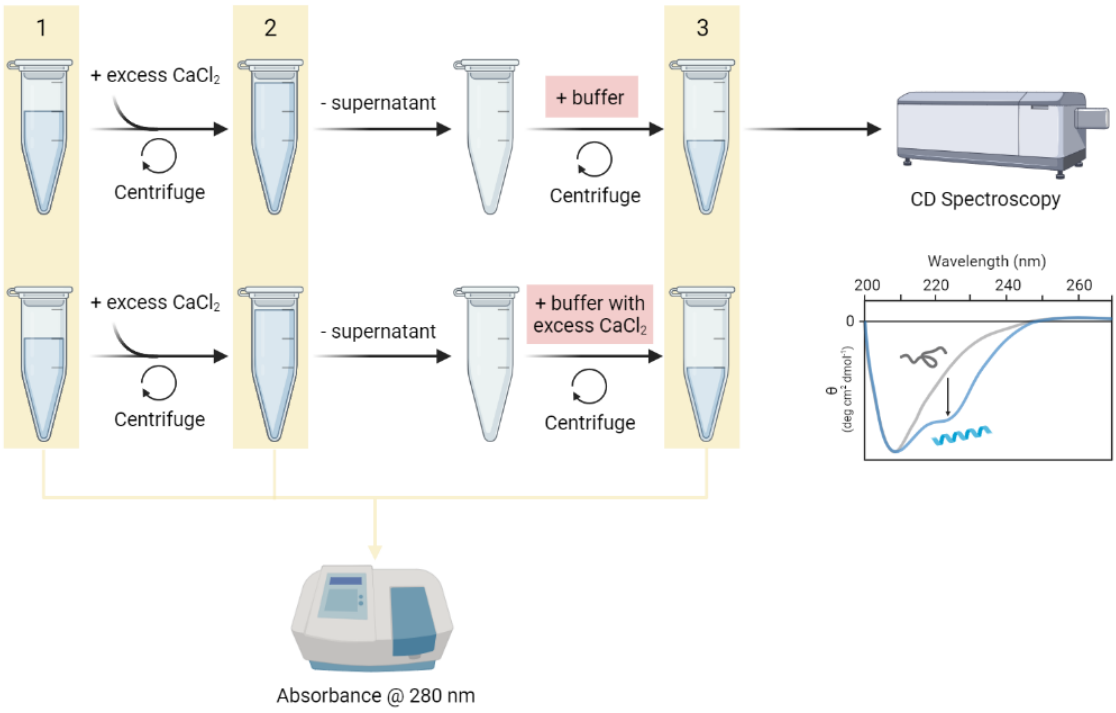


Figure 4.6. Schematic Diagram of the Procedure Used to Evaluate the Reversibility of Oc Precipitation. First, Oc samples were prepared in the appropriate buffer, centrifuged, and initial in-solution protein concentrations were determined by measuring absorbance of the supernatant at 280 nm (1). All samples were then titrated with a final concentration of 500 mM $CaCl_2$ and centrifuged to separate in-solution Oc from the precipitate. The in-solution Oc concentration was again determined by measuring absorbance at 280 nm (2), allowing for the nanomoles of Oc precipitated to be calculated. The remaining in-solution Oc and buffer solution was removed from the samples, and then either i) buffer, or ii) buffer containing 500 mM $CaCl_2$ was added, and the samples were mixed thoroughly by pipetting. Finally, the samples were centrifuged again to separate in-solution Oc from precipitate, and the concentration of in-solution Oc was determined by measuring absorbance at 280 nm (3). The Oc samples reconstituted in buffer without $CaCl_2$ added were used for further downstream analysis using CD spectroscopy.

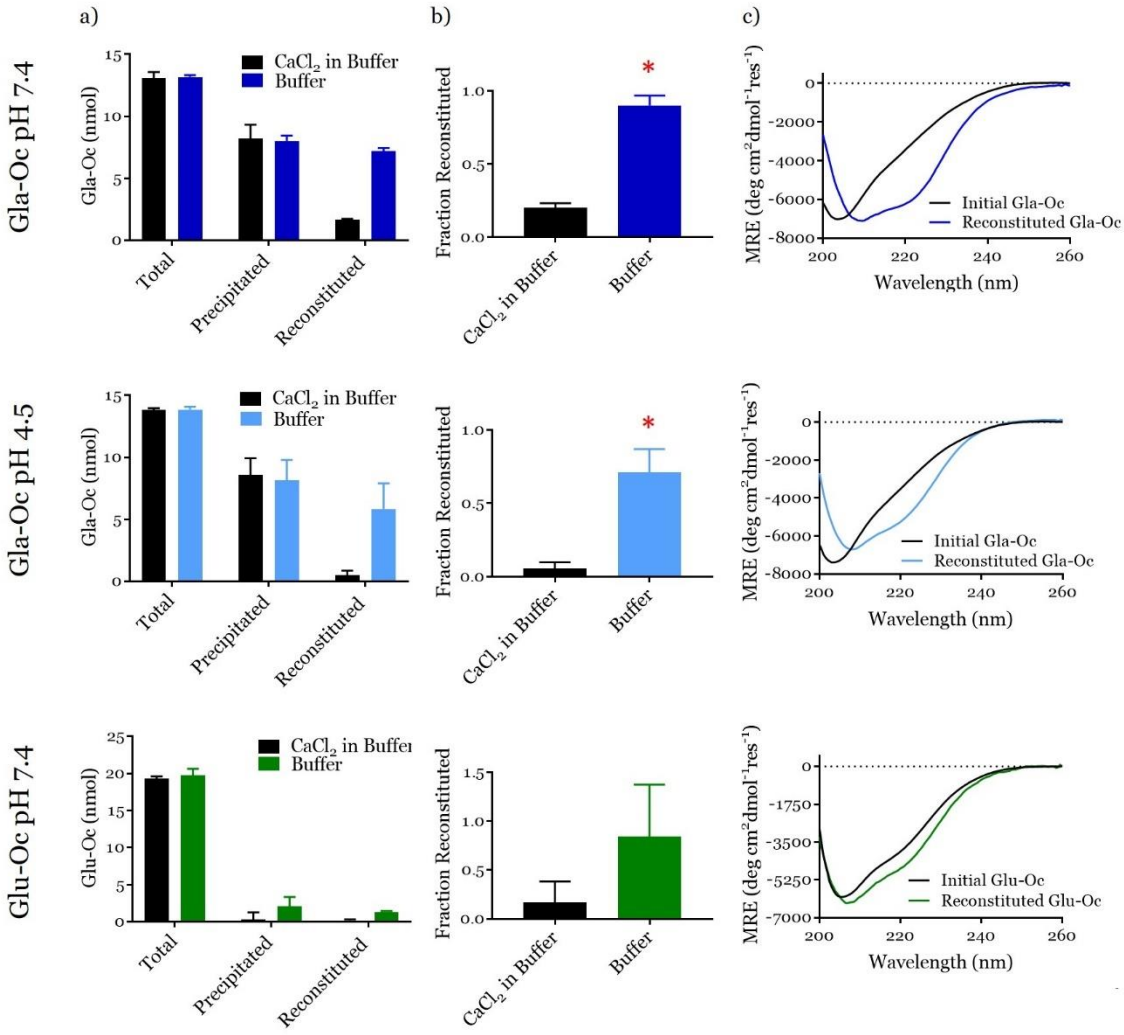


Figure 4.7. Reversibility of Oc Filamentation. a) Nanomoles (nmol) of in-solution Oc prior to addition of 500 mM Ca^{2+} (Total), precipitated Oc after addition of 500 mM Ca^{2+} and centrifugation of sample (Precipitated), and in-solution Oc after mixing the precipitate thoroughly and centrifuging again (Reconstituted) with either 500 mM Ca^{2+} in buffer (black bars) or buffer without Ca^{2+} added (coloured bars). b) Fraction of Oc reconstituted with 500 mM Ca^{2+} in buffer, or buffer only, with the difference being statistically significantly different for Gla-Oc at pH 7.4 and 4.5 (Welch's unpaired t-test; p-value < 0.05). c) CD spectra of Oc prior to addition of 500 mM Ca^{2+} (Initial), and after reconstitution of Oc samples with buffer without Ca^{2+} added (Reconstituted). Experiment completed for Gla-Oc at pH 7.4 (dark blue data sets), Gla-Oc at pH 4.5 (light blue data sets), and Glu-Oc at pH 7.4 (green data sets). n = 3 and error bars represent SD.

Table 4.1. Fraction of Oc Precipitated with 500 mM Ca^{2+} at Various pH Values. The data presented in **Figure 4.2** is presented here, with the average fraction of precipitated Gla-Oc, or Glu-Oc, in 5 mM ammonium acetate + 150 mM NaCl (AA), pH 4.5 – 6.5, and 20 mM Tris + 150 mM NaCl (Tris), pH 6.0 – 7.4, presented. SD = standard deviation, n = 3. P-values < 0.05 are indicated by the asterisk (*).

		Fraction Precipitated				
		Gla-Oc		Glu-Oc		
	pH	Avg	SD	Avg	SD	p-value
AA	4.5	0.121	0.056	0.015	0.068	0.2636
	5.0	0.143	0.060	0.197	0.068	0.4169
	5.5	0.244	0.059	0.223	0.070	0.8026
	6.0	0.262	0.018	0.162	0.054	0.1377
	6.5	0.253	0.031	0.195	0.011	0.1385
Tris	6.0	0.300	0.042	0.213	0.027	0.1160
	6.5	0.458	0.019	0.193	0.070	0.0354*
	7.0	0.418	0.011	0.173	0.080	0.0377*
	7.4	0.371	0.008	0.233	0.029	0.0171*

CHAPTER 5

DETERMINATION OF THE GLA-OC : ALUMINUM EQUILIBRIUM BINDING CONSTANT

5.1 Introduction

Aluminum (Al) is the most abundant metal in the Earth's crust and was made accessible primarily due to mining during the industrial revolution and bioavailable during acid rain. Al is ubiquitous throughout the environment, found in soil, dust particles, plants, and natural water systems [187]. However, natural forms of Al in the soil, for example aluminosilicates, are poorly soluble at neutral pH, and therefore are not readily absorbed. The use of Al as a clarifying agent in drinking water (~ 70 μg/L), in many foods including processed grains, salt, where it is used as an anti-clumping agent, tea, and processed cheese (≤ 50 mg/slice), and as a packing agent has dramatically increased human exposure to Al [93]. Of the Al humans are exposed to, about 0.1 % is absorbed by the gastrointestinal (GI) tract and 0.012 % is absorbed transdermally (e.g., 50-75 mg/day exposure due to aluminum chlorohydrate present in antiperspirants). Sources of citrate, for example in orange juice, increase Al bioavailability, whereas silicates and phosphates, for example phosphates present in dairy products, decrease bioavailability due to their low solubility and thus decreased absorption by the GI tract. The vast majority of Al, 95 %, is filtered out of the body through the kidneys by glomerular filtration, with the remaining 5 % distributed throughout the body with half of this being stored in the bones with a half-life of 7 years [93].

Although there is no known biological function of Al, the effects of Al on normal brain function span over 200 biological processes. Al affects protein phosphorylation and dephosphorylation, causes abnormal protein aggregation and oligomerization, which

affects the normal protein degradation process, and alters gene expression [188]. These perturbations induce apoptosis of neurons and glial cells, ultimately resulting in neuronal signaling imbalances [189]. Chronic low dose Al exposure is also suspected to be linked to neurodegeneration and Alzheimer's disease [189]. Additionally, acute exposure to Al is widely accepted as toxic, with chronic kidney disease (CKD) patients undergoing dialysis in the 1970's and 1980's experiencing dialysis-induced encephalopathy due to Al contamination in dialysis solutions from tap water [93, 99, 190].

In patients with CKD, Al has also been shown to cause osteodystrophy presented as osteomalacia and adynamic bone disease due to reduced osteoblast activity and bone mineralization [191]. *In vivo* studies demonstrate Al reduces bone mineral density and mineral and trace element content in rats [192]. Studies *in vitro* revealed Al causes osteoblast apoptosis by increasing phosphorylation, and thus activity, of the oxidative stress-mediated c-Jun N-terminal kinase (JNK), increasing mRNA and protein expression of c-Jun, increasing expression of the pro-apoptotic caspase 3, caspase 9, bax, and factor-related apoptosis ligand genes and decreasing the anti-apoptotic Bcl-2 gene expression [193]. In addition, Al inhibits hydroxyapatite crystal growth *in vitro* and is incorporated into the bone matrix resulting in structural changes to the mineral deposits [98]. Al^{3+} has also been shown to bind to Oc, with the K_d for single-site binding to bovine Oc being approximately 83 nM [95].

Rather unexpectedly, during analysis of a cohort of non-diabetic CKD patient plasma samples to evaluate the correlation between CKD and glucose homeostasis and circulating osteocalcin carboxylation and fragmentation status described by Kratz, *et al* [45], it was discovered that these purified Oc samples were complexed with Al^{3+} with about 5 to 50 % of 3-Gla-Oc bound by Al^{3+} (**Figure 5.1**). After completion of follow-up control

experiments, it was determined here that the Al^{3+}-bound Oc observed in these patient samples was, at least to some degree, due to Al contamination in the buffers used during the Oc purification procedure (**Figure 5.2**). Thus, whether the Oc bound to Al^{3+} in these patient samples was indicative of circulating Oc bound to Al^{3+} within the patients' blood could not be confidently determined here.

Based on the previously reported equilibrium dissociation constant for bovine Oc binding to Al^{3+}, this suggested that Oc would be significantly bound by Al^{3+} during *in vitro* studies. Thus, the main motivation for this work was to determine the dissociation constant for Al^{3+} binding to the ***human*** unmodified form of Oc, Glu-Oc, and the vitamin K dependent post-translationally modified form, 3-Gla-Oc. Due to the low solubility of Al^{3+}, and the high level of total Al (cAl) present in the environment, a novel approach to quantitatively assess metal binding to Oc was developed here. The inclusion of citrate, a known competitive binder of Al^{3+}, was crucial to improve overall Al solubility, and to overcome experimental limits of detection and Al contamination issues in the system. Therefore, the method described here was developed to indirectly quantify the free concentration of Al^{3+} at equilibrium using circular dichroism (CD) spectroscopy to measure the concentration of Al^{3+}-bound 3-Gla-Oc (AlOc) and inductively-coupled plasma mass spectrometry (ICP-MS) to quantify cAl for each Al sample titration. Using the experimentally determined AlOc and cAl values, known total citrate (cCit) and total osteocalcin (cOc) concentrations, the pH of the system, and the association constants (K_as) for all Al-citrate, and Al-OH, complexes it was possible to solve a complex system of twenty linear equilibrium and mass balance equations to calculate the concentration of free Al^{3+} for each individual Al titration. Quantitative knowledge of free Al^{3+} values was necessary for proper non-linear regression curve fitting using a single-site binding model and estimation of the dissociation constant for Al^{3+} coordination with monomeric 3-Gla-Oc to

be K_{dAlOc} = 1.0 ± 0.12 nM at pH 4.5. A pH of 4.5 is relevant to Oc physiologically due to the decrease in pH down to pH 4.5 at the osteoclast resorption lacunae and increased Oc release into blood plasma during bone resorption [41]. Due to the qualitatively assessed lower affinity of Al^{3+} binding to Glu-Oc, along with the challenges faced with Al solubility at higher pH, and Glu-Oc solubility challenges at lower pH, complimentary K_{dAlOc} values for Gla-Oc binding Al^{3+} at pH 7.4 and for Glu-Oc binding Al^{3+} at pH 4.5 and 7.4 could not be determined.

The work presented here not only relates to the biological implications of Al^{3+} binding to Oc but is also suggestive that caution should be taken when performing *in vitro* studies of Oc, due to the ability of naturally occurring background concentration of Al^{3+} to act as a competitive binder to Oc.

5.2 Materials and Methods

5.2.1 Materials

Both the Des-Gla synthetic human osteocalcin (Glu-Oc), Cat # AS-65307-025, and [Gla 17, 21, 24] – synthetic human osteocalcin (Gla-Oc), Cat # AS-22830, were purchased from AnaSpec, Co. Aluminum chloride hexahydrate (99 % purity), ammonium acetate (≥ 98 % purity), ammonium hydroxide, citric acid monohydrate (≥ 99.5 % purity), guanidine hydrochloride, HEPES, lead(II) nitrate (99.999 % trace metals basis), sodium chloride (≥ 99.5 % purity), sodium hydroxide (minimum 98 % purity), trifluoroacetic acid (≥ 99 % purity), Tris base (≥ 99.9 %), and Tris hydrochloride (≥ 99.0 % purity) were purchased from Sigma-Aldrich, Co. Nitric acid (67-70 %), ARISTAR PLUS for trace metal analysis, was purchased from VWR. A 1 mm pathlength quartz cuvette with a 10 mm thick base, to reduce sample volume, and concentrated cuvette cleaner solution were purchased from Starna Cells, Inc. The 0.2 µm filters were purchased from Pall, Corp. Slide-A-Lyzer MINI dialysis chambers, 3.5 K MWCO, LC-MS grade acetonitrile, and MSIA tips,

underivatized, were purchased from Thermo Scientific. Mouse anti-human osteocalcin monoclonal antibody (clone # 2H9F111F8) was purchased from Bio-Rad. UV 96-well microplates were purchased from Greiner.

5.2.2 Preparation of Osteocalcin Stocks

When preparing Oc stock solutions in 20 mM Tris + 150 mM NaCl, pH 7.4, the Gla-Oc lyophilized powder was reconstituted initially with 40 µL of 100 mM NaOH, swirled gently to ensure full reconstitution of the protein powder, and then the final volume was brought to 500 µL with 20 mM Tris + 150 mM NaCl, pH 7.4. The pH was verified using short range pH paper and adjusted with small volumes of HCl if necessary. For the Oc stock solutions in 5 mM ammonium acetate, or 20 µM citrate, pH 4.5, Gla-Oc vials were reconstituted by first adding 40 µL of 1 % ammonium hydroxide, the vials were swirled gently to ensure all powder was fully reconstituted, and the final volume was brought to 500 µL with the appropriate buffer. The pH was verified prior to dilution using short range pH paper. For the Glu-Oc vials, the reconstitution technique was identical as described for Gla-Oc vials above, except the volume of NaOH/ammonium hydroxide and final volume were half that of the Gla-Oc vials (20 µL, 250 µL) to account for there being half the weight of lyophilized protein (~ 0.25 mg vs. ~ 0.5 mg). Stock Oc sample concentrations were then determined by measuring absorbance at 280 nm using a Thermo Scientific Multiskan Go UV-vis spectrophotometer, with Oc's extinction coefficient known, $\varepsilon = 1.33\ (mg/mL)^{-1}cm^{-1}$ [104].

5.2.3 Mass Spectral Analysis by Infusion, MSIA Tip Extraction & LC-MS

Infusion samples contained Gla-Oc (4 µM), human insulin (4 µM), and Al in 5 mM ammonium acetate, pH 7.4, with molar ratios of Al:osteocalcin ranging from 0.2:1 to 10:1. Human insulin was included as an internal standard to help differentiate between osteocalcin-specific and non-specific metal binding. Each sample was incubated at 25 °C

in the dark until binding equilibrium had taken place, and directly infused onto a Bruker maXis 4G (Q-TOF) MS at 5 µL/min using electrospray ionization to detect Al-bound Gla-Oc.

Synthetic serum was made using 40 mg/mL human serum albumin (HSA), 30 µM transferrin, 99 µM citrate, and 37.5 nM $AlCl_3 \cdot 6H_2O$ in 1 x phosphate buffered saline (PBS) with 1 mM $CaCl_2$ and 0.5 mM $MgCl_2$, pH 7.4. Synthetic serum was also prepared omitting $AlCl_3 \cdot 6H_2O$. Synthetic serum (750 µL) was spiked with either Gla-Oc or Al-bound Gla-Oc (2.5 nM) and the resultant samples were purified by MSIA tip extraction.

A cohort of CKD patient plasma samples were obtained for the SUGAR study published elsewhere by Kratz, *et al* [45], where full details on sample collection, storage, and analysis are described. Briefly, 87 participant samples were analyzed, 50 of which were in the CKD group, while the remaining samples were obtained from 37 age, gender, and race matched healthy controls. Samples were collected in tubes containing ethylenediaminetetraacetic acid (EDTA), plasma separated within 60 minutes, and stored at -70 °C until MSIA tip experiments were completed.

Mouse anti-human Oc antibody was coupled to MSIA tips following the protocol published previously [194] and stored at 4 °C until use. The tips were mounted on a Beckman Multimek 96-channel automated pipetting robot and pre-rinsed with 150 µL of HEPES buffered saline (HBS) 50 times, 150 µL of 50 mM HCl 50 times, and 150 µL HBS 50 times. Osteocalcin capture (from plasma samples or the synthetic Gla-Oc control samples) was completed by pipetting 150 µL of the samples 1500 times. The MSIA tips were washed 10 times with 150 µL of 100 mM Tris, pH 7.0, 150 µL of distilled water 10 times, 150 µL of 25 % acetonitrile in 1.5 M ammonium acetate 10 times, and 150 µL of distilled water 10 times. Samples were eluted in 12 µL of 0.4 % (v/v) TFA and dispensed

into a 96-conical well polypropylene autosampler tray immediately prior to LC-ESI-MS injection. A Dionex Ultimate 3000 HPLC (Thermo Scientific) was used to load samples onto a peptide concentration-and-desalting trap. Osteocalcin was eluted off the trap using 20 % aqueous solvent (0.1 % formic acid in water) and 80 % acetonitrile. The Bruker maXis 4G (Q-TOF) MS was operated in positive ion TOF-only mode. Spectra were deconvoluted and baseline subtracted using Bruker DataAnalysis v4.2 software and peak intensities were tabulated to analyze relative abundance of the mass spectral peaks.

5.2.4 Circular Dichroism Spectroscopy

In-solution Gla-Oc secondary structure in the presence and absence of Al, or Pb, was evaluated by preparing 20 μM Gla-Oc in 20 mM Tris + 150 mM NaCl, pH 7.4. Gla-Oc was titrated with either 0 – 45 μM $AlCl_3$, or 0 – 16.67 μM $PbCl_2$, and centrifuged for 5 minutes at 17, 000 x g prior to analysis by circular dichroism (CD). Absorbance at 280 nm was measured to determine the in-solution Oc concentrations to convert instrument mdeg units to mean residue ellipticity (MRE = $degcm^2dmol^{-1}res^{-1}$). The general operation and cuvette cleaning procedures follow the same protocol described in **4.2.4**.

Final CD spectra used to generate binding curve data were generated by preparing 100 μL final samples containing 50 μM Gla-Oc in 5 mM citrate + 150 mM NaCl, pH 4.5, with targeted cAl values of 0, 0.015, 0.025, 0.035, 0.05, 0.1, 0.13, 0.175, 0.25, 0.35, 0.5, 0.75, 1, 1.25, 1.5, 2, 2.5, 3, 3.5, and 3.96 mM. Al stocks were prepared at 2 x the targeted final concentrations using a 5 mM Al stock in 5 mM citrate + 150 mM NaCl, pH 4.5, to prepare serial dilutions using the buffer. All samples, buffers, and Al stocks were prepared in falcon tubes and microcentrifuge tubes that had been pre-washed with 2 % HNO_3 in 18.2 Ω-cm dIH_2O, and air dried overnight prior to use. All Al stocks and buffers were further prepared in freshly obtained 18.2 Ω-cm dIH_2O, stored in a Teflon bottle, using

trace metal grade reagents. All Gla-Oc samples had a corresponding 400 µL blank sample that was prepared with cAl and buffer to match the final cAl concentrations found in the Gla-Oc samples. Final Gla-Oc and Al samples were incubated at room temperature for 18 hrs using a plate shaker set to 250 rpm. Prior to CD analysis, samples were centrifuged for 5 min at 13, 300 rpm. Blank measurements were obtained using 100 µL aliquots of the corresponding cAl sample prior to each Gla-Oc with cAl titration sample analysis. Absorbance was measured at 280 nm, followed by a CD wavelength scan from 195 – 265 nm.

The cAl in the blank samples were then measured by ICP-MS by diluting 200 µL of these samples in 1.8 mL for a final 2 % HNO_3. Al standards were run prior to sample injections, with a rinse, blank check, and check standard run in between every 5 sample injections.

Sample spectra were blank subtracted and converted to MRE using the 50 µM Gla-Oc concentration and the number of residues being 49. To determine the fraction of Gla-Oc bound by Al^{3+}, the change in ellipticity at 222 nm ($\Delta\Theta$) was calculated by subtracting the ellipticity value at 222 nm of the 0 mM cAl sample from the ellipticity value at 222 nm obtained for each subsequent Gla-Oc sample titrated with Al. These $\Delta\Theta$ values were plotted against the cAl values measured by ICP-MS and the data were fit to **Equation 5.1**, which was described by Dowd, *et al* [29], to estimate the maximum change in ellipticity ($\Delta\Theta_{max}$) using the Matlab script in **Appendix B**. The fraction of Gla-Oc bound for each sample was then calculated using **Equation 5.2**. The concentration of AlOc was determined by multiplying the fraction of Gla-Oc bound by the cOc, and this value, along with the ICP-MS measured cAl values, were input into the Matlab script to solve for Al^{3+}

and estimate the K_{dAlOc} for each individual sample run. The data were fit using non-linear regression (NLR) curve fitting, as described in **5.2.8**.

$$\Delta\Theta = \Delta\Theta_{max} * \frac{cAl}{cAl + K_{App}} \quad \textbf{Equation 5.1}$$

$$\text{Fraction Oc Bound} = \frac{\Delta\Theta}{\Delta\Theta_{max}} \quad \textbf{Equation 5.2}$$

5.2.5 Absorbance Spectrophotometry

The methods used here follow nearly an identical protocol as described in **4.2.3**, and therefore will not be described in full here, although the few changes are mentioned below. For the Oc precipitation with Al titration experiments, Oc samples were prepared at roughly 20 μM starting protein concentration in 300 μL sample volume in 20 mM Tris + 150 mM NaCl, pH 7.4, or 50 μM Gla-Oc in 20 μM citrate, pH 4.5. All samples were completed in triplicate. Samples were titrated with Al to arrive at final Al concentrations ranging from 0 – 6.5 mM in the samples prepared in 20 mM Tris + 150 mM NaCl, pH 7.4, or 0 – 30 μM Al for samples prepared in 20 μM citrate, pH 4.5. Control samples to account for reduced in-solution Oc concentrations due to dilution alone were completed with titrations of buffer without Al added, with corresponding buffer blanks. These samples were used to determine the total concentration of Oc during each sample titration, and to calculate the fraction of Oc precipitated when compared to concentrations of Oc in the Al titrated samples.

5.2.6 Dialysis

To optimize the dialysis incubation time and ensure equilibrium binding was achieved, 75 μL aliquots of 50 μM Gla-Oc in 20 μM citrate, pH 4.5, were added to the sample

chamber compartment of 3 kDa molecular weight cut-off microdialysis chambers, and 1.225 mL aliquots of 8 μM Al in 20 μM citrate, pH 4.5, were added to the dialysate compartment in a 2 mL 96-well plate that had been acid washed with 2 % HNO_3 overnight and air dried prior to use. Samples were completed in quadruplicate and incubated for 0, 1, 2, 4, 6, 8, 12, 18, 24, or 48 hours at room temperature using a plate shaker. Complementary control samples where the sample chambers were filled with 75 μL of 20 μM citrate, pH 4.5, without Gla-Oc added, were prepared to measure cAl, and incubated for 48 hrs. The results were plotted with Al_{ICP} values against time to determine the time to reach equilibrium in the current experimental set-up, with all future experiments completed with an 18-hour incubation period.

5.2.7 ICP-MS

Prior to any dialysis and ICP-MS sample runs, all falcon tubes, microcentrifuge tubes, and 2 mL 96-well plates were soaked in 2 % HNO_3 prepared in ultra-pure 18.2 Ω dIH_2O overnight. Any tubes, and the 2 mL 96-well plates, that were to be used for buffer and sample preparation and dialysis incubation then had the 2 % HNO_3 discarded and were air dried prior to use, to ensure complete vaporization of the HNO_3 before proceeding. Falcon tubes used to hold final ICP-MS samples could be stored in 2 % HNO_3 until use. Once samples had undergone the dialysis procedure, each sample's dialysate was aliquoted into acid washed microcentrifuge tubes for storage during transport of samples to the ICP-MS core facility. Once at the core facility, the 2 % HNO_3 in the acid washed falcon tubes used for final storage of ICP-MS samples was discarded, and 1 mL of 4 % HNO_3 was added to each tube, along with 1 mL of the samples. These samples were then stored at room temperature until ICP-MS injection.

As a quality control verification, 1 mL Al samples were prepared at 0, 0.94, 1.88, 17.9, 22.6, and 27.8 μM, and mixed 1:1 with 4 % HNO_3 to make final 2 % HNO_3 solutions. Samples were completed in triplicate and injected onto a Thermo Scientific Quadrupole ICP-MS (iCAP-Q), where the experimental Al concentrations (Al_{ICP}) were determined based on an Al standard curve prepared separately by the core facility.

5.2.8 Matlab Simulations

To generate data for non-linear regression curve fitting and to simulate data to assist with optimization of ICP-MS sample conditions, Matlab scripts were prepared that, given several experimentally determined parameters as input, output the speciation of Al in citrate buffer (including the concentration of free Al^{3+}) as well as the K_d for Oc:Al^{3+} binding. These Matlab scripts are based on multiple equilibrium chemistry and include equations that describe equilibrium association/dissociation for all known acid/base and metal-ligand binding reactions as well as mass balance equations for all solution components. Acid/base and metal-ligand association equilibrium constants along with the references from which they were obtained are listed in **Table 5.1**. The equations for association constants for citrate-H complexes are listed below:

$$K_{HCit2-} = \frac{HCit^{2-}}{Cit^{3-} * H^{+}} \qquad \textbf{Equation 5.3}$$

$$K_{H2Cit-} = \frac{H_2Cit^{-}}{HCit^{2-} * H^{+}} \qquad \textbf{Equation 5.4}$$

$$K_{H3Cit} = \frac{H_3Cit}{H_2Cit^{-} * H^{+}} \qquad \textbf{Equation 5.5}$$

The equations for association constants for citrate-Al complexes are:

$$K_{AlCit} = \frac{AlCit}{Al^{3+} * Cit^{3-}} \qquad \textbf{Equation 5.6}$$

$$K_{AlCitH+} = \frac{\mathrm{AlCitH^+}}{\mathrm{AlCit}*\mathrm{H^+}}$$ **Equation 5.7**

$$K_{AlCitH-} = \frac{\mathrm{AlCit}}{\mathrm{AlCitH^-}*\mathrm{H^+}}$$ **Equation 5.8**

$$K_{AlCit23-} = \frac{\mathrm{AlCit_2^{3-}}}{\mathrm{AlCit}*\mathrm{Cit^{3-}}}$$ **Equation 5.9**

$$K_{Al(OH)CitH-^2-} = \frac{\mathrm{AlCit^-H_-}}{\mathrm{Al(OH)(CitH_-)^{2-}}*\mathrm{H^+}}$$ **Equation 5.10**

$$\beta_{Al3(OH)4Cit3^4-} = \frac{\mathrm{Al_3(OH)_4Cit_3^{4-}}*[\mathrm{H^+}]^4}{[\mathrm{Al^{3+}}]^3*[\mathrm{Cit^{3-}}]^3}$$ **Equation 5.11**

$$\beta_{Al3(OH)4(CitH-)3^7-} = \frac{\mathrm{Al_3(OH)_4(CitH_-)_3{}^{7-}}*[\mathrm{H^+}]^7}{[\mathrm{Al^{3+}}]^3*[\mathrm{Cit^{3-}}]^3}$$ **Equation 5.12**

The equations for association constants for Al-OH complexes are:

$$K_{Al(OH)\hat{}2+} = \frac{\mathrm{Al(OH)^{2+}}}{\mathrm{Al^{3+}}*\mathrm{OH^-}}$$ **Equation 5.13**

$$K_{Al(OH)2\hat{}+} = \frac{\mathrm{Al(OH)_2^+}}{\mathrm{Al(OH)^{2+}}*\mathrm{OH^-}}$$ **Equation 5.14**

$$K_{Al(OH)3} = \frac{\mathrm{Al(OH)_3}}{\mathrm{Al(OH)_2^+}*\mathrm{OH^-}}$$ **Equation 5.15**

$$K_{Al(OH)4\hat{}-} = \frac{\mathrm{Al(OH)_4^-}}{\mathrm{Al(OH)_3}*\mathrm{OH^-}}$$ **Equation 5.16**

$$\beta_{Al2(OH)24\hat{}+} = \frac{\mathrm{Al_2(OH)_2^{4+}}}{[\mathrm{Al^{3+}}]^2*[\mathrm{OH^-}]^2}$$ **Equation 5.17**

$$\beta_{Al3(OH)4^5+} = \frac{\mathrm{Al_3(OH)_4^{5+}}}{[\mathrm{Al^{3+}}]^3*[\mathrm{OH^-}]^4}$$ **Equation 5.18**

The equations for dissociation constants for Al-Oc complexes are:

$$K_{dAlOc} = \frac{\mathrm{Al^{3+} * Oc}}{\mathrm{AlOc}}$$ **Equation 5.19**

$$K_{dAlOc2} = \frac{\mathrm{AlOc * Oc}}{\mathrm{AlOc_2}}$$ **Equation 5.20**

The mass balance equation for citrate is given by the following equation, where cCit is the total concentration of all citrate species:

$$\mathrm{cCit} = \mathrm{Cit^{3-}} + \mathrm{HCit^{2-}} + \mathrm{H_2Cit^-} + \mathrm{H_3Cit} + \mathrm{AlCit} + \mathrm{AlCitH^+} + \mathrm{AlCitH^-} + (2 * \mathrm{AlCit_2^{3-}}) + \mathrm{Al(OH)(CitH\_)^{2-}} + (3 * \mathrm{Al_3(OH)_4Cit_3^{4-}}) + (3 * \mathrm{Al_3(OH)_4(CitH\_)_3}^{7-})$$

Equation 5.21

For single-site binding of Al to Oc, the mass balance equation for aluminum is given by the following equation, where cAl is the total concentration of all Al species:

$$\mathrm{cAl} = \mathrm{Al^{3+}} + \mathrm{AlCit} + \mathrm{AlCitH^+} + \mathrm{AlCitH^-} + \mathrm{AlCit_2^{3-}} + \mathrm{Al(OH)^{2+}} + \mathrm{Al(OH)_2^+} + \mathrm{Al(OH)_3} + \mathrm{Al(OH)_4^-} + (2 * \mathrm{Al_2(OH)_2^{4+}}) + \left(3 * \mathrm{Al_3(OH)_4^{5+}}\right) + \mathrm{Al(OH)(CitH\_)^{2-}} + (3 * \mathrm{Al_3(OH)_4Cit_3^{4-}}) + (3 * \mathrm{Al_3(OH)_4(CitH\_)_3}^{7-}) + \mathrm{AlOc}$$

Equation 5.22

For two-site binding of Al to Oc, the mass balance equation for aluminum is given by the following equation, where cAl is the total concentration of all Al species:

$$\mathrm{cAl} = \mathrm{Al^{3+}} + \mathrm{AlCit} + \mathrm{AlCitH^+} + \mathrm{AlCitH^-} + \mathrm{AlCit_2^{3-}} + \mathrm{Al(OH)^{2+}} + \mathrm{Al(OH)_2^+} + \mathrm{Al(OH)_3} + \mathrm{Al(OH)_4^-} + (2 * \mathrm{Al_2(OH)_2^{4+}}) + \left(3 * \mathrm{Al_3(OH)_4^{5+}}\right) + \mathrm{Al(OH)(CitH\_)^{2-}} + (3 * \mathrm{Al_3(OH)_4Cit_3^{4-}}) + (3 * \mathrm{Al_3(OH)_4(CitH\_)_3}^{7-}) + \mathrm{AlOc} + \mathrm{AlOc_2}$$

Equation 5.23

The mass balance equation for osteocalcin is given by the following equation, where cOc is the total concentration of all Oc species:

$$\mathrm{cOc} = \mathrm{Oc} + \mathrm{AlOc}$$ **Equation 5.24**

And finally, for the two-site binding system, the mass balance equation for osteocalcin is given by the following equation, where cOc is the total concentration of all Oc species:

$$cOc = Oc + AlOc + (2 * AlOc_2)$$ **Equation 5.25**

For the Matlab script used to simulate data during optimization of the protocol and validation of the non-linear regression technique (**Appendix C**), the constants in the system include the association constants listed in **Table 5.1**, which cannot be modified, the concentration of H^+ and OH^-, which depends on the pH, the total concentration of Al (cAl), citrate (cCit), and osteocalcin (cOc), as well as the dissociation constants for Al binding to Oc (K_{dAlOc} and K_{dAlOc2}). During the generation of simulated data for the single-site binding model, these parameters may be set to generate an output of data including the free concentration of Al^{3+} in the system (Al^{3+}), fraction of Oc bound by Al^{3+}, and the concentration of Al not bound by Oc (Al_{ICP}), which can be described by **Equation 5.26** below:

$$Al_{ICP} = Al^{3+} + AlCit + AlCitH^+ + AlCitH^- + AlCit_2^{3-} + Al(OH)^{2+} + Al(OH)_2^+ + Al(OH)_3 + Al(OH)_4^- + (2 * Al_2(OH)_2^{4+}) + \left(3 * Al_3(OH)_4^{5+}\right) + Al(OH)(CitH\_)^{2-} + (3 * Al_3(OH)_4Cit_3^{4-}) + (3 * Al_3(OH)_4(CitH\_)_3{}^{7-}$$

Equation 5.26

Note that Al_{ICP} would be the concentration measured by ICP-MS during actual experiments, since all forms of Al that are not bound by Oc would be present in the dialysate (see **Figure 5.5** for illustration).

For the single-site binding system, **Equations 5.3 – 5.19**, and **Equations 5.21, 5.22,** and **5.24** are solved by Matlab simultaneously. For the two-site binding system, **Equations 5.3 – 5.21** and **Equations 5.23** and **5.25** are solved by Matlab simultaneously. When using these equations to solve for free Al^{3+} concentrations in the

single-site binding system, experimentally determined cAl, AlOc, and known pH, cCit, and cOc values are input, while the K_{dAlOc} is unknown, and a calculated value can be output with each individual data set (see **Appendix D** for the Matlab script). For the two-site binding system, which was not used for the final data set since it was determined experimentally that Al^{3+} binding followed a single-site binding model, it is also necessary to output the free concentration of osteocalcin (Oc), as these values are necessary for downstream non-linear regression analysis.

5.2.9 Non-Linear Regression Analysis

Non-linear regression (NLR) curve fitting analysis of all data was completed using APMonitor [195], following tutorials found on their website for use of their software within Matlab. During initial validation of the NLR technique, data were simulated following the method described in **5.2.8** to generate data for both single-site binding and two-site binding systems. These data were fit using NLR using either the model derived for single-site binding, where the fraction of Oc bound by Al^{3+} is described by:

$$\text{Fraction Oc Bound} = y = \frac{AlOc}{Oc+AlOc}$$

Substituting AlOc using the equation used to describe the dissociation constant for Al^{3+} binding to Oc (**Equation 5.19**) into the equation used to describe the fraction of Oc bound:

$$y = \frac{\frac{Al^{3+}*Oc}{K_{dAlOc}}}{Oc+\frac{Al^{3+}*Oc}{K_{dAlOc}}}$$

Which can be simplified to give the equation used for NLR curve fitting for the single-site binding model:

$$y = \frac{Al^{3+}}{K_{dAlOc}+Al^{3+}}$$ **Equation 5.27**

For the two-site binding situation, the fraction of Oc bound by Al^{3+} is described by:

$$\text{Fraction Oc Bound} = z = \frac{AlOc+(2*AlOc_2)}{Oc+AlOc+(2*AlOc_2)}$$

The dissociation constant equations that describe the formation of AlOc (**Equation 5.19**) and $AlOc_2$ (**Equation 5.20**) are used to substitute the AlOc and $AlOc_2$ terms in the fraction Oc equation described above to give:

$$z = \frac{\frac{Al^{3+}*Oc}{K_{dAlOc}}+2*\frac{\frac{Al^{3+}*Oc}{K_{dAlOc}}*Oc}{K_{dAlOc2}}}{Oc+\frac{Al^{3+}*Oc}{K_{dAlOc}}+2*\frac{\frac{Al^{3+}*Oc}{K_{dAlOc}}*Oc}{K_{dAlOc2}}}$$

Which can be simplified to give the equation used for NLR curve fitting for the two-site binding model:

$$z = \frac{(Al^{3+}*K_{dAlOc2})+(2*Al^{3+}*Oc)}{(K_{dAlOc}*K_{dAlOc2})+(Al^{3+}*K_{dAlOc2})+(2*Al^{3+}*Oc)}$$ **Equation 5.28**

Note that K_{dAlOc} and K_{dAlOc2} must be specified to be ≥ 0, to eliminate negative solutions. All measured, or simulated, values for fraction of Oc bound by Al^{3+}, Al^{3+}, and Oc in the two-site binding model, must be provided for successful fitting of the data to take place, with output values for predicted fraction of Oc bound by Al^{3+}, K_{dAlOc} and K_{dAlOc2}. The Matlab script used to fit data using APMonitor is found in **Appendix E**. To further validate the final fitted K_d values, data were also fit using cftool in Matlab (**Appendix F**), as well as the NLR curve fitting option in GraphPad Prism 7 software, which was used for final figure generation.

5.2.10. Chrome Azurol S Experiments

To qualitatively monitor for the formation of large, unaccounted-for polymeric forms of Al over time, 2 mL samples containing 187.5 μM chrome azurol S (CAS) were incubated with 0, 0.05, 0.1, 0.25, 0.5, 0.75, 1, 1.25, 1.5, 2, 2.5, 3, 3.5, or 3.96 mM cAl in 5 mM citrate + 150 mM NaCl, pH 4.5, at room temperature for 0.5, 1, 2, 4, 8, 18, 24, 48, 72, and 168 hrs. At each time point, samples were vortexed, and 200 μL aliquots of each sample were added to a 96-well microplate, with absorbance measured at 568 nm. Note that the desired 567.5 nm wavelength could not be achieved due to spectrophotometer limitations. Absorbance measurements were blank subtracted using the average of 3 replicates of the 0 cAl samples. n = 3. The time point at which absorbance readings were no longer stable (i.e., flat over time) were used for each cAl sample to evaluate which data points in the final Gla-Oc with Al incubation samples to exclude from the ΔΘ versus cAl plot, as well as the final binding curve and NLR analysis.

5.3 Results

5.3.1 Determination of the Source of Al^{3+}-bound Gla-Oc in Patient Plasma Samples

During analysis of a cohort of clinical plasma samples by mass spectrometric immunoassay (MSIA) tip extraction and LC-MS (**Figure 5.1a**), a significant fraction of Gla-Oc was observed to bind to Al^{3+}. A representative mass spectrum obtained from MSIA tip extraction of Oc from a patient plasma sample is shown in **Figure 5.1b**, highlighting the isotopically resolved mass spectral peaks associated with free 3-Gla-Oc, 3-Gla-Oc bound to Al^{3+}, and 3-Gla-Oc bound to Mn^{2+}. In this sample, Al^{3+}-bound 3-Gla-Oc constitutes ~ 40 % of the total 3-Gla-Oc present within the sample. The mass spectrum presented here is not a unique finding within the cohort, with the 87 patient plasma

samples presenting 5-50 % of 3-Gla-Oc bound to Al^{3+} (**Figure 5.1c**). To verify that the LC-MS system did not serve as a source of Al^{3+}, synthetic 3-Gla-Oc was directly injected onto the LC-MS system and no Al^{3+} adducts were observed. In contrast, the relative abundance of the Mn^{2+} adduct varied depending on the concentration of osteocalcin flowing through the spray needle at any given instant. This suggested that the Mn^{2+} ions were likely derived from the stainless steel spray needle due to oxidative electrochemical processes that naturally arise when the ion source is operated in positive ion mode [196]. The spray needle was made from 316L stainless steel, which may contain up to 2.0 % Mn by mass [197].

Although the LC-MS was determined not to be the source of Al^{3+}-bound 3-Gla-Oc, whether the MSIA tip extraction process affected observed Al^{3+}-bound 3-Gla-Oc levels was still uncertain. To address this question, synthetic Gla-Oc was spiked into HBS-N buffer, synthetic serum, or 2 M ammonium acetate buffer (used as a wash buffer during the MSIA extraction procedure) and purified by MSIA tip extraction. The HBS-N buffer contributed more Al^{3+} than the 2 M ammonium acetate buffer, with about 30 % of Gla-Oc bound to Al^{3+} after MSIA tip extraction of Gla-Oc spiked into HBS-N buffer (**Figure 5.2a**). When Gla-Oc was spiked into synthetic serum without Al^{3+} added, the percentage of Gla-Oc bound to Al^{3+} was about 3 %. This stark difference in Al^{3+}-bound Gla-Oc levels obtained from different sample conditions suggested Al^{3+} was present in the buffers, but also raised the question whether other components in the synthetic serum were sequestering Al^{3+}, and therefore reducing the levels bound to Gla-Oc. To further support this finding, Al^{3+}-bound Gla-Oc (i.e., Gla-Oc pre-fortified to be nearly saturated with Al^{3+}) was spiked into HBS-N buffer and synthetic serum (without added Al^{3+}) and purified by MSIA tip extraction. Approximately 76 % of Gla-Oc was in the Al^{3+}-bound form after injection onto the LC-MS (**Figure 5.2b**). Similarly, the Al^{3+}-bound Gla-Oc spiked into HBS-N samples

was not depleted after sample processing. The Al^{3+}-bound Gla-Oc levels from the synthetic serum spiked sample were depleted to 24 % of the total Gla-Oc, suggesting components in the synthetic serum sequester Al^{3+} and limit its binding to Gla-Oc. This is expected since, at equilibrium, transferrin binds ~ 90 % and citrate binds ~ 10 % of the Al^{3+} in plasma, with only trace quantities bound to other Al^{3+} binders or free in solution [93].

5.3.2 Circular Dichroism & Oc Precipitation Studies

The generation of binding affinity data for Gla-Oc binding to Al^{3+} was first attempted using the previously described CD technique [29], which was used to determine the binding affinity for bovine Oc binding to Ca^{2+} (K_d = 1.25 mM) and to Pb^{2+} (K_d = 85 nM), by monitoring the change in ellipticity at 222 nm. The initial attempt to obtain similar CD spectra for human Gla-Oc binding to Al^{3+} (**Figure 5.3a**) was unsuccessful, as the ellipticity values at 222 nm did not change in the presence of increasing titrations of Al. Attempts at replicating Pb titration CD spectra with Gla-Oc were also unsuccessful and absorbance measurements at 280 nm after each Pb titration revealed significant loss of in-solution Gla-Oc compared to the expected change in concentration due to dilution alone (**Figure 5.3b** & **5.3c**).

The results presented in **Figure 5.3** alluded to the possibility that Gla-Oc was precipitating out of solution with increasing titrations of Al or Pb. To further validate this hypothesis for Al, Gla-Oc (**Figure 5.4a**) and Glu-Oc (**Figure 5.4b**) were titrated with increasing Al concentrations, samples were centrifuged, and the absorbance at 280 nm of the supernatants were measured. The results demonstrate the stark difference between the in-solution Oc concentrations when titrated with Al, compared to the expected concentrations due to dilution alone, and when samples were titrated with buffer only. The fractions of Oc precipitated with increasing Al titrations were then plotted for Gla-Oc

(**Figure 5.4c**) and Glu-Oc (**Figure 5.4d**), and may reflect the binding affinity of Oc for precipitated Al in the form of $Al(OH)_{3(s)}$.

5.3.3 Initial Citrate Concentration Optimization and Titration Data

Since both CD and direct infusion MS, discussed in **Chapter 2**, did not provide adequate results to accurately determine the binding affinity of Gla-Oc for Al^{3+}, the next approach tested aimed to quantitate the unbound concentration of Al^{3+} with each Al titration using dialysis and inductively-coupled plasma mass spectrometry (ICP-MS). With a 1:1 binding model for Gla-Oc binding to Al^{3+}, this would enable calculation of the fraction of Gla-Oc bound by Al^{3+}, allowing for calculation of K_{dAlOc} for the binding interaction. The dialysis set-up is illustrated in **Figure 5.5** and shows the components of the system that are found in the sample chamber, where Oc is added, and in the dialysate, which includes free Al^{3+} and all Al-citrate and Al-OH complexes, present in both the sample chamber and dialysate at equal concentrations once equilibrium has been achieved.

Loss of in-solution Oc concentrations is likely a reflection of Oc adsorption to precipitated $Al(OH)_3$, due to the low solubility of Al^{3+} at pH 7.4 (~ 19 – 82.21 fM, depending on the $K_{sp} = 3 \times 10^{-34}$ [198] or $K_{sp} = 1.3 \times 10^{-33}$ [108]). The low solubility of Al^{3+} could be improved by reducing the pH of solution to 4.5, with the calculated solubility limit of Al^{3+} being approximately 9.5 – 41.2 µM. If the lowest solubility limit is assumed, using the solubility product constant described by Skoog, *et al* [198], then considering the affinity constants for Al-OH complexes, this sets the maximum cAl in the system to be about 10.4 µM to ensure total maintenance of Al in solution. To further improve solubility of Al within the biologically relevant pH range for Oc within the bone, a competitive binder for Al^{3+}, citrate, was included in solution. The affinity constants for Al-citrate complexes are known [199, 200], enabling the calculation of the dissociation constant for Oc binding

to Al^{3+} in the presence of this competitive binder. Inclusion of citrate in the system does not affect the solubility limit for free Al^{3+}, but rather increased total citrate concentrations, cCit, affect the cAl at which point this Al^{3+} solubility limit is achieved. This approach creates a reservoir of soluble but "releasable" Al^{3+} in which citrate-bound Al^{3+} can be released as free Al^{3+} when the existing pool of free Al^{3+} is depleted due to binding to Oc. These maximum cAl values were calculated for 0 – 500 μM cCit for each reported K_{sp} value (**Figure 5.6a**). With just 20 μM citrate, pH 4.5, the calculated maximum cAl where Al^{3+} reaches the solubility limit was determined to be approximately 30.1 μM. Prior to Gla-Oc and Al titration sample preparation, an initial set of samples with cAl ranging from 0 – 27.8 μM were prepared and injected onto the ICP-MS and demonstrated both the precision and accuracy of measurements obtained (**Figure 5.6b**). These results are suggestive that cAl remained soluble at the concentrations tested here, resulting in accurate serial dilution of the stocks to prepare these samples. The samples where no Al was added revealed there to be ~ 1 μM cAl present as a contaminant, even after extensive acid washing procedures took place, revealing a lower limit for cAl in the system. In addition, Gla-Oc did not precipitate to any significant degree during Al titrations when the solution contained 20 μM citrate, pH 4.5, further supporting the maintenance of in-solution cAl and cOc in the experimental set-up (**Figure 5.6c**).

Next, the dialysis incubation time was optimized with Gla-Oc samples incubated for 0 – 48 hours, with equilibrium taking place after 8 hours (**Figure 5.7a**). All future samples were incubated for 18 hours to ensure equilibrium had taken place. An initial set of Gla-Oc titrations with Al were completed to yield the binding curve depicted in **Figure 5.7b**, where the fraction of Gla-Oc bound is plotted against the Al_{ICP} values. To convert Al_{ICP} values to Al^{3+}, the equations discussed in **5.2.8** for the single-site binding model were used to solve for Al^{3+} using experimentally determined cAl and AlOc concentrations

(where the AlOc values were determined by subtracting the Al_{ICP} values from the cAl values) and the known cOc and pH. Inclusion of all data points yields the binding curve depicted in **Figure 5.7c**, where the data points highlighted in grey were excluded from the final binding curve due to the high variability found between replicate data points in **Figure 5.7d**, with these data shown not to fit to the single-site binding model. Along with the individual K_{dAlOc} estimates presented in **Table 5.2**, these results are suggestive that the data do not fit to the single-binding model in **Equation 5.27**. Thus, both the experimental design and the non-linear regression model had to be modified for successful binding affinity constant characterization to take place. Since the data appeared to plateau at ~ 0.5 fraction bound, as seen in **Figure 5.7b**, this led to the hypothesis that the single-site binding assumption may be invalid, with the concentration of Al^{3+} bound to Gla-Oc not equal to the concentration of Gla-Oc bound to Al^{3+} (ie, not a 1:1 binding system). Since multiplying the concentration of Al^{3+} bound by a factor of 2 when calculating the concentration of Gla-Oc bound seemed to result in the fraction of Gla-Oc bound reaching a plateau at ~ 1, this led to the hypothesis that binding may follow a 2:1 Gla-Oc:Al^{3+} binding model. As discussed in the methods section, the development of Matlab simulations and NLR curve fitting analysis was developed successfully for this scenario. However, as will be seen in section **5.3.5**, since $AlOc_2$ (2:1 Gla-Oc:Al^{3+} complex) was not observed experimentally, this suggested that this plateau of the binding data at ~ 0.5 fraction of Gla-Oc bound by Al^{3+} was not caused by the presence of an $AlOc_2$ dimer.

5.3.4 Binding Simulations & Non-linear Regression Validation

Due to the challenges faced with initial attempts at obtaining binding data using the dialysis and ICP-MS technique, it was necessary to use the Matlab scripts to produce simulated binding data to assist in further experimental plans (**Appendix C**), and to

validate the downstream NLR curve fitting technique (**Appendix D**) prior to use with experimental data. The equations used in both the single-site, and two-site, binding models are described in the methods section **5.2.8**, and used along with defined values for cAl, cOc, cCit, and dissociation (K_{AlOc} and K_{AlOc2}) and citrate and hydroxide association (**Table 5.1**) constant parameters, it was possible to generate simulated single-site (**Figure 5.8a**) and 2:1 Gla-Oc:Al^{3+} (**Figure 5.9a**) binding data.

To fit the data using the single-site (**Equation 5.27**) and 2:1 Gla-Oc:Al^{3+} (**Equation 5.28**) binding models, the data must be plotted with free Al^{3+} concentrations on the x-axis, rather than Al_{ICP} values. To test the ability to calculate Al^{3+} values with these simulated data sets, the values that would be known in an experimental setting were used to solve for Al^{3+} and estimate K_d values using the Matlab scripts (**Appendix D**). For the single-site binding simulated data, values that would be experimentally known are cAl, cOc, cCit, H^+, OH^-, and Al_{ICP}. For the single-site binding situation, the concentration of Al^{3+}-bound Oc (AlOc), and the fraction of Oc bound, were then calculated and these values were used along with the same equations used to simulate the data to determine the free Al^{3+} concentrations and to estimate K_{dAlOc} for each Al titration (**Table 5.3**). The data were then fit using the NLR equation, **Equation 5.27**, to obtain predicted fraction of Oc bound values (**Table 5.4**) and to estimate a single K_{dAlOc} value for the given data set (**Figure 5.8b**).

Since the initial experimental data did not fit the single-site binding model, the data simulated using a 2:1 Gla-Oc:Al^{3+} binding model were used to validate the ability to successfully solve for free Al^{3+} concentrations using experimentally known values cAl, cOc, cCit, H^+, OH^-, Al_{ICP}. Due to the additional unknown parameter K_{dAlOc2}, another parameter must be input into the Matlab script to successfully solve for Al^{3+} and to estimate both

K_{dAlOc} and K_{dAlOc2}. The free concentration of Oc (Oc) must also be calculated for successful non-linear regression curve fitting of the data. This can be done in an experimental situation by measuring the concentration of $AlOc_2$ (eg, by SEC), since knowledge of the concentrations of this additional parameter allows for an additional parameter to be solved by the script successfully. Therefore, the $AlOc_2$ values were taken from the simulated data and, along with the other parameters previously listed, were input into the Matlab script to solve for Al^{3+}, free Oc, K_{dAlOc} and K_{dAlOc2} for each cAl titration (**Table 5.5**). The free Al^{3+}, free Oc, and fraction of Oc bound values were then used to fit the data using **Equation 5.28** to solve for predicted fraction of Oc bound values (**Table 5.6**) and single K_{dAlOc} and K_{dAlOc2} values for the given data set (**Figure 5.9b**).

To make an initial estimate for K_{dAlOc} and K_{dAlOc2} from the initial experimental data shown in **Figure 5.7b**, the assumption that the concentration of Oc bound is equal to twice the concentration of Al bound was made, since $AlOc_2$ measurements were not made on this data set. This is, of course, an unvalidated assumption, and was purely made to assist with cCit optimizations prior to proceeding with experimental data acquisition. With this assumption, the dissociation constants were determined to be K_{dAlOc} = 0.0051789 μM and K_{dAlOc2} = 1.3686 μM (**Figure 5.10a**) and were successfully predicted by the two-site binding non-linear regression model when simulated data were generated with these defined K_d values (**Figure 5.10b**). Due to the lower limit of ~ 1 μM cAl, the initial experimental data started with a minimum of ~ 0.56 fraction of Gla-Oc bound by Al^{3+}. To obtain a more complete binding curve, free Al^{3+} concentrations must reach low enough quantities for the fraction of Gla-Oc bound by Al^{3+} to be minimized. For high affinity binders, such is the case with Gla-Oc for Al^{3+}, this poses a challenge due to limits of detection of instrumentation. This issue was resolved using citrate in solution, due to the μM concentrations of Al_{ICP} measured in the system, compared to the pM to nM

concentrations of free Al^{3+} that would otherwise need to be measured to obtain relevant binding curve data. Simulated data where the initial cAl is 1 µM resulted in reduced Al^{3+} and fraction of Oc bound values when cCit was increased (**Figure 5.10c**), with values below 0.2 fraction of Gla-Oc bound when cCit was set to 500 µM (**Figure 5.10d**). This demonstrated that the ~ 1 µM cAl limit effects on Al^{3+} and fraction Oc bound values could be overcome with increased cCit values and gave a starting point for the experimental design going forward. Knowledge of the concentrations of free Al^{3+} and free Oc with each cAl titration eliminated the "titration effect" (**Figure 5.11**), which is described in the review by Jarmoskaite, *et al.* [201], where the assumption of total ligand is equivalent to the free concentration of ligand is no longer valid as the concentration of protein is increased well above its binding affinity for the ligand. Calculation of these parameters (free Al^{3+} and free Oc) enabled the use of concentrations of cAl and cOc within limits of detection in the system, and overcame the titration effect.

Since it was later found that Al^{3+} binding to 3-Gla-Oc follows the single-site binding model, the two-site binding model will not be discussed further, but may be used in the future for more complex binding systems. For both the single-site and two-site binding models, the NLR curve fitting technique proved to be successful at calculating the free Al^{3+} and K_d values, as they closely matched the simulated data used when optimizing the Matlab scripts.

5.3.5 $AlOc_2$ was not Observed by Direct Infusion MS

To experimentally validate the presence of the $AlOc_2$ dimer, 8 µM Gla-Oc in 20 µM citrate, pH 4.5, was titrated with 80 µM cAl and incubated for 18 hrs at room temperature, followed by direct infusion onto the MS. The mass spectrum revealed low abundance of Oc dimers (**Figure 5.12a**), with a 7+ charge state, which matches the results presented in

Chapter 2. The charge deconvoluted mass spectrum for the monomeric Gla-Oc (**Figure 5.12b**) showed the base peak to be the 3-Gla-Oc bound to one Al^{3+}. The charge deconvoluted mass spectrum for the dimeric Gla-Oc species (**Figure 5.12c**) did not show the isotope peak cluster associated with the 2x3-Gla-Oc dimer or the Al^{3+}+2x3-Gla-Oc ($AlOc_2$) dimer. However, isotope peaks corresponding to $2Al^{3+}$+2x2-Gla-Oc, $2Al^{3+}$+2-Gla+3-Gla, and $2Al^{3+}$+2x3-Gla-Oc (collectively referred to as Al_2Oc_2) were observed. The low abundance of these dimers, however, was suggestive that the primary binding event is the formation of monomeric Gla-Oc bound to a single Al^{3+}, with the calculated fractional abundance of the Al_2Oc_2 of 0.012, determined by dividing the sum of the mass spectral peak counts associated with Al_2Oc_2 species by the sum of the total mass spectral peak counts for all Gla-Oc species across all charge states. Attempts at further validation of this Al_2Oc_2 dimer, or any other Oc_2 dimer, by size exclusion chromatography (SEC) revealed no detectable dimer present in the chromatogram (data not shown). Due to the low fractional abundance of the observed Al_2Oc_2 dimer species, all future experiments were performed with the single-site binding assumption.

5.3.6 Final CD & ICP-MS Experiments

With the single-site binding assumption, it was possible to re-visit the use of CD spectroscopy as a technique to quantify the fraction of Oc bound by Al^{3+} now that the Al and Oc precipitation issues had been resolved. Thus, Gla-Oc samples were prepared in 5 mM citrate + 150 mM NaCl, pH 4.5, and titrated with 0 – 3.96 mM cAl. The change in ellipticity ($\Delta\Theta$) at 222 nm for each sample was plotted against the cAl values experimentally determined by ICP-MS and the data were fit to **Equation 5.1** to estimate the $\Delta\Theta_{max}$ = -2213 (**Figure 5.13b**) using the Matlab script shown in **Appendix B**. Measured cAl values closely matched the expected cAl values, as shown in **Figure 5.14**.

At pH values greater than or equal to ~ 4.5 and at concentrations greater than or equal to high micromolar/low millimolar, Al in aqueous solution can slowly (over periods of hours to days) form large, undefined soluble polymeric species that are not accounted for by the species in **Equation 5.22** [202]. If this were to occur it would become impossible to accurately calculate Al speciation and the Oc:Al^{3+} equilibrium binding constant (K_d) using the multiple equilibrium chemistry approach described above. Fortunately, the formation of these large polymeric species can be monitored using a spectrophotometric technique. Chrome azurol S (CAS) is a dye that absorbs in the 500-600 nm region when it binds to free Al^{3+} and its small polymeric forms (**Table 5.1**) that occur upon the initial dissolution of Al salts in water [203, 204]. Values excluded from the NLR curve fit are in grey and were excluded based on the time Al stocks were incubated prior to CD sample titration and analysis, and the cAl values that were found not to maintain linear absorbance readings at that particular time point, as determined by the CAS experiments (**Figure 5.15**). The final binding curve data were then plotted in **Figure 5.13c** and fit to the single-site binding model in **Equation 5.27** to calculate K_{dAlOc} = 1.0 ± 0.12 nM with R^2 = 0.855. Individual K_{dAlOc} estimates for each sample are shown in **Table 5.7**. The CD spectra for Gla-Oc incubated with 0 – 0.75 mM cAl were overlayed in **Figure 5.13d**, depicting an increase in α-helical content with increased cAl titrations. Due to the slow (days long) formation of polymeric Al species that could not be quantitatively accounted for, further increases in cAl titrations resulted in an unaccountable loss in free Al^{3+} and therefore a decrease in α-helical structure from 0.75 – 3.96 mM cAl, with the 3.96 mM cAl titrated Gla-Oc sample closely matching the sample titrated with 0 mM cAl (**Figure 5.13e**).

5.4 Discussion

Initial attempts at quantifying the binding affinity of Al^{3+} for Gla-Oc and Glu-Oc using direct infusion native MS (**Chapter 2**) proved unsuccessful due to i) the harsh ESI source,

ii) the presence of competitive binding metals in the spray needle, and iii) spray needle clogging due to protein/metal precipitation during the infusion process. The challenges faced when seeking to quantify the binding affinity of Al^{3+} for human Gla-Oc and Glu-Oc were again brought to attention when the CD results for Pb^{2+} binding Oc obtained by Dowd, *et al* [29] were not able to be reproduced here. The solubility of Pb^{2+} in water at pH 7.4 is roughly 0.23 µM (K_{sp} = 1.43 x 10^{-20} [108] for $Pb(OH)_2$), suggesting that Pb^{2+} should precipitate as $Pb(OH)_2$ at the µM concentrations tested here.

As mentioned previously, the precipitation of Oc with increased Al titrations was likely a reflection of the low solubility limit of Al^{3+} at pH 7.4. For this reason, all future experiments were conducted at pH 4.5, and with the inclusion of citrate in the system this enabled successful binding curve data to be obtained for Gla-Oc binding to Al^{3+}. Since Oc functions *in vivo* within the pH range of 4.5 – 7.4, the values reported here are considered biologically relevant. The use of a higher citrate concentration (5 mM) allowed for proper buffering conditions. Additionally, CD spectroscopy can tolerate physiological salt concentrations (150 mM NaCl), which was thus a far more physiologically ideal experimental design than the dialysis and ICP-MS technique. Use of this technique also enabled confirmation of increased α-helical secondary structure induced by Al^{3+} binding, matching previously published literature for Ca^{2+} and Pb^{2+} binding to Oc.

In contrast to Ca^{2+} and Pb^{2+} binding reported in the literature, an apparent decrease in the affinity of Al^{3+} binding to Gla-Oc with further increased cAl titrations was observed. This could be explained by the increase in higher-order Al complexation and polymer formation [202], and thus untrackable sequestration of Al from the pool available to interact with Gla-Oc in solution. This posed a challenge when estimating the $\Delta\Theta_{max}$, as it could not be further validated experimentally. Attempts at validating the $\Delta\Theta_{max}$ at pH 4.0 where Al is more soluble, revealed that Gla-Oc precipitates (**Figure 5.16a**), and is

thus unstable, at this lower pH, which is the same pH as its known pI of 4.0 [110].

Due to the observation that Glu-Oc precipitates below pH 5.0 (**Figure 5.16b**), complementary binding experiments for Glu-Oc binding Al^{3+} at pH 4.5 were not possible. This, along with the challenges faced with Al when increasing the pH, both low solubility and increased higher-order Al complexation [95, 203-205], and the qualitative knowledge of lower affinity binding of Glu-Oc to Al^{3+} compared to Gla-Oc (**Chapter 2**), made determination of the K_{dAlOc} for Glu-Oc binding Al^{3+} not feasible. Precipitation of Oc at, and below, the pI is expected due to the loss of surface charge of the protein (**Figure 5.17**), and since the overall charge of the Glu/Gla residues involved in metal binding will depend on the pH of the solution, a dissociation constant for Glu-Oc binding to Al^{3+} at pH 5.0 would not be directly comparable to the value obtained for Gla-Oc binding to Al^{3+} at pH 4.5.

The K_{dAlOc} for human Gla-Oc binding to Al^{3+} reported here is significantly lower than the value reported in the literature, K_{dAlOc} = 83 nM for bovine Oc at pH 5, likely for the following reasons: i) human Oc may bind with higher affinity compared to bovine Oc, ii) Gla-Oc binding is parsed here (Rowatt, *et al* [95] did not parse Gla-Oc and Glu-Oc binding), with qualitative evidence supporting a higher affinity compared to Glu-Oc, and iii) the effect of competitive binders, including the formation of Al-OH complexes (both soluble and insoluble), was accounted for here, while this was not taken into account previously in the literature. Taken together, the results presented here are expected to better reflect the true affinity of human Gla-Oc for Al^{3+}.

Since calculation of these K_d values depends on the accuracy of the previously reported association constants, the values presented here are reported with errors calculated based on the reported K_a values exhibiting 0 and ±10 % error in the published values. Based on these error propagation calculations, the calculated K_{dAlOc} value is not

greatly affected by this error propagation (**Figure 5.18**).

The overall success of this method to quantify Al^{3+} binding to Gla-Oc is in large part due to the inclusion of a competitive binder, citrate, in the system. This enabled the increase in the solubility limit of cAl, overcoming issues related to high limits of detection, and preventing precipitation of $Al(OH)_3$ and Gla-Oc during binding studies. With previously reported K_a values for Al-OH and Al-citrate complexes, it was possible to parse between true free Al^{3+} and the remaining Al in various complexed species. The use of citrate also alleviated challenges faced with obtaining data points associated with the lower fraction of Gla-Oc bound in the binding curve data that would otherwise be impossible to obtain due to Al contamination in the system. Finally, the concentration of citrate, and therefore cAl, was increased to include millimolar quantities for improved buffer capacity in the system.

5.5 Conclusions

Through the use of a competitive binder, decreased pH in the system, and a complex system of linear equations, the binding affinity of Al^{3+} for Gla-Oc was successfully estimated here. This method could be adopted to determine the binding affinity of Gla-Oc for other metals, including more complicated systems, as exemplified in the two-site binding system discussed in the methods section. The nanomolar dissociation constant calculated for Al^{3+} binding to Gla-Oc suggests that the Al^{3+} that accumulates in bone may compete for Oc binding, and suggests it is important to consider Al as a potential contaminant during *in vitro* studies of Oc. Since Oc functions to assist with proper formation of hydroxyapatite within the bone, but when bound to bone also serves as a chemoattractant from bone-resorbing osteoclast precursor cells [54, 206-209], the binding of Al^{3+} to Gla-Oc offers a possible explanation for the fact that animal and tissue culture models have shown that Al can impact bone mineralization, leading to osteoblast

apoptosis and reduced bone mineral density [191-193]. As mentioned previously, the distribution of Al^{3+} at equilibrium in plasma is such that the majority is bound to transferrin (~ 90 %) and citrate (~ 10 %), with only trace quantities available to bind to other components [93]. However, with the high affinity binding of Al^{3+} to 3-Gla-Oc presented here, it can be hypothesized that Al^{3+} found within the bone may bind to 3-Gla-Oc, and, following bone dissolution/resorption, potentially remain bound in blood plasma to be excreted by glomerular filtration. In fact, the ability of much of the Al^{3+} to remain bound to 3-Gla-Oc during the several hours long MSIA purification procedure suggests that the rapid (several minutes [44]) excretion rate of Oc from plasma into urine should limit the degree of Al^{3+} transfer from 3-Gla-Oc to transferrin and citrate, and thus prevent this toxic metal's redistribution to other tissues throughout the body. Taken together with the γ-carboxylation dependency of increased binding affinity of Oc for Al^{3+}, this suggests that this process of removal of toxic metals from the body also has a nutritional dependence (vitamin K intake). Since binding kinetics for the Al^{3+} and 3-Gla-Oc interaction were not studied quantitatively here, and the effect of γ-carboxylation status of Oc on Al^{3+} excretion was not tested *in vivo*, this is an important avenue of future research to evaluate the validity of this hypothesis. If this is validated in the future, the significance of adequate nutritional intake of vitamin K on the regulation of toxic metal deposition to more sensitive tissues throughout the body will be bolstered by these findings.

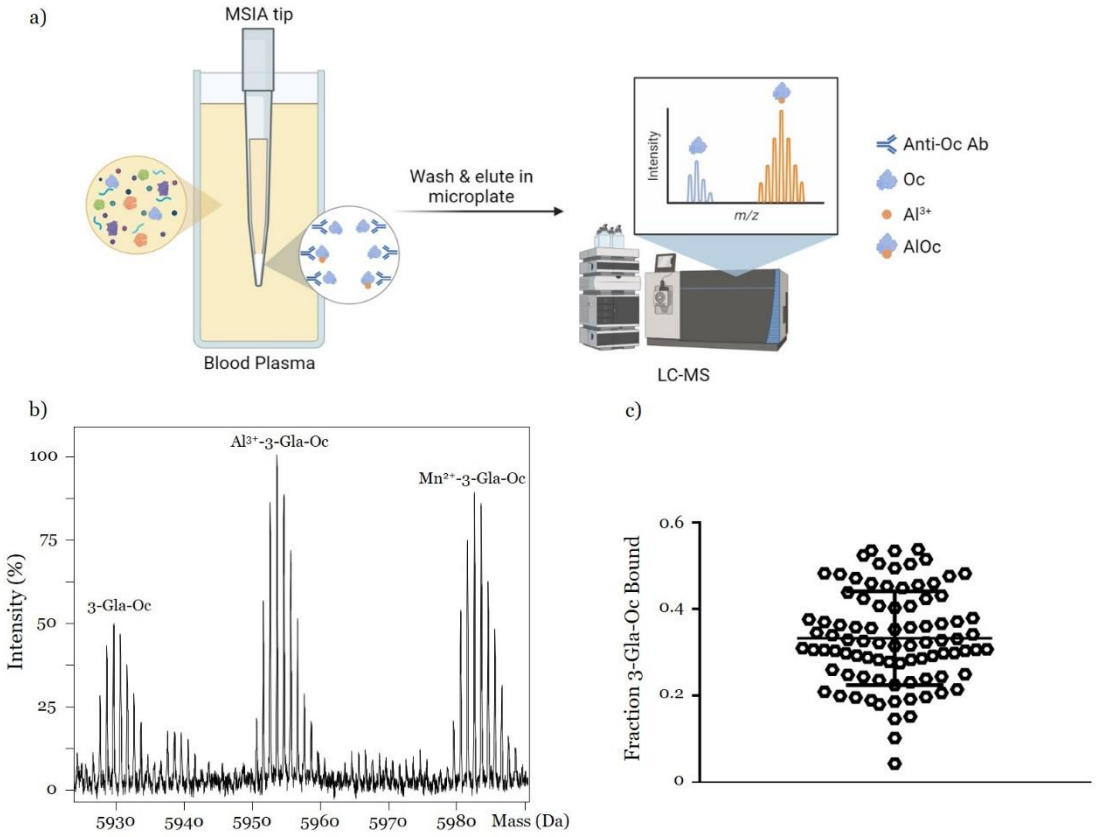


Figure 5.1. Al^{3+} Bound to Osteocalcin in Patient Plasma Samples. a) MSIA tip purification using anti-Oc antibody (Ab) and LC-MS injection workflow. b) Deconvoluted mass spectrum of an LC-MS injected purified patient plasma sample showing 3-Gla-Oc, 3-Gla-Oc bound to one Al^{3+}, and 3-Gla-Oc bound to one Mn^{2+}. c) Fraction of 3-Gla-Oc bound to Al^{3+} after purification and LC-MS sample injection of patient plasma samples. Mean shown by black line with error bars corresponding to one standard deviation from the mean.

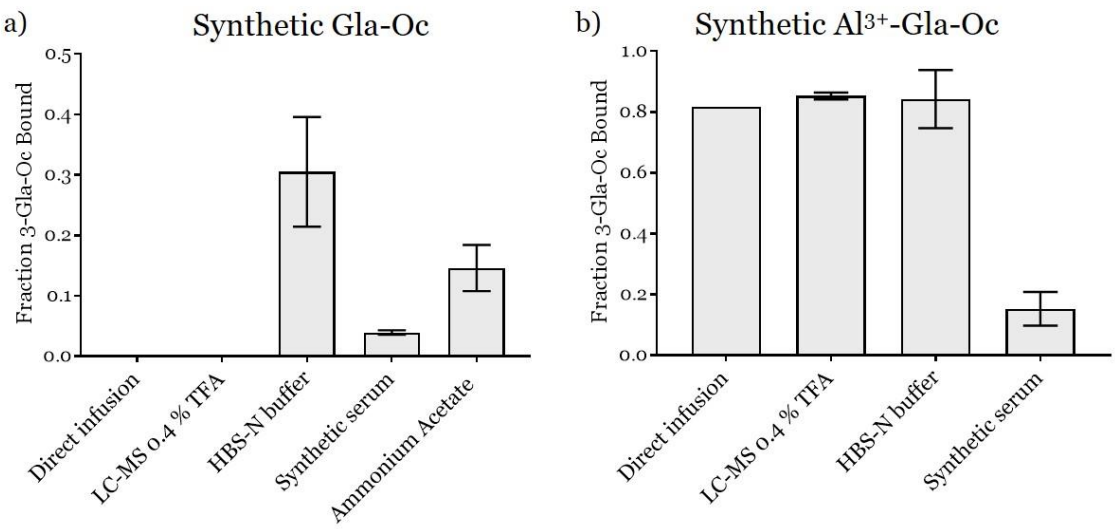


Figure 5.2. Synthetic Gla-Oc Control Experiments to Assess the Degree of Gain and Loss of Al^{3+}-Bound Gla-Oc Obtained During MSIA Tip Purification and LC-MS Injection Procedures. a) Synthetic Gla-Oc, or b) synthetic Gla-Oc pre-incubated with Al, was directly infused onto the MS, injected onto the LC-MS, or spiked into HBS-N buffer, synthetic serum, or ammonium acetate buffer (used as a rinse buffer during MSIA tip procedure), and purified by MSIA tip extraction, followed by LC-MS injection. The fraction of 3-Gla-Oc bound by Al^{3+} was quantified, with n = 2 replicates and error bars representing SD.

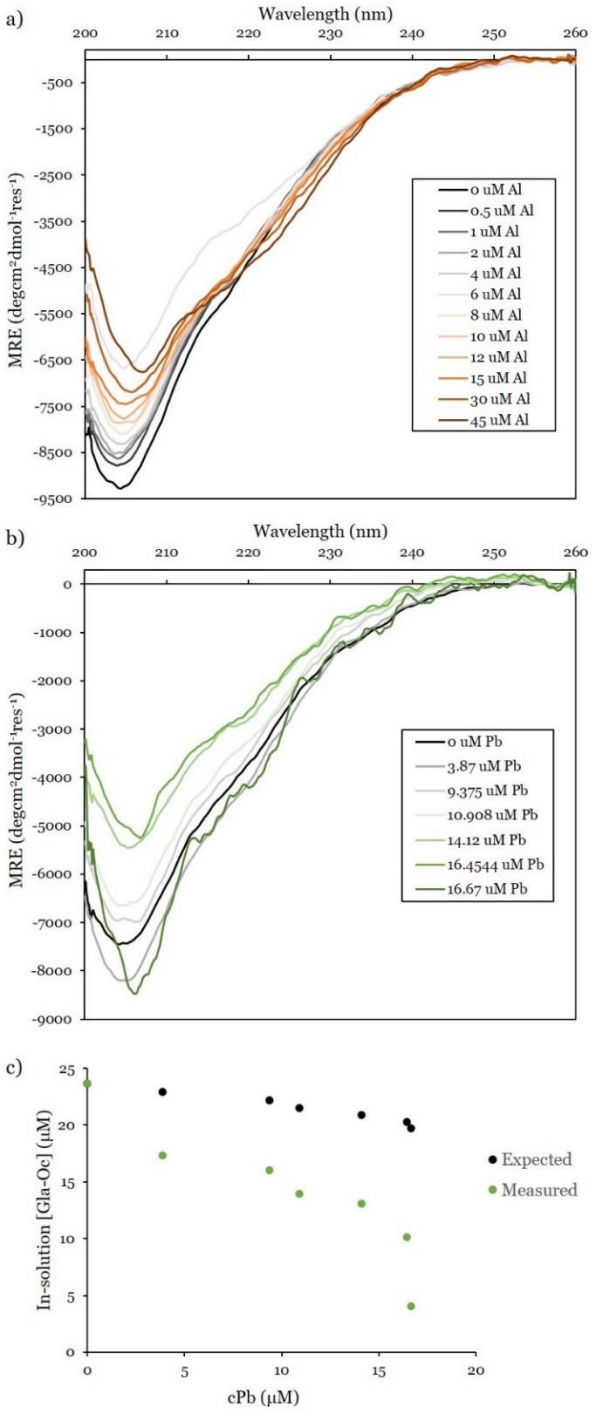


Figure 5.3. Circular Dichroism to Obtain the Dissociation Constant for Gla-Oc Binding Al^{3+} and Pb^{2+} Reveals Protein Precipitation. CD spectra of Gla-Oc with a) 0 – 45 μM Al and b) 0 – 16.67 μM Pb titrations from 200 – 260 nm wavelengths. MRE = molar residue ellipticity, $degcm^2dmol^{-1}res^{-1}$. c) Expected sample Gla-Oc concentrations due to dilution (black) versus measured concentrations (green) with increasing titrations of Pb.

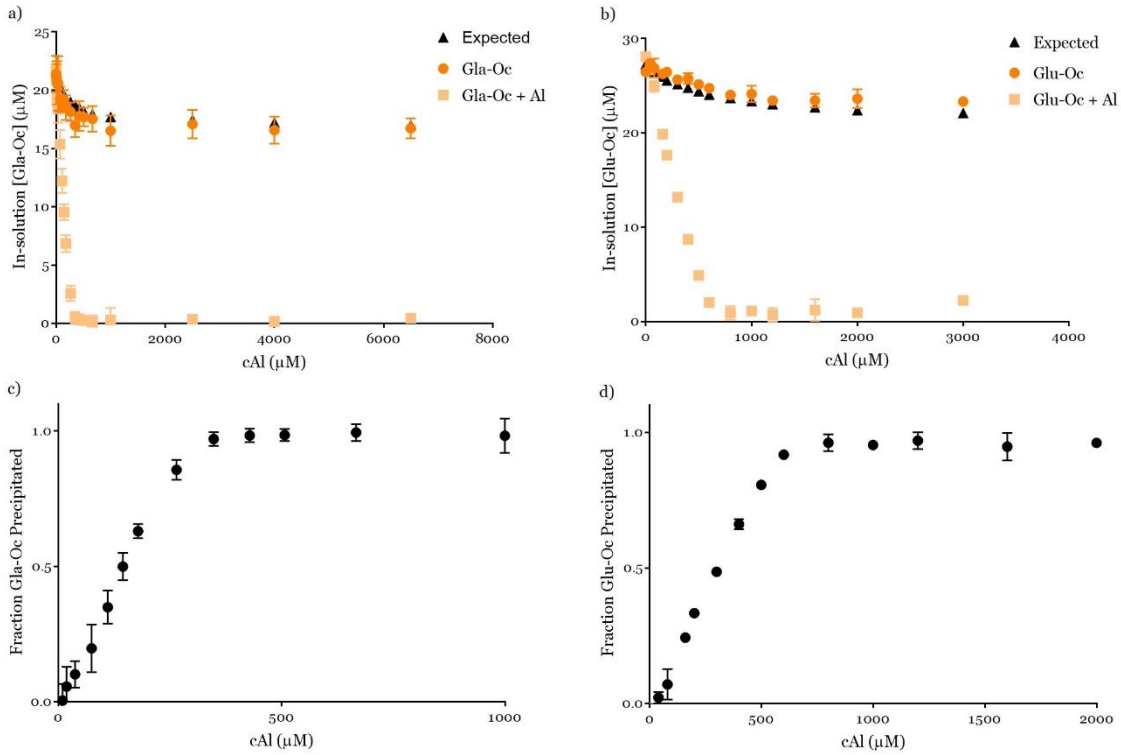


Figure 5.4. Loss of In-Solution Osteocalcin in the Presence of Increased Al Titrations. a) Comparison between the expected concentration of Gla-Oc due to dilution alone (black triangles), the concentration of Gla-Oc in samples titrated with 20 mM Tris + 150 mM NaCl, pH 7.4 buffer (dark orange circles), and the concentration of Gla-Oc in samples titrated with increasing concentrations of Al in buffer (light orange squares). b) Complementary experiments testing in-solution Glu-Oc concentrations. The fraction of c) Gla-Oc or d) Glu-Oc precipitated was then plotted against the total concentration of Al in the samples. n = 3.

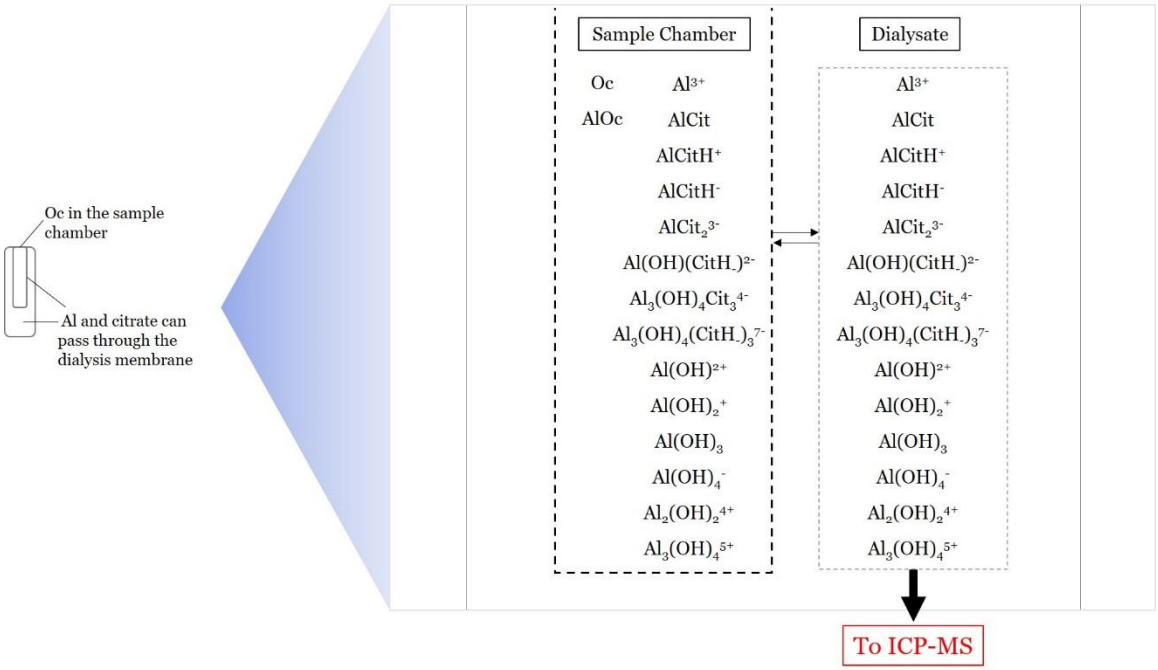


Figure 5.5. Dialysis Set-up Used for Binding Affinity Studies of Osteocalcin Binding to Al^{3+}. On the left is a simplified depiction of the set-up, where the Oc is placed in the sample chamber compartment and the Al is added to the dialysate and capable of passing through the dialysis membrane. The right-hand side of the figure further specifies the components of the system that are found in the sample chamber, which includes all forms of Oc, free Al^{3+}, Al bound citrate complexes, and Al hydroxide complexes. The molecular weight cut-off of 3 kDa of the dialysis membrane prevents all forms of Oc to diffuse into the dialysate, while the free Al^{3+}, Al bound citrate complexes, and the Al hydroxide complexes can pass freely through the membrane. Once sufficient sample incubation time is completed such that equilibrium may be achieved between the concentration of free Al^{3+}, Al bound citrate complexes, and the Al hydroxide complexes in the dialysate matches the concentration in the sample chamber, an aliquot of the dialysate solution is prepared for ICP-MS to determine $[Al]_{ICP}$.

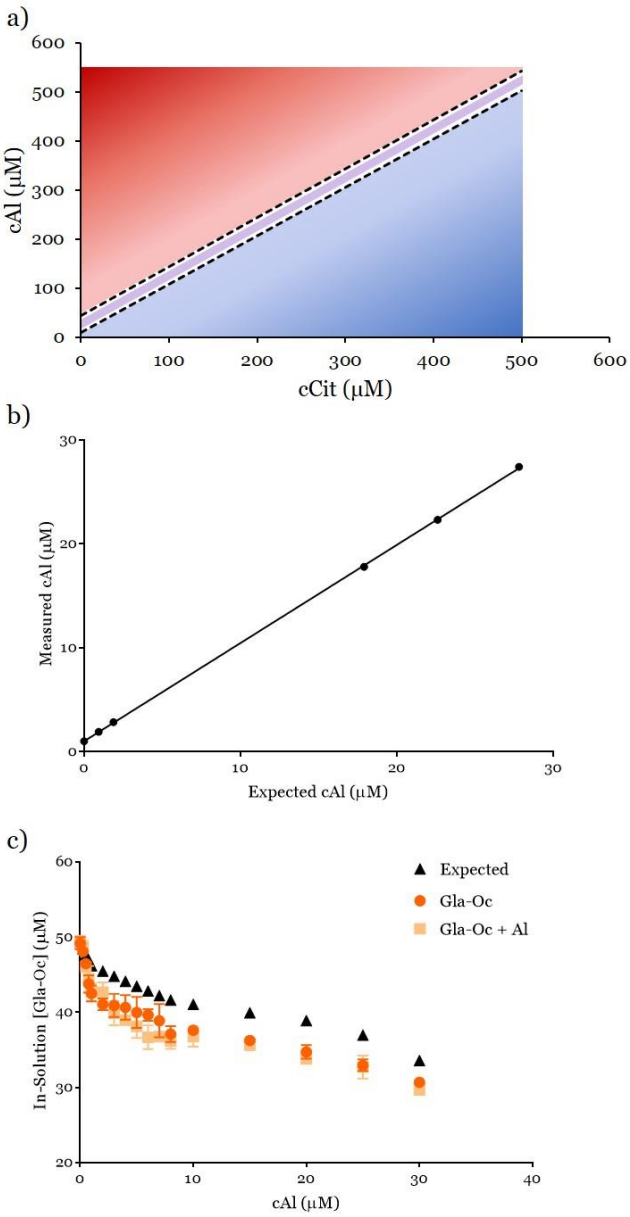


Figure 5.6. Citrate Assists in Maintaining Both Al and Gla-Oc in Solution. a) The upper (K_{sp} = 1.3 x 10^{-33}) and lower (K_{sp} = 3 x 10^{-34}) dashed lines represent the calculated maximum cAl that remain soluble at increasing concentrations of citrate, with the purple region representing the cAl values between these limits. Below the lower dashed line represents points where cAl and cCit values result in maintenance of Al in solution (blue gradient), whereas the region above the dashed lines represents cAl and cCit combinations that result in Al precipitation (red gradient). b) Quality control experiment to verify the precision and accuracy of Al concentrations (µM) between the expected and ICP-MS-measured values in 20 µM citrate, pH 4.5. c) Expected in-solution Gla-Oc concentrations (black triangles) due to dilution alone compared to the in-solution Gla-Oc concentrations experimentally measured for samples titrated with 20 µM citrate (dark orange circles) and 20 µM citrate containing 0 – 30 µM Al (light orange squares). n = 3.

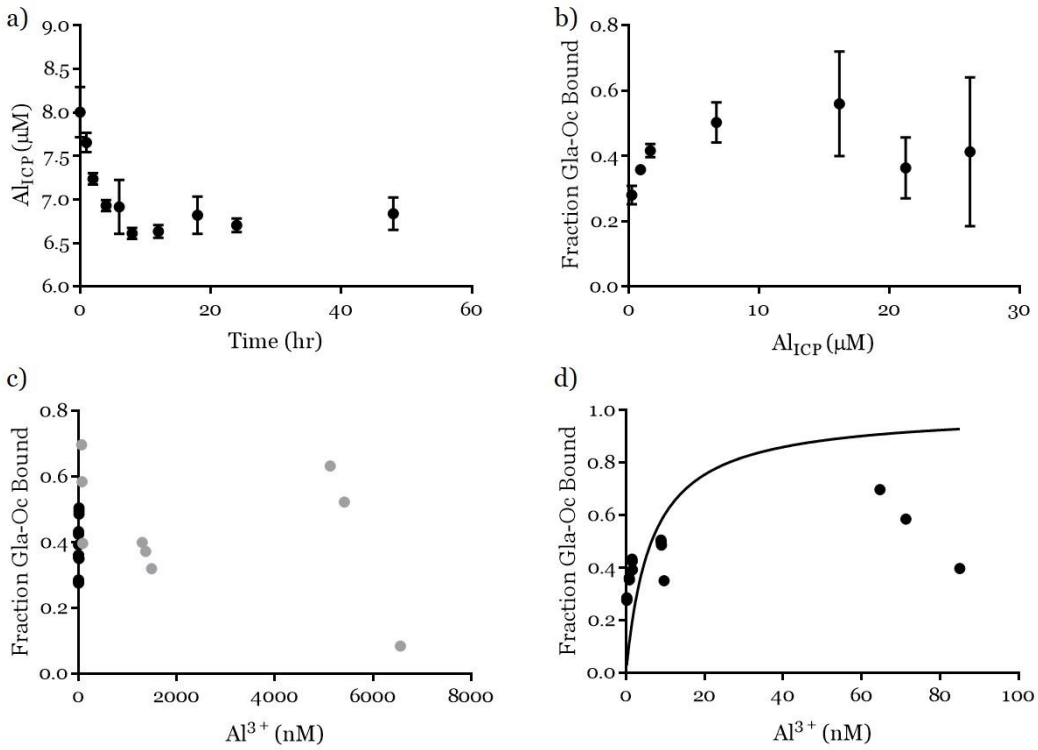


Figure 5.7. Sample Time Course Equilibrium, and Initial Dialysis and ICP-MS Experiment with Single-Site Binding Assumption. a) Time course to establish time to reach equilibrium with Gla-Oc binding Al^{3+} following the dialysis and ICP-MS procedure. n = 4. b) Initial binding curve, where the fraction of Gla-Oc bound was plotted against the Al concentrations measured by ICP-MS (Al_{ICP}), with n = 3. c) Fraction of Gla-Oc bound was plotted against the Al^{3+} concentrations calculated for each individual data point, with the data points in black shown not to fit to the single-site binding equation in d). Greyed data points are the individual replicates for cAl concentrations of 17.8, 22.3, and 27.4 μM.

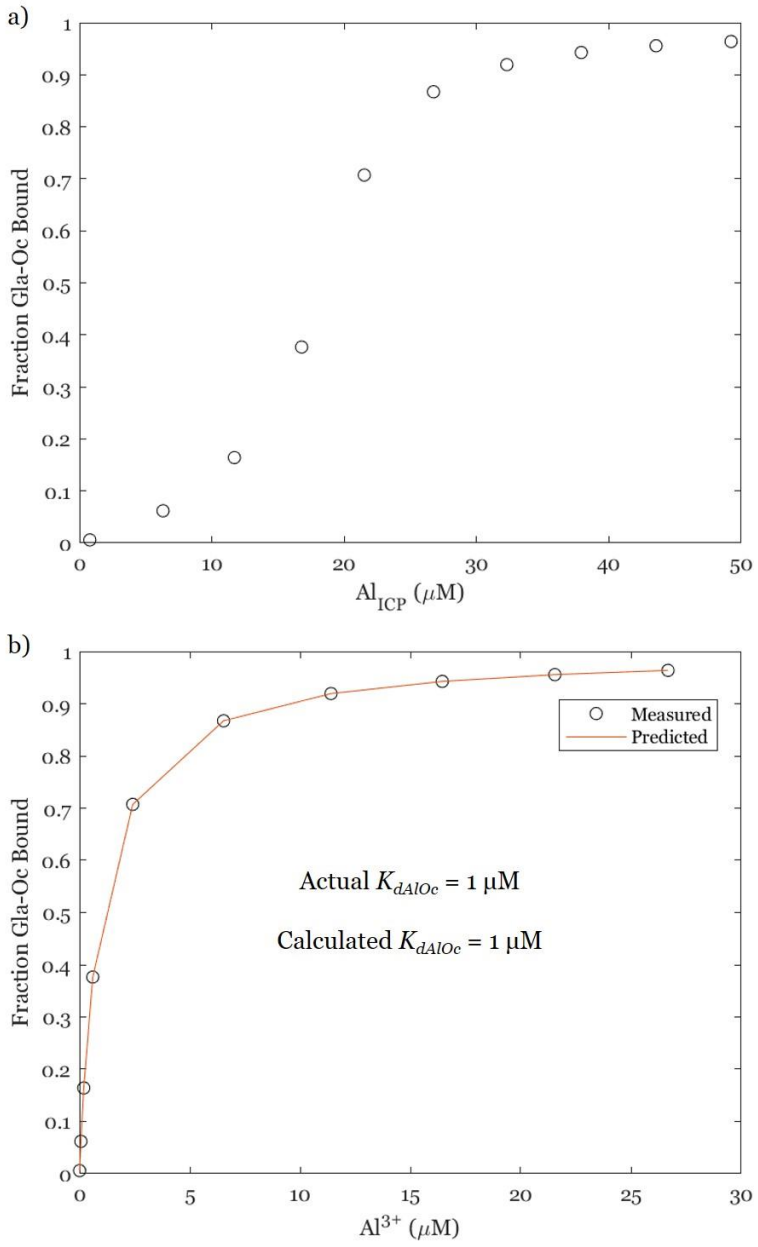


Figure 5.8. Single-site Binding Simulated Data and Non-Linear Regression Curve Fitting Validation. a) Simulated data when K_{dAlOc} = 1 µM for Gla-Oc binding to Al^{3+}, with the fraction of Gla-Oc bound plotted against the concentration of Al measured by ICP-MS (Al_{ICP}). b) The concentration of free Al^{3+} was calculated and plotted on the x-axis, enabling non-linear regression (NLR) curve fitting of the simulated data using the single-site binding model described in **Equation 5.27**. The fraction of Gla-Oc bound determined when developing simulated data (measured) was plotted with the fraction of Gla-Oc bound determined during the NLR curve fitting analysis (predicted). The actual dissociation constant assigned to the system when simulating the data (K_{dAlOc} = 1 µM) was then compared to the dissociation constant determined by NLR (K_{dAlOc} = 1 µM) to validate the approach before proceeding to the more complex 2:1 Gla-Oc: Al^{3+} binding system.

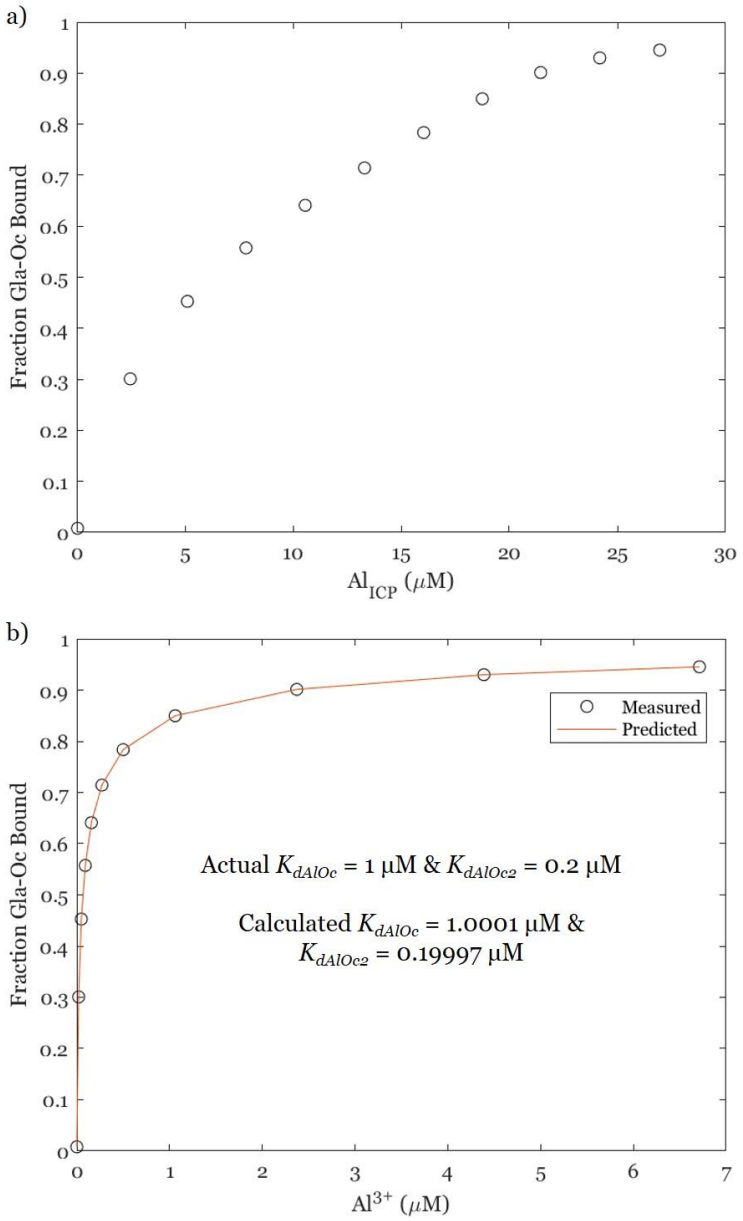


Figure 5.9. 2:1 Gla-Oc:Al^{3+} Binding Simulated Data and Non-Linear Regression Curve Fitting Validation. a) Simulated data when K_{dAlOc} = 1 µM and K_{dAlOc2} = 0.2 µM for two Gla-Oc binding to one Al^{3+}, with the fraction of Gla-Oc bound plotted against the concentration of Al measured by ICP-MS ($[Al]_{ICP}$). b) The concentration of free Al^{3+} was calculated and plotted on the x-axis, enabling non-linear regression (NLR) curve fitting of the simulated data using the two-site binding model described in **Equation 5.28**. The fraction of Gla-Oc bound determined when developing simulated data (measured) was plotted with the fraction of Gla-Oc bound determined during the NLR curve fitting analysis (predicted). The actual dissociation constants assigned to the system when simulating the data (K_{dAlOc} = 1 µM & K_{dAlOc2} = 0.2 µM) were then compared to the dissociation constants determined by NLR (K_{dAlOc} = 1.0001 µM & K_{dAlOc2} = 0.19997 µM) to validate the approach.

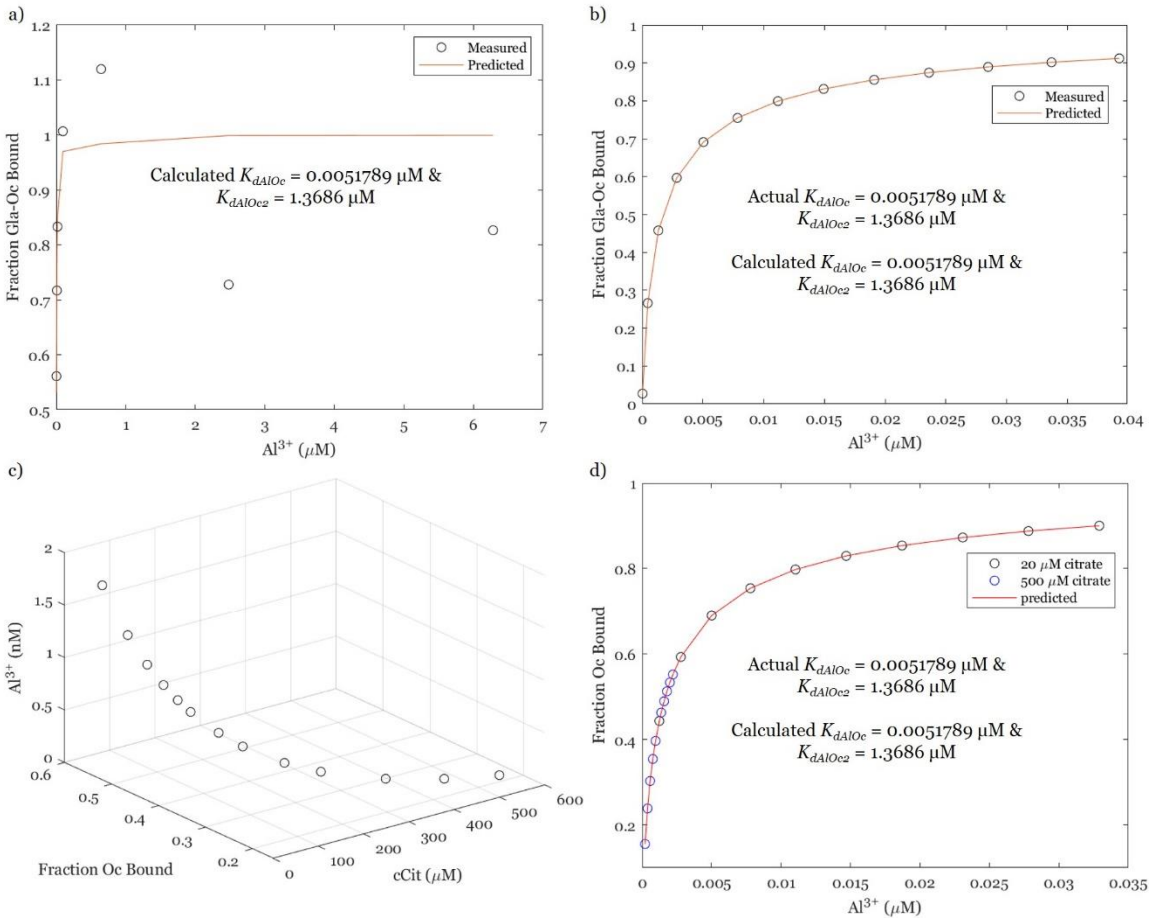


Figure 5.10. Initial Dialysis and ICP-MS Experiment and Citrate Concentration Optimization. a) Initial dialysis and ICP-MS data plotted with the assumption that the concentration of Oc bound is equal to 2 x the concentration of Al^{3+} bound and fit using NLR with the two-site binding model described in **Equation 5.28**. This was completed to make initial estimates of the binding affinities and to assist with further protocol optimization. b) Simulated data generated with K_{dAlOc} and K_{dAlOc2} defined based on the initial estimates determined in a) demonstrate the capability of the NLR technique to accurately predict the binding affinities. c) Simulated data generated using the estimated K_{dAlOc} and K_{dAlOc2} values, cAl = 1 µM, cOc = 2.885 µM, and cCit = 1 x 10^{-6} – 500 µM demonstrating the effect of increased cCit concentrations on Al^{3+} concentration and the fraction of Oc bound. d) Simulated data with the same K_{dAlOc} and K_{dAlOc2} values, minimum cAl = 1 µM, cOc = 2.885 µM, and cCit = 20 µM (black circles) or 500 µM (blue circles). Data were verified to fit using NLR curve fitting using the two-site binding model.

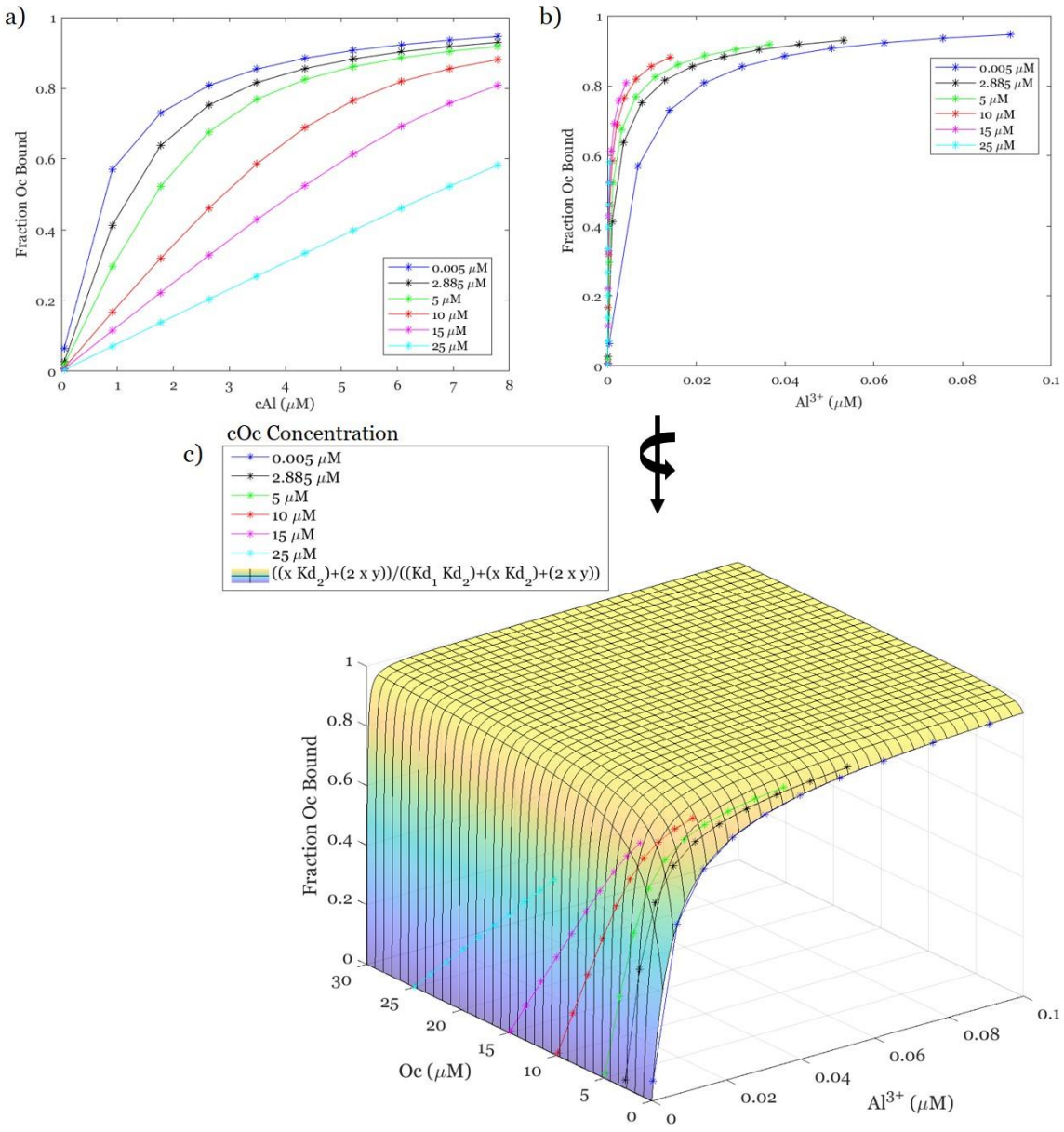


Figure 5.11. Overcoming the Titration Effect by Determining Al^{3+} and Oc Concentrations. The titration effect is exemplified in a), where increasing cOc concentrations, combined with the assumption that cAl ≈ Al^{3+} dramatically affects the binding curve observed, and thus also the calculated K_d. Plotting the fraction of Oc bound versus free Al^{3+} in b) at first glance does not appear to alleviate the titration effect issues, although taken together with known free Oc concentrations in c), each cOc data set is found to lie on the surface defined by the two-site binding equation described by **Equation 5.28**, along with the assigned K_{dAlOc} = 0.0051789 μM and K_{dAlOc2} = 1.3686 μM. Data sets were generated using cCit = 20 μM, cAl = 0.05 – 7.7945 μM, cOc = 0.005, 2.885, 5, 10, 15, and 25 μM, and the K_d values described above.

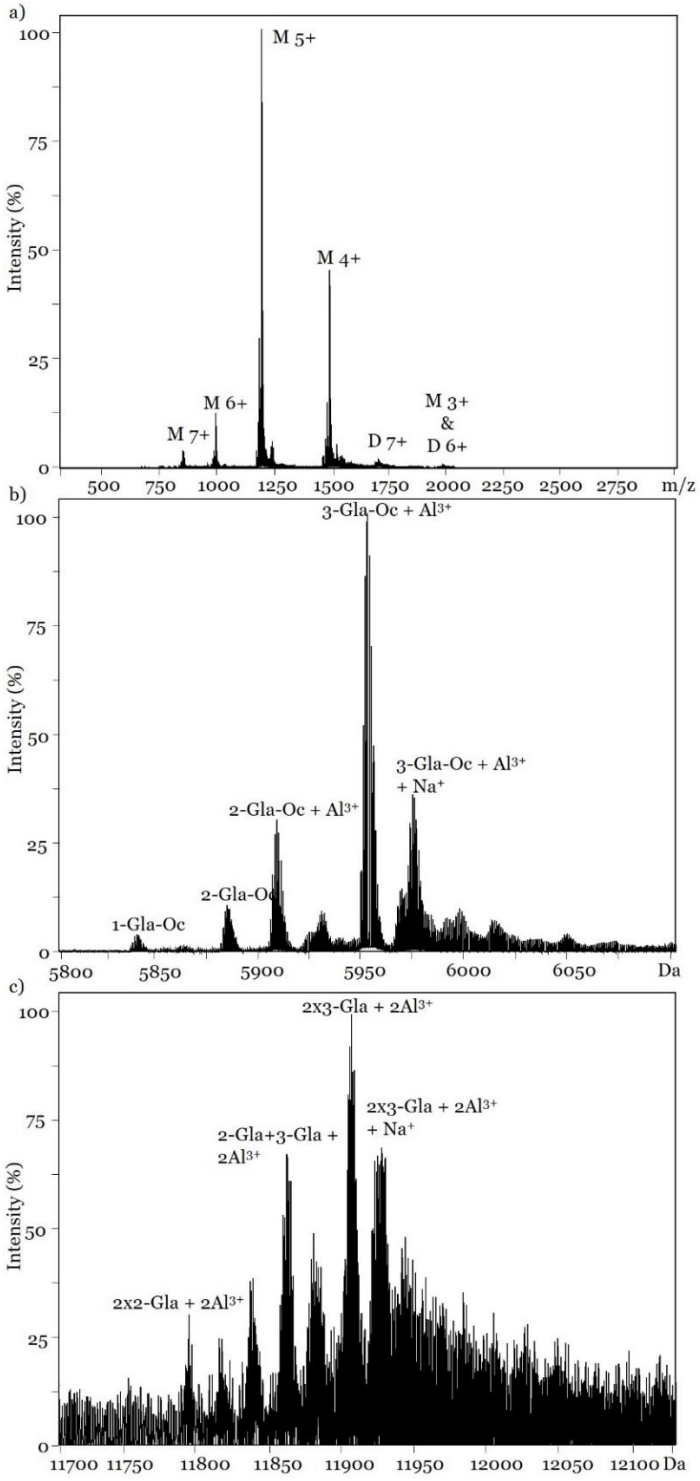


Figure 5.12. Direct Infusion MS of Gla-Oc Sample Incubated with Al Reveals Low-Abundance Al_2Oc_2 Dimer. a) Full direct infusion MS spectrum of monomer (M) and dimer (D) charge states. Charge deconvoluted spectrum in a) focusing on the b) monomer, or c) dimer, Gla-Oc mass spectral peaks.

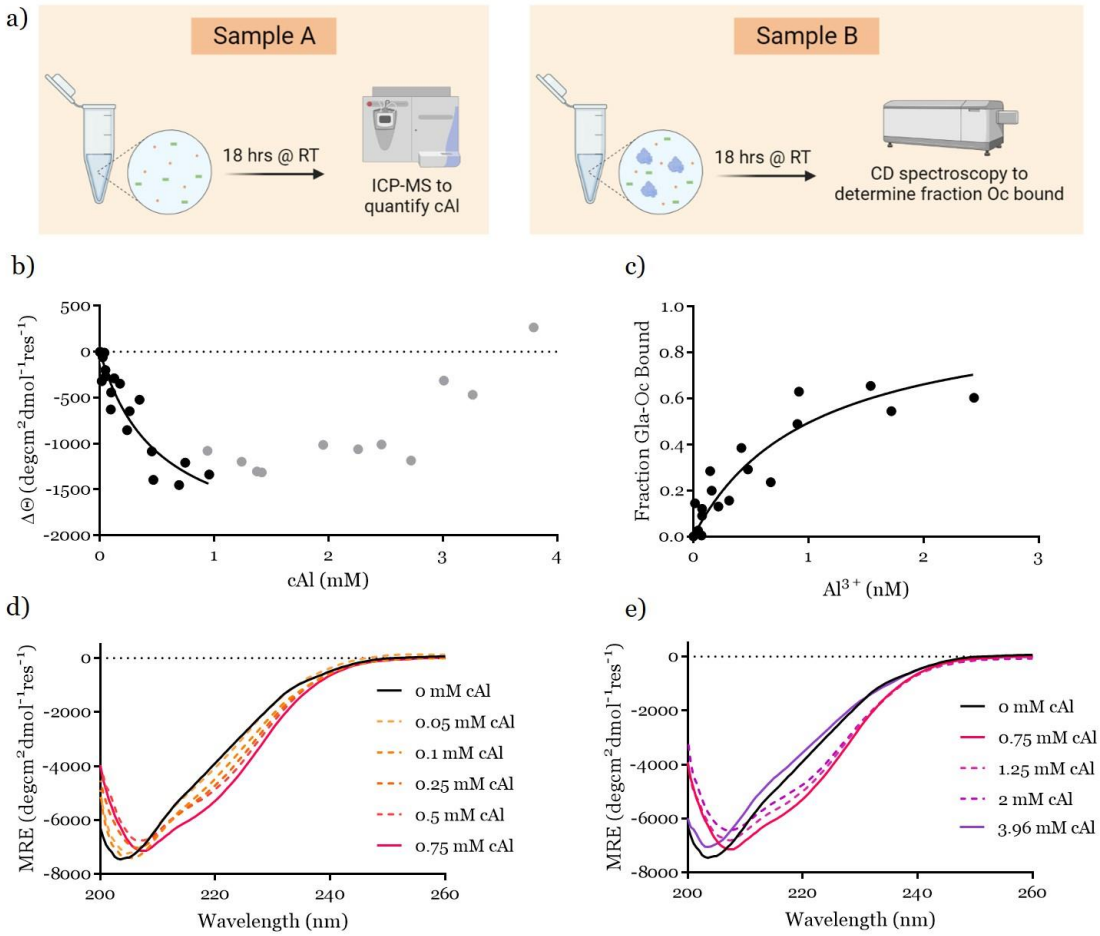


Figure 5.13. Binding Data for Al^{3+} Binding to Gla-Oc Using CD and ICP-MS. a) Experimental set-up consisted of side-by-side samples containing 5 mM citrate + 150 mM NaCl, pH 4.5, cAl titrations, without (Sample A), or with (Sample B), 50 µM Gla-Oc. Sample A was used as both a buffer blank during CD spectroscopy, and to quantify cAl by ICP-MS. Sample B was analyzed by CD spectroscopy. b) The change in mean residue ellipticity ($\Delta\Theta$) was plotted against measured cAl values to estimate the maximum change in mean residue ellipticity ($\Delta\Theta_{max}$), with black data points included in the estimate. Grey data points were excluded based on results from the CAS experiment. c) Binding curve based on data points included in b), and single-site binding assumption. Overlayed CD spectra for Gla-Oc titrated with d) 0, 0.05, 0.1, 0.25, 0.5, and 0.75 mM cAl, and e) 0, 0.75, 1.25, 2, and 3.96 mM cAl, from 200 – 260 nm.

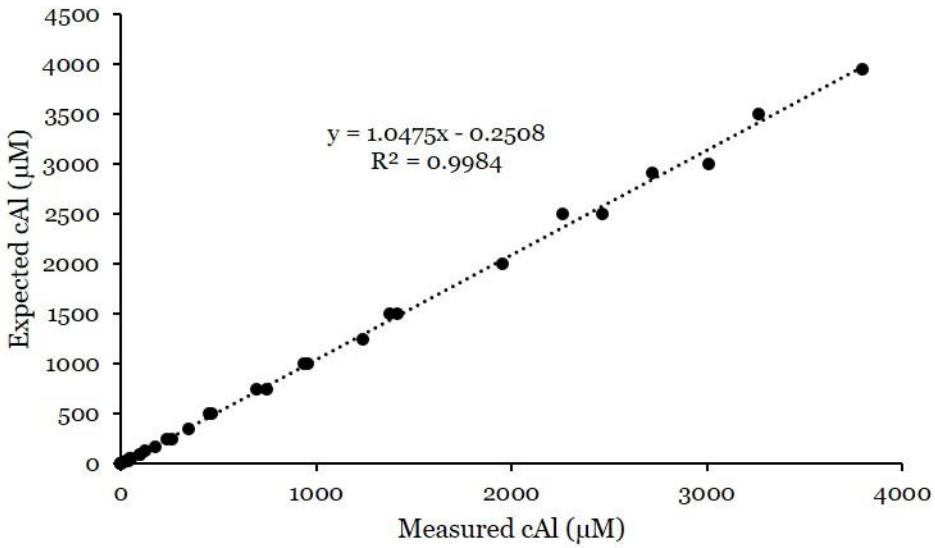


Figure 5.14. Comparison of Expected and Measured cAl Values by ICP-MS.

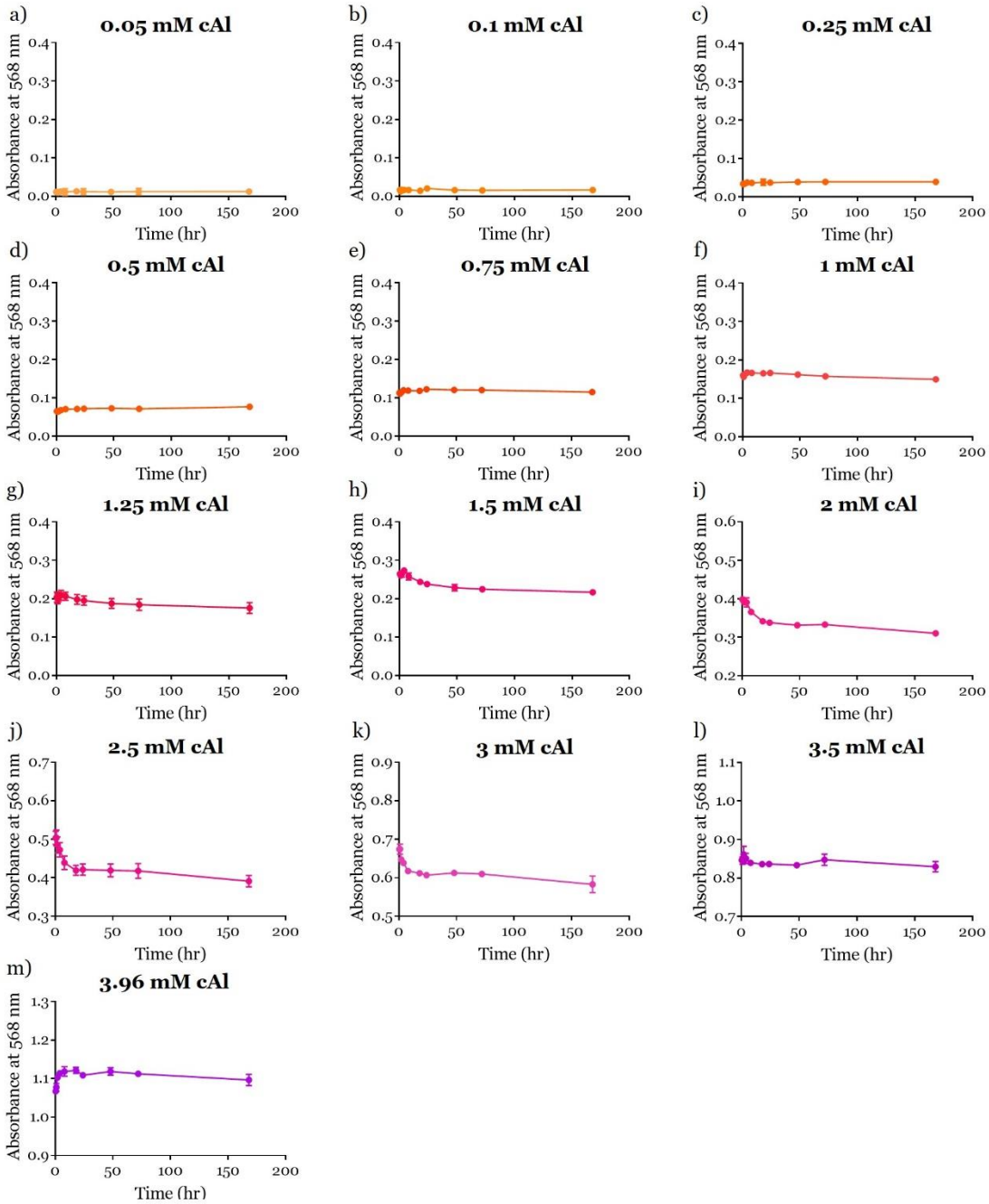


Figure 5.15. Chrome Azurol S Timecourse to Evaluate Changes in Al^{3+} Concentrations Over Time. Total cAl concentrations of 0.05 – 3.96 mM (a – m) were incubated with 187.5 µM chrome azurol S (CAS) at room temperature for 0.5, 1, 2, 4, 8, 18, 24, 48, 72, and 168 hr, with absorbance of sample aliquots measured at 568 nm and blank subtracted using 0 mM cAl samples containing 187.5 µM CAS. n = 3 replicates per time point.

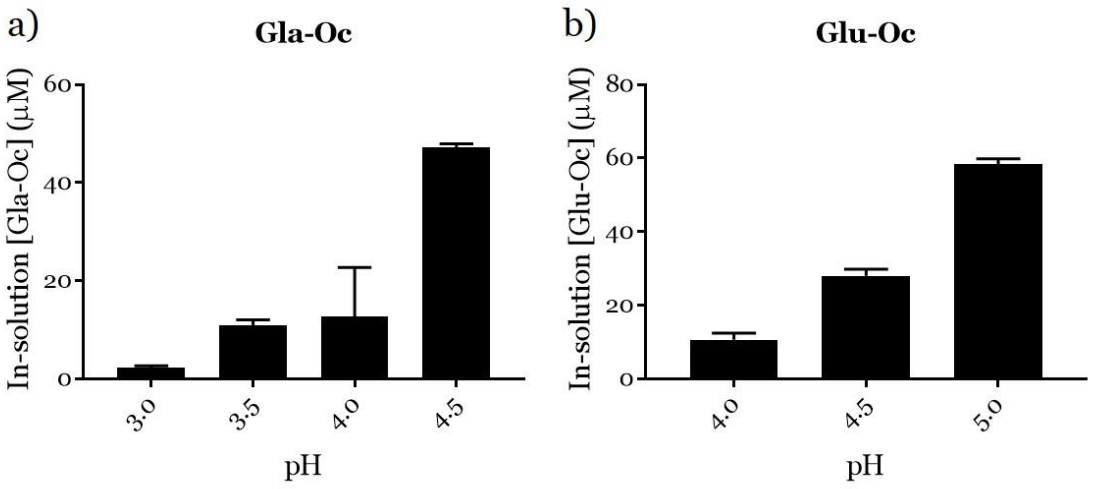


Figure 5.16. In-Solution Oc Concentrations Change Due to Changes in pH. a) Gla-Oc in-solution concentrations at pH 3.0, 3.5, 4.0, and 4.5, and b) Glu-Oc in-solution concentrations at pH 4.0, 4.5, and 5.0 after 18 hrs incubation at room temperature. All Gla-Oc and Glu-Oc comparison pH samples were prepared from the same stock solution, with a 50 μM target concentration in the diluted samples. Samples were centrifuged for 5 min at 13, 300 rpm, with white precipitates observed in Gla-Oc pH 3.0, 3.5, and 4.0, and Glu-Oc pH 4.0 and 4.5 samples. n = 2 and error bars represent SD.

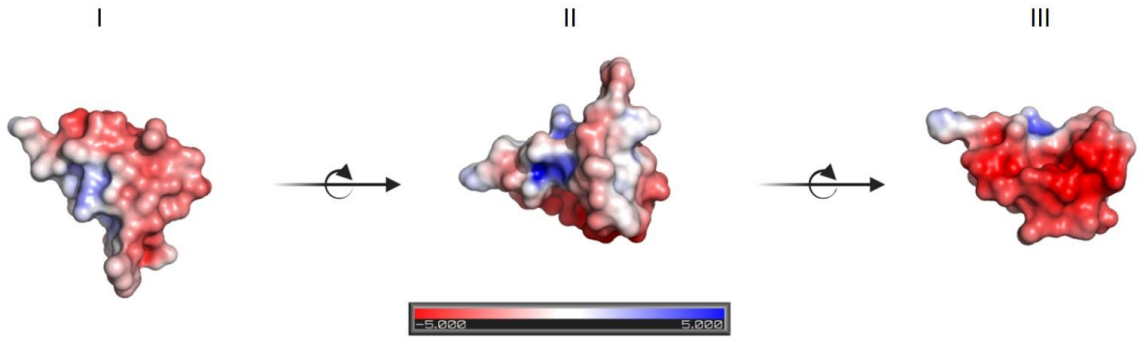


Figure 5.17. Surface Charge Plot of Gla-Oc at pH 7.4. Rotations of the porcine Gla-Oc crystal structure, depicting negative surface charge (red) and positive surface charge (blue), with the Gla residues shown in rotation III. Crystal structure obtained from PDB 1q8h.

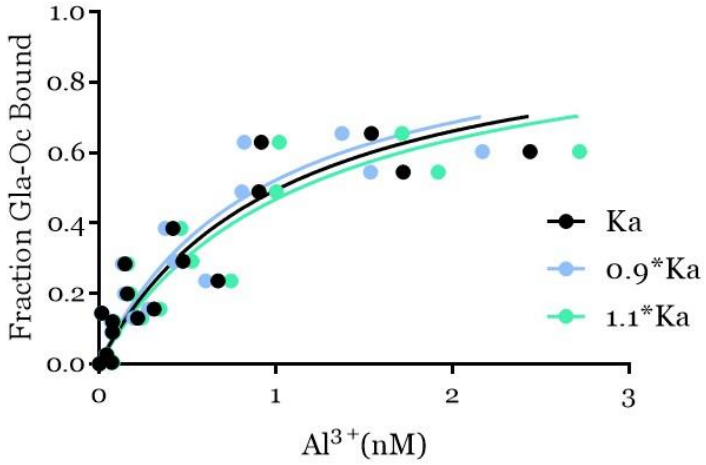


Figure 5.18. Error Propagation with 10 % Error in Literature K_a Values. The calculated K_{dAlOc} = 1.021 ± 0.1225 nM (black data set), based on solving for Al^{3+} using the Al-OH and Al-citrate complex K_a values reported previously [199, 200], and the effect of introducing -10 % (light blue data set), and +10 % (light green data set) error to the reported K_a values on the calculated K_{dAlOc} = 0.9135 ± 0.1094 nM (-10 % error), and K_{dAlOc} = 1.134 ± 0.1364 nM (+10 % error).

Table 5.1. Known Association Constants (K_a) and Their References.

Complexes	Quotient	K_a (M)	Reference
Citrate-H	$\frac{HCit^{2-}}{Cit^{3-} * H^+}$	$K_{HCit2-} = 5.012 \times 10^5$	[199]
	$\frac{H_2Cit^-}{HCit^{2-} * H^+}$	$K_{H2Cit-} = 2.239 \times 10^4$	[199]
	$\frac{H_3Cit}{H_2Cit^- * H^+}$	$K_{H3Cit} = 7.943 \times 10^2$	[199]
Citrate-Al	$\frac{AlCit}{Al^{3+} * Cit^{3-}}$	$K_{AlCit} = 9.55 \times 10^7$	[199]
	$\frac{AlCitH^+}{AlCit * H^+}$	$K_{AlCitH+} = 8.71 \times 10^2$	[199]
	$\frac{AlCit}{AlCitH^- * H^+}$	$K_{AlCitH-} = 2.042 \times 10^3$	[199]
	$\frac{AlCit_2^{3-}}{AlCit * Cit^{3-}}$	$K_{AlCit23-} = 1.0 \times 10^5$	[200]
	$\frac{AlCit^-H_-}{Al(OH)(CitH_-)^{2-} * H^+}$	$K_{Al(OH)CitH-^2-} = 1.698 \times 10^6$	[199]
	$\frac{Al_3(OH)_4Cit_3^{4-} * [H^+]^4}{[Al^{3+}]^3 * [Cit^{3-}]^3}$	$\beta_{Al3(OH)4Cit3^4-} = 2.692 \times 10^{14}$	[199]
	$\frac{Al_3(OH)_4(CitH_-)_3{}^{7-} * [H^+]^7}{[Al^{3+}]^3 * [Cit^{3-}]^3}$	$\beta_{Al3(OH)4(CitH-)3^7-} = 1.23 \times 10^{-11}$	[199]
Al-OH	$\frac{Al(OH)^{2+}}{Al^{3+} * OH^-}$	$K_{Al(OH)^2+} = 2.951 \times 10^8$	[199]
	$\frac{Al(OH)_2^+}{Al(OH)^{2+} * OH^-}$	$K_{Al(OH)2^+} = 2.138 \times 10^8$	[199]
	$\frac{Al(OH)_3}{Al(OH)_2^+ * OH^-}$	$K_{Al(OH)3} = 7.943 \times 10^7$	[199]
	$\frac{Al(OH)_4^-}{Al(OH)_3 * OH^-}$	$K_{Al(OH)4^-} = 6.309 \times 10^6$	[199]
	$\frac{Al_2(OH)_2^{4+}}{[Al^{3+}]^2 * [OH^-]^2}$	$\beta_{Al2(OH)2^4+} = 1 \times 10^{20}$	[199]
	$\frac{Al_3(OH)_4^{5+}}{[Al^{3+}]^3 * [OH^-]^4}$	$\beta_{Al3(OH)4^5+} = 1 \times 10^{42}$	[199]

Table 5.2. Individually Calculated K_{dAlOc} Values for Dialysis & ICP-MS Data.

	K_{dAlOc} (nM)		
cAl (μM)	Rep 1	Rep 2	Rep 3
1.028	0.454	0.497	0.546
1.928	1.442	1.451	1.525
2.859	2.573	2.017	2.122
8.005	9.109	9.530	8.729
17.788	129.100	28.035	50.559
22.319	1957.200	3163.200	2303.800
27.403	2980.400	4938.400	71554.000

Table 5.3. Simulated and Solved Values for Single-Site Binding with K_d = 1 µM. The simulated total concentration of Al (cAl), concentration of Al not bound by Oc ($[Al]_{ICP}$), and concentration of free Al^{3+} ($[Al^{3+}]_s$) using **Equations 5.3 -5.19**, **5.21**, **5.22**, and **5.24** and K_d = 1 µM are depicted in the left set of columns. From these data, it is possible to calculate the concentration of Oc bound by Al^{3+} ([AlOc]) and the fraction of Oc bound by Al^{3+}, which is shown in the middle set of columns. The simulated cAl, $[Al]_{ICP}$, and calculated [AlOc], along with the defined cCit = 20 µM, cOc = 2.885 µM, $[H^+]$ = 31.6 µM, $[OH^-]$ = 0.000316 µM, and these same equations are used to verify the ability to accurately calculate the concentration of free Al^{3+} in the system, and to estimate K_d for each cAl tested. Concentrations in µM.

Simulated			Calculated		Solved	
cAl	$[Al]_{ICP}$	$[Al^{3+}]_s$	[AlOc]	Fraction Bound	$[Al^{3+}]_{free}$	K_d
0.8	0.783157	0.005872	0.016843	0.005838	0.005872	1
6.49	6.311615	0.065907	0.178385	0.061832	0.065907	1
12.18	11.70656	0.19632	0.473437	0.164103	0.19632	1
17.87	16.78365	0.603981	1.08635	0.376551	0.60398	1
23.56	21.51982	2.414914	2.040176	0.707167	2.4149	1
29.25	26.74802	6.532218	2.501979	0.867237	6.5322	1
34.94	32.28761	11.40251	2.652386	0.919371	11.403	1
40.63	37.91037	16.44584	2.719631	0.94268	16.446	1
46.32	43.56295	21.54841	2.757053	0.955651	21.548	1
52.01	49.22924	26.67763	2.780764	0.96387	26.678	1

Table 5.4. Comparison of Simulated Fraction Bound, and Fraction Bound Estimated by Non-Linear Regression, for Single-Site Binding with K_d = 1 µM. The simulated concentration of free Al^{3+} values ($[Al^{3+}]_s$) were plotted against the simulated fraction bound values and subjected to non-linear regression curve fitting using a single-site binding model. The fraction bound estimated by non-linear regression (NLR) curve fitting were not statistically significantly different from the simulated data ($p > 0.05$), and the dissociation constant determined from NLR was K_d = 1 µM. Concentrations in µM.

$[Al^{3+}]_s$	Simulated Fraction Bound	NLR Fraction Bound
0.005872	0.005838	0.005838
0.065907	0.061832	0.061832
0.19632	0.164103	0.1641
0.603981	0.376551	0.37655
2.414914	0.707167	0.70717
6.532218	0.867237	0.86724
11.40251	0.919371	0.91937
16.44584	0.94268	0.94268
21.54841	0.955651	0.95565
26.67763	0.96387	0.96387

Table 5.5. Simulated and Solved Values for 2:1 Gla-Oc:Al^{3+} Binding with K_{d1} = 1 µM and K_{d2} = 0.2 µM. The simulated total concentration of Al (cAl), concentration of Al not bound by Oc ($[Al]_{ICP}$), concentration of Oc dimer bound to one Al^{3+} ($[AlOc_2]$), concentration of free Al^{3+} ($[Al^{3+}]_s$), and concentration of free Oc ($[Oc]_s$) using **Equations 5.3 – 5.21**, **5.23**, and **5.25** and K_d = 1 µM are depicted in the left set of columns. The simulated cAl, $[Al]_{ICP}$, and $[AlOc_2]$, and these same equations are used to verify the ability to accurately calculate the concentration of free Al^{3+} ($[Al^{3+}]$), free Oc ($[Oc]_{free}$), and the fraction of Oc bound in the system, and to estimate K_{d1} and K_{d2} for each cAl tested, which are depicted in the right set of columns. Concentrations in µM.

Simulated					Solved				
cAl	$[Al]_{ICP}$	$[AlOc_2]$	$[Al^{3+}]_s$	$[Oc]_s$	$[Al^{3+}]_{free}$	$[Oc]_{free}$	Fraction Bound	K_{d1}	K_{d2}
0.05	0.037981	0.011233	0.000274	2.861748	0.000274	2.8617	0.00806	0.99881	0.20024
2.93	2.475581	0.41343	0.020321	2.017153	0.020321	2.0172	0.30081	1.0001	0.19999
5.81	5.118001	0.61419	0.049281	1.578807	0.049281	1.5788	0.45275	0.99995	0.20001
8.69	7.827395	0.74577	0.091518	1.276625	0.091518	1.2766	0.5575	1	0.2
11.57	10.56395	0.84322	0.157212	1.035723	0.15721	1.0357	0.64099	1	0.2
14.45	13.30803	0.91895	0.270627	0.824089	0.27063	0.82408	0.71436	0.99999	0.2
17.33	16.0436	0.97429	0.499931	0.624314	0.49993	0.62431	0.7836	1	0.2
20.21	18.75441	0.99599	1.060415	0.433416	1.0604	0.43342	0.84977	1	0.2
23.09	21.45217	0.96228	2.371294	0.284887	2.3713	0.28489	0.90125	1	0.2
25.97	24.18482	0.89758	4.388656	0.202248	4.3887	0.20224	0.9299	0.99996	0.19999
28.85	26.95715	0.83446	6.712002	0.157686	6.712	0.15769	0.94534	1	0.20001

Table 5.6. Comparison of Simulated Fraction Bound, and Fraction Bound Estimated by Non-Linear Regression, for 2:1 Gla-Oc:Al^{3+} Binding with K_{d1} = 1 µM and K_{d2} = 0.2 µM. The simulated concentration of free Al^{3+} values ($[Al^{3+}]_s$) and free Oc values ($[Oc]_s$) were plotted against the simulated fraction bound values and subjected to non-linear regression curve fitting using a single-site binding model. Concentrations in µM.

$[Al^{3+}]_s$	$[Oc]_s$	Simulated Fraction Bound	NLR Fraction Bound
0.000274	2.861748	0.00806	0.00806
0.020321	2.017153	0.300814	0.30082
0.049281	1.578807	0.452753	0.45275
0.091518	1.276625	0.557496	0.5575
0.157212	1.035723	0.640997	0.641
0.270627	0.824089	0.714354	0.71435
0.499931	0.624314	0.7836	0.7836
1.060415	0.433416	0.849769	0.84977
2.371294	0.284887	0.901252	0.90125
4.388656	0.202248	0.929897	0.92989
6.712002	0.157686	0.945343	0.94534

Table 5.7. Individually Calculated K_{dAlOc} Values for CD & ICP-MS Data.

	cAl (μM)	K_{dAlOc} (nM)
Day 1	0.527	-
	48.081	0.773
	95.615	0.367
	238.235	0.666
	454.492	0.941
	692.475	0.808
	956.569	1.598
Day 2	0.316	-
	49.047	0.552
	98.519	0.637
	260.276	1.147
	467.331	0.539
	745.936	1.437
Day 3	0.275	-
	15.527	0.085
	25.171	1.517
	40.237	15.142
	125.845	1.448
	174.936	1.676
	346.455	2.176

CHAPTER 6

CONCLUSIONS AND FUTURE DIRECTIONS

The work presented in this dissertation offers new insights into the qualitative binding properties of human Oc binding to various divalent and trivalent metal cations that are present in the bone. Novel metal binding partners described here include Mn^{2+}, Fe^{3+}, and Cr^{3+}, while the previously determined binding of Ca^{2+}, Mg^{2+}, Pb^{2+}, and Al^{3+} to Oc were reproduced here. More importantly, all metal cation binding was found qualitatively to be Gla-dependent, and thus vitamin K dependent. Future studies should focus on the biological significance of Oc binding metals other than Ca^{2+} to gain a better understanding of both the function of Oc in the bone, and potentially in circulation, and how normal Oc function may be disrupted by abnormal quantities of these competitive binding metals in bone. Quantitative equilibrium binding studies of Oc, parsing between Glu-Oc and Gla-Oc, with these metals will reveal the extent to which these metals may compete with Ca^{2+} for binding to Oc, and thus, combined with *in vivo* studies, the degree to which each metal may affect normal physiological function in a manner that depends on vitamin K intake.

The observation of higher-order homomeric complexes of Gla-Oc and Glu-Oc in the absence of NaCl *in vitro*, as mentioned previously, is unlikely to be of any biological significance due to the tight control of physiological levels of Na^{+} at approximately 136 mM in the ECF [138]. These studies did, however, highlight the importance of use of near physiological NaCl concentrations during *in vitro* work to maintain proper protein quaternary structure.

The ability of human Oc to form filaments in the presence of increased titrations of Ca^{2+} was shown to be Gla- and pH-dependent, offering an explanation for the conditions that govern filamentation. Additionally, the reversibility of filamentation was demonstrated here, further supporting the potential physiological relevance of

filamentation. Based on these findings, it was hypothesized that Oc filamentation may act as a storage mechanism for Oc and/or Ca^{2+} in the bone. Future work addressing this hypothesis should focus on *in vivo* studies to confirm Oc filament structures occur within the bone. Additional work to obtain cryo-EM 3-D reconstructions of the filaments would also be beneficial to understand the mechanism by which filamentation occurs.

Finally, through the use of a combination of CD spectroscopy and ICP-MS, it was possible to quantify the equilibrium dissociation constant for Al^{3+} binding to human 3-Gla-Oc (K_d = 1.0 ± 0.12 nM) in the presence of the competitive binder, citrate, with known association constants for Al-citrate, and Al-OH, complexes [199, 200]. Generation of a Matlab script to solve the complex system of twenty linear equilibrium and mass balance equations enabled non-linear regression curve fitting of the data to a single-site binding model. Analyzing Gla-Oc binding at pH 4.5 is physiologically relevant since this is the pH of the osteoclast resorption lacunae during bone resorption [41]. Complementary experiments to obtain the equilibrium dissociation constant for Al^{3+} binding human Glu-Oc were not possible due to Glu-Oc precipitation at pH 4.5, which is just below its known pI = 4.6 [110]. Qualitatively, the binding of Gla-Oc to Al^{3+} occurs with higher affinity than Glu-Oc, reinforcing the vitamin K dependence of this interaction. The high affinity of Al^{3+} binding to human Gla-Oc is highly suggestive of this toxic metal's capability to compete with Ca^{2+} for binding to Gla-Oc in the bone. Due to the known deposition of Al in the bone, the Gla-dependent nature of this binding interaction, and the ability of Oc to be released from the bone into circulation and excreted by glomerular filtration in the kidneys, future studies on this interaction should focus on the kinetics of this binding interaction and evaluate *in vivo* the effect of suppression of γ-carboxylation of Oc on excretion of Al^{3+} into the urine.

Together, the work presented here offered further insight into the binding and structural properties of Oc. It is through gradual, systematic studies that discoveries are made, and the work presented here improves the understanding of this highly abundant bone protein to build on past research, and act as a steppingstone for future, research in this field.

REFERENCES

[1] Prigodich R V & Vesely M R. (1997). Characterization of the complex between bovine osteocalcin and type I collagen. *Arch Biochem Biophys*. **345**, 339-41.

[2] Desbois C, Hogue D A, Karsenty G. (1994). The mouse osteocalcin gene cluster contains three genes with two separate spatial and temporal patterns of expression. *J Biol Chem*. **269**, 1183-90.

[3] Sato M & Tada N. (1995). Preferential expression of osteocalcin-related protein mRNA in gonadal tissues of male mice. *Biochem Biophys Res Commun*. **215**, 412-21.

[4] Al Rifai O, Chow J, Lacombe J, *et al*. (2017). Proprotein convertase furin regulates osteocalcin and bone endocrine function. *J Clin Invest*. **127**, 4104-17.

[5] Nielsen-Marsh C M, Richards M P, Hauschka P V, *et al*. (2005). Osteocalcin protein sequences of Neanderthals and modern primates. *Proc Natl Acad Sci U S A*. **102**, 4409-13.

[6] Rehder D S, Gundberg C M, Booth S L & Borges C R. (2015). Gamma-carboxylation and fragmentation of osteocalcin in human serum defined by mass spectrometry. *Mol Cell Proteomics*. **14**, 1546-55.

[7] Poser J W & Price P A. (1979). A method for decarboxylation of gamma-carboxyglutamic acid in proteins. Properties of the decarboxylated gamma-carboxyglutamic acid protein from calf bone. *J Biol Chem*. **254**, 431-6.

[8] Al Rifai O, Julien C, Lacombe J, *et al*. (2020). The half-life of the bone-derived hormone osteocalcin is regulated through *O*-glycosylation in mice, but not in humans. *Elife*. **9**, e61174.

[9] Negre-Salvayre A, Salvayre R, Auge N, *et al*. (2009). Hyperglycemia and glycation in diabetic complications. *Antioxid Redox Signal*. **11**, 3071-109.

[10] Singh V P, Bali A, Singh N, *et al*. (2014). Advanced glycation end products and diabetic complications. *Korean J Physiol Pharmacol*. **18**, 1-14.

[11] Rubin M R, Paschalis E P, Poundarik A, *et al*. (2016). Advanced glycation endproducts and bone material properties in type 1 diabetic mice. *PLoS ONE*. **11**, e0154700.

[12] Gundberg C, Anderson M, Dickson I, *et al*. (1986). "Glycated" osteocalcin in human and bovine bone. The effect of age. *J Biol Chem*. **261**, 14557-61.

[13] Thomas C J, Cleland T P, Zhang S, *et al*. (2017). Identification and characterization of glycation adducts on osteocalcin. *Anal Biochem*. **525**, 46-53.

[14] Tavakol M & Vaughan T J. (2020). The structural role of osteocalcin in bone biomechanics and its alteration in Type-2 diabetes. *Sci Rep*. **10**, 17321.

[15] Shearer M J, Shearer P N & Newman P. (2008). Metabolism and cell biology of vitamin K. *Thromb Haemost*. **100**, 530-47.

[16] Fusaro M, Mereu M C, Aghi A, *et al.* (2017). Vitamin K and bone. *Clin Cases Miner Bone Metab*. **14**, 200-6.

[17] Palermo A, Tuccinardi D, D'Onofrio L, *et al.* (2017). Vitamin K and osteoporosis: myth or reality? *Metabolism*. **70**, 57-71.

[18] Akbari S & Rasouli-Ghahroudi A A. (2018). Vitamin K and bone metabolism: a review of the latest evidence in preclinical studies. *Biomed Res Int*. **2018**, 4629383.

[19] Rodríguez-Olleros Rodríguez C & Díaz Curiel M. (2019). Vitamin K and bone health: A review on the effects of vitamin K deficiency and supplementation and the effect of non-vitamin K antagonist oral anticoagulants on different bone parameters. *J Osteoporos*. **2019**, 2069176.

[20] Booth S L, Tucker K L, Chen H, *et al.* (2000). Dietary vitamin K intakes are associated with hip fracture but not with bone mineral density in elderly men and women. *Am J Clin Nutr*. **71**, 1201-8.

[21] Rejnmark L, Vestergaard P, Charles P, *et al.* (2006). No effect of vitamin K1 intake on bone mineral density and fracture risk in perimenopausal women. *Osteoporos Int*. **17**, 1122-32.

[22] Booth S L, Broe K E, Gagnon D R, *et al.* (2003). Vitamin K intake and bone mineral density in women and men. *Am J Clin Nutr*. **77**, 512-6.

[23] Booth S L, Broe K E, Peterson J W, *et al.* (2004). Associations between vitamin K biochemical measures and bone mineral density in men and women. *J Clin Endocrinol Metab*. **89**, 4904-9.

[24] Feskanich D, Weber P, Willett W C, *et al.* (1999). Vitamin K intake and hip fractures in women: a prospective study. *Am J Clin Nutr*. **69**, 74-9.

[25] Hao G, Zhang B, Gu M, *et al.* (2017). Vitamin K intake and the risk of fractures: a meta-analysis. *Medicine (Baltimore)*. **96**, e6725.

[26] Fonseca H & Moreira-Gonçalves D. (2014). Bone quality: the determinants of bone strength and fragility. *Sports Med*. **44**, 37-53.

[27] Hauschka P V & Carr S A. (1982). Calcium-dependent alpha-helical structure in osteocalcin. *Biochemistry*. **21**, 2538-47.

[28] Cristiani A, Maset F, De Toni L, *et al.* (2014). Carboxylation-dependent conformational changes of human osteocalcin. *Front Biosci (Landmark Ed)*. **19**, 1105-16.

[29] Dowd T L, Rosen J F, Mints L, Gundberg C M. (2001). The effect of Pb^{2+} on the structure and hydroxyapatite binding properties of osteocalcin. *Biochim Biophys Acta*. **1535**, 153-63.

[30] Wians F H, Krech K E, Hauschka P V. (1983). Effects of magnesium and calcium on osteocalcin adsorption to hydroxyapatite. *Magnesium*. **2**, 83-92.

[31] Talmage R V & Mobley H T. (2009). The concentration of free calcium in plasma is set by the extracellular action of noncollagenous proteins and hydroxyapatite. *Gen Comp Endocrinol*. **162**, 245-50.

[32] Hoang Q Q, Sicheri F, Howard A J, Yang D S. (2003). Bone recognition mechanism of porcine osteocalcin from crystal structure. *Nature*. **425**, 977-80.

[33] Dowd T L, Rosen J F, Li L, *et al*. (2003). The three-dimensional structure of bovine calcium ion-bound osteocalcin using 1H NMR spectroscopy. *Biochemistry*. **42**, 7769-79.

[34] Malashkevich V N, Almo S C & Dowd T L. (2013). X-ray crystal structure of bovine 3 Glu-Osteocalcin. *Biochemistry*. **52**, 8387-92.

[35] Chen L, Jacquet R, Lowder E, *et al*. (2015). Refinement of collagen-mineral interaction: A possible role for osteocalcin in apatite crystal nucleation, growth and development. *Bone*. **71**, 7-16.

[36] Bronckers A L, Gay S, Dimuzio M T, *et al*. (1985). Immunolocalization of gamma-carboxyglutamic acid containing proteins in developing rat bones. *Coll Relat Res*. **5**, 273-81.

[37] Bronckers A L, Gay S, Finkelman R D, *et al*. (1987). Immunolocalization of Gla proteins (osteocalcin) in rat tooth germs: comparison between indirect immunofluorescence, peroxidase-antiperoxidase, avidin-biotin-peroxidase complex, and avidin-biotin-gold complex with silver enhancement. *J Histochem Cytochem*. **35**, 825-30.

[38] McKee M D, Glimcher M J & Nanci A. (1992). High-resolution immunolocalization of osteopontin and osteocalcin in bone and cartilage during endochondral ossification in the chicken tibia. *Anat Rec*. **234**, 479-92.

[39] Nudelman F, Pieterse K, George A, *et al*. (2010). The role of collagen in bone apatite formation in the presence of hydroxyapatite nucleation inhibitors. *Nat Mater*. **9**, 1004-9.

[40] Toroian D, Lim J E & Price P A. (2007). The size exclusion characteristics of type I collagen: implications for the role of noncollagenous bone constituents in mineralization. *J Biol Chem*. **282**, 22437-47.

[41] Eastell R & Hannon R A. (2008). Biomarkers of bone health and osteoporosis risk. *Proc Nutr Soc*. **67**, 157-62.

[42] Seibel M J. (2003). Biochemical markers of bone remodeling. *Endocrinol Metab Clin North Am*. **32**, 83-113.

[43] Hannemann A, Friedrich N, Spielhagen C, *et al*. (2013). Reference intervals for serum osteocalcin concentrations in adult men and women from the study of health in Pomerania. *BMC Endocr Disord*. **13**, 11.

[44] Price P A, Williamson M K & Lothringer J W. (1981). Origin of the vitamin K-dependent bone protein found in plasma and its clearance by kidney and bone. *J Biol Chem*. **256**, 12760-6.

[45] Kratz M, Zelnick L R, Trenchevska O, *et al*. (2021). Relationship between chronic kidney disease, glucose homeostasis, and plasma osteocalcin carboxylation and fragmentation. *J Ren Nutr*. **31**, 248-56.

[46] Sim M, Lewis J R, Prince R L, *et al*. (2020). The effects of vitamin K-rich green leafy vegetables on bone metabolism: A 4-week randomized controlled trial in middle-aged and older individuals. *Bone Rep*. **12**, 100274.

[47] del Pino J, Martín-Gómez E, Martín-Rodríguez M, *et al*. (1990). Influence of sex, age, and menopause in serum osteocalcin (BGP) levels. *Klin Wochenschr*. **69**, 1135-8.

[48] Singh S, Kumar D & Lal A K. (2015). Serum osteocalcin as a diagnostic biomarker for primary osteoporosis in women. *J Clin Diagn Res*. **9**, RC04-RC07.

[49] Liu Z, Chen R, Jiang Y, *et al*. (2019). A meta-analysis of serum osteocalcin level in postmenopausal osteoporotic women compared to controls. *BMC Musculoskelet Disord*. **20**, 532.

[50] Riquelme-Gallego B, García-Molina L, Cano-Ibáñez N, *et al*. (2020). Circulating undercarboxylated osteocalcin as estimator of cardiovascular and type 2 diabetes risk in metabolic syndrome patients. *Sci Rep*. **10**, 1840.

[51] Hauschka P V, Lian J B & Gallop P M. (1975). Direct identification of the calcium-binding amino acid, γ-carboxyglutamate, in mineralized tissue. *Proc Natl Acad Sci U S A*. **72**, 3925-9.

[52] Price P A, Otsuka A S, Poser J W, *et al*. (1976). Characterization of a γ-carboxyglutamic acid-containing protein from bone. *Proc Natl Acad Sci U S A*. **73**, 1447-51.

[53] Ducy P, Desbois C, Boyce B, *et al*. (1996). Increased bone formation in osteocalcin-deficient mice. *Nature*. **382**, 448-52.

[54] Chenu C, Colucci S, Grano M, *et al*. (1994). Osteocalcin induces chemotaxis, secretion of matrix proteins, and calcium-mediated intracellular signaling in human osteoclast-like cells. *J Cell Biol*. **127**, 1149-58.

[55] Ivaska K K, Hentunen T A, Vääräniemi J, *et al*. (2004). Release of intact and fragmented osteocalcin molecules from bone matrix during bone resorption *in vitro*. *J Biol Chem*. **279**, 18361-9.

[56] Gundberg C M & Weinstein R S. (1986). Multiple immunoreactive forms of osteocalcin in uremic serum. *J Clin Invest*. **77**, 1762-7.

[57] Gundberg C M, Grant F D, Conlin P R, *et al.* (1991). Acute changes in serum osteocalcin during induced hypocalcemia in humans. *J Clin Endocrinol Metab*. **72**, 438-43.

[58] Boskey A L, Gadaleta S, Gundberg C, *et al.* (1998). Fourier transform infrared microspectroscopic analysis of bones of osteocalcin-deficient mice provides insight into the function of osteocalcin. *Bone*. **23**, 187-96.

[59] Aarden E M, Wassenaar A M, Alblas M J, *et al.* (1996). Immunocytochemical demonstration of extracellular matrix proteins in isolated osteocytes. *Histochem Cell Biol*. **106**, 495-501.

[60] Nishimoto S K, Araki N, Robinson F D, *et al.* (1992). Discovery of bone gamma-carboxyglutamic acid protein in mineralized scales. The abundance and structure of Lepomis macrochirus bone gamma-carboxyglutamic acid protein. *J Biol Chem*. **267**, 11600-5.

[61] Kavukcuoglu N B, Patterson-Buckendahl P & Mann A B. (2009). Effect of osteocalcin deficiency on the nanomechanics and chemistry of mouse bones. *J Mech Behav Biomed Mater*. **2**, 348-54.

[62] Rammelt S, Neumann M, Hanisch U, *et al.* (2005). Osteocalcin enhances bone remodeling around hydroxyapatite/collagen composites. *J Biomed Mater Res A*. **73**, 284-94.

[63] Ritter N M, Farach-Carson M C, Butler W T. (1992). Evidence for the formation of a complex between osteopontin and osteocalcin. *J Bone Miner Res*. **7**, 877-85.

[64] Kaartinen M T, Pirhonen A, Linnala-Kankkunen A, *et al.* (1997). Transglutaminase-catalyzed cross-linking of osteopontin is inhibited by osteocalcin.

[65] Poundarik A A, Diab T, Sroga G E, *et al.* (2012). Dilatational band formation in bone. *Proc Natl Acad Sci U S A*. **109**, 19178-83.

[66] Bailey S, Karsenty G, Gundberg C, *et al.* (2018). Osteocalcin and osteopontin influence bone morphology and mechanical properties. *Ann N Y Acad Sci*. **1409**, 79-84.

[67] Poundarik A A, Boskey A, Gundberg C, *et al.* (2018). Biomolecular regulation, composition and nanoarchitecture of bone mineral. *Sci Rep*. **8**, 1191.

[68] Moriishi T, Ozasa R, Ishimoto T, *et al.* (2020). Osteocalcin is necessary for the alignment of apatite crystallites, but not glucose metabolism, testosterone synthesis, or muscle mass. *PLoS Genet*. **16**, e1008586.

[69] Hung P S, Kuo Y C, Chen H G, *et al.* (2013). Detection of osteogenic differentiation by differential mineralized matrix production in mesenchymal stromal cells by Raman spectroscopy. *PLoS ONE*. **8**, e65438.

[70] Termine J D & Posner A S. (1966). Infrared analysis of rat bone: age dependency of amorphous and crystalline mineral fractions. *Science*. **153**, 1523-5.

[71] Tsao Y-T, Huang Y-J, Wu H-H, *et al.* (2017). Osteocalcin mediates biomineralization during osteogenic maturation in human mesenchymal stromal cells. *Int J Mol Sci*. **18**, 159.

[72] Simon P, Grüner D, Worch H, *et al.* (2018). First evidence of octacalcium phosphate@osteocalcin nanocomplex as skeletal bone component directing collagen triple-helix nanofibril mineralization. *Sci Rep*. **8**, 13696.

[73] Song Q, Jiao K, Tonggu L, *et al.* (2019). Contribution of biomimetic collagen-ligand interaction to intrafibrillar mineralization. *Sci Adv*. **5**, eaav9075.

[74] Lee N K, Sowa H, Hinoi E, *et al.* (2007). Endocrine regulation of energy metabolism by the skeleton. *Cell*. **130**, 456-69.

[75] Oury F, Sumara G, Sumara O, *et al.* (2011). Endocrine regulation of male fertility by the skeleton. *Cell*. **144**, 796-809.

[76] Pi M, Wu Y & Quarles L D. (2011). GPRC6A mediates responses to osteocalcin in beta-cells in vitro and pancreas in vivo. *J Bone Miner Res*. **26**, 1680-3.

[77] Pi M, Chen L, Huang M-Z, *et al.* (2008). GPRC6A null mice exhibit osteopenia, feminization and metabolic syndrome. *PLoS ONE*. **3**, e3858.

[78] Diegel C R, Hann S, Ayturk U M, *et al.* (2020). An osteocalcin-deficient mouse strain without endocrine abnormalities. *PLoS Genet*. **16**, e1008361.

[79] Lambert L J, Challa A K, Niu A, *et al.* (2016). Increased trabecular bone and improved biomechanics in an osteocalcin-null rat model created by CRISPR/Cas9 technology. *Dis Models Mech*. **9**, 1169-79.

[80] Smajilovic S, Clemmensen C, Johansen L D, *et al.* (2013). The L-alpha-amino acid receptor GPRC6A is expressed in the islets of Langerhans but is not involved in L-arginine-induced insulin release. *Amino Acids*. **44**, 383-90.

[81] Wellendorph P, Johansen L D, Jensen A A, *et al.* (2009). No evidence for a bone phenotype in GPRC6A knockout mice under normal physiological conditions. *J Mol Endocrinol*. **42**, 215-23.

[82] Jorgensen C V, Gasparini S J, Tu J, *et al.* (2019). Metabolic and skeletal homeostasis are maintained in full locus GPRC6A knockout mice. *Sci Rep*. **9**, 5995.

[83] Komori T. (2020). Functions of osteocalcin in bone, pancreas, testis, and muscle. *Int J Mol Sci*. **21**, 7513.

[84] Freeman H C, Hugill A, Dear N T, *et al.* (2006). Deletion of nicotinamide nucleotide transhydrogenase: A new quantitative trait locus accounting for glucose intolerance in C57BL/6J mice. *Diabetes*. **55**, 2153-6.

[85] Aston-Mourney K, Wong N, Kebede M, *et al.* (2007). Increased nicotinamide nucleotide transhydrogenase levels predispose to insulin hypersecretion in a mouse strain susceptible to diabetes. *Diabetologia*. **50**, 2476-85.

[86] Toye A A, Lippiat J D, Proks P, *et al.* (2005). A genetic and physiological study of impaired glucose homeostasis control in C57BL/6J mice. *Diabetologia*. **48**, 675-86.

[87] Fontaine D A & Davis D B. (2016). Attention to background strain is essential for metabolic research: C57BL/6 and the International Knockout Mouse Consortium. *Diabetes*. **65**, 25-33.

[88] Nicholson A, Reifsnyder P C, Malcolm R D, *et al.* (2010). Diet-induced obesity in two C57BL/6 substrains with intact or mutant nicotinamide nucleotide transhydrogenase (Nnt) gene. *Obesity*. **18**, 1902-5.

[89] Bilotta F L, Arcidiacono B, Messineo S, *et al.* (2018). Insulin and osteocalcin: further evidence for a mutual cross-talk. *Endocrine*. **59**, 622-32.

[90] Oury F, Khrimian L, Denny C A. (2013). Maternal and offspring pools of osteocalcin influence brain development and functions. *Cell*. **155**, 228-41.

[91] Dudev T, Lim C. (2014). Competition among metal ions for protein binding sites: determinants of metal ion selectivity in proteins. *Chem Rev*. **114**, 538-56.

[92] Mushak P. (1993). New directions in the toxicokinetics of human lead exposure. *Neurotoxicology*. **14**, 29-42.

[93] Yokel R A & McNamara P J. (2001). Aluminum toxicokinetics: an updated minireview. *Pharmacol Toxicol*. **88**, 159-67.

[94] Dowd T L, Rosen J F, Gundberg C M, Gupta R K. (1994). The displacement of calcium from osteocalcin at submicromolar concentrations of free lead. *Biochim Biophys Acta*. **1226**, 131-7.

[95] Rowatt E, Sorensen E S, Triffit J, Viess A, Williams R J. (1997). An examination of the binding of aluminum to protein and mineral components of bone and teeth. *J Inorg Biochem*. **68**, 235-8.

[96] Ellis D E, Terra J, Warschkow O, *et al.* (2006). A theoretical and experimental study of lead substitution in calcium hydroxyapatite. *Phys Chem Chem Phys*. **8**, 967-76.

[97] Monir A U, Gundberg C M, Yagerman S E, *et al.* (2010). The effect of lead on bone mineral properties from female adult C57/BL6 mice. *Bone*. **47**, 888-94.

[98] Degeratu C N, Mabilleau G, Cincu C & Chappard D. (2013). Aluminum inhibits the growth of hydroxyapatite crystals developed on a biomimetic methacrylic polymer. *J Trace Elem Med Biol*. **27**, 346-51.

[99] Ellis H A, McCarthy J H & Herrington J. (1979). Bone aluminum in haemodialysed patients and in rats injected with aluminum chloride: relationship to impaired bone mineralisation. *J Clin Pathol*. **32**, 832-44.

[100] Dermience M, Lognay G, Mathieu F & Goyens P. (2015). Effects of thirty elements on bone metabolism. *J Trace Elem Med Biol*. **32**, 86-106.

[101] Bae Y J & Kim M H. (2008). Manganese supplementation improves mineral density of the spine and femur and serum osteocalcin in rats. *Biol Trace Elem Res*. **124**, 28-34.

[102] Balogh E, Paragh G & Jeney V. (2018). Influence of iron on bone homeostasis. *Pharmaceuticals (Basel)*. **11**, 107.

[103] Guggenbuhl P, Filmon R, Mabilleau G, Baslé M F, Chappard D. (2008). Iron inhibits hydroxyapatite crystal growth in vitro. *Metabolism*. **57**, 903-10.

[104] Tracy R P, Andrianorivo A, Riggs B L & Mann K G. (1990). Comparison of monoclonal and polyclonal antibody-based immunoassays for osteocalcin: a study of sources of variation in assay results. *J Bone Miner Res*. **5**, 451-61.

[105] Ding J & Anderegg R J. (1995). Specific and nonspecific dimer formation in the electrospray ionization mass spectrometry of oligonucleotides. *J Am Soc Mass Spectrom*. **6**, 159-64.

[106] Wang W, Kitova E N & Klassen J S. (2003). Influence of solution and gas phase processes on protein-carbohydrate binding affinities determined by nanoelectrospray Fourier transform ion cyclotron resonance mass spectrometry. *Anal Chem*. **75**, 4945-55.

[107] Nettleton E J, Tito P, Sunde M, *et al.* (2000). Characterization of the oligomeric states of insulin in self-assembly and amyloid fibril formation by mass spectrometry. *Biophys J*. **79**, 1053-65.

[108] Rumble J R. (2020). CRC Handbook of Chemistry and Physics, 101st ed. *CRC Press*: Boca Raton, FL.

[109] Konermann L. (2017). Addressing a common misconception: ammonium acetate as neutral pH "buffer" for native electrospray mass spectrometry. *J Am Soc Mass Spectrom*. **28**, 1827-35.

[110] Nishimoto S K & Price P A. (1985). The vitamin K-dependent bone protein is accumulated within cultured osteosarcoma cells in the presence of the vitamin K antagonist warfarin. *J Biol Chem*. **260**, 2832-6.

[111] Wintersteiner O & Abramson H A. (1933). The isoelectric point of insulin. *J Biol Chem*. **99**, 741-53.

[112] Song W J, Sontz P A, Ambroggio X I, Tezcan F A. (2014). Metals in protein-protein interfaces. *Annu Rev Biophys*. **43**, 409-31.

[113] Jeffrey P D. (1986). Self-association of des-(B26-B30)-insulin. The effect of Ca^{2+} and some other divalent cations. *Biol Chem Hoppe Seyler*. **367**, 363-9.

[114] Duboué-Dijon E, Delcroix P, Martinez-Seara H, *et al.* (2018). Binding of divalent cations to insulin: Capillary electrophoresis and molecular simulations. **122**, 5640-8.

[115] Straub S G & Sharp G W. (2002). Glucose-stimulated signaling pathways in biphasic insulin secretion. *Diabetes Metab Res Rev*. **18**, 451-63.

[116] Henquin J C. (2000). Triggering and amplifying pathways of regulation of insulin secretion by glucose. *Diabetes*. **49**, 1751-60.

[117] Kostov K. (2019). Effects of magnesium deficiency on mechanisms of insulin resistance in type 2 diabetes: Focusing on the processes of insulin secretion and signaling. *Int J Mol Sci*. **20**, 1351.

[118] Bryan M R, Nordham K D, Rose D I R, *et al.* (2020). Manganese acts upon insulin/IGF receptors to phosphorylate AKT and increase glucose uptake in Huntington's disease cells. *Mol Neurobiol*. **57**, 1570-93.

[119] Scudeller L A, Srinivasan S, Rossi A M, *et al.* (2017). Orientation and conformation of osteocalcin adsorbed onto calcium phosphate and silica surfaces. *Biointerphases*. **12**, 02D411.

[120] Coffman F D & Dunn M F. (1988). Insulin-metal ion interactions: the binding of divalent cations to insulin hexamers and tetramers and the assembly of insulin hexamers. *Biochemistry*. **27**, 6179-87.

[121] Dunn M F. (2005). Zinc-ligand interactions modulate assembly and stability of the insulin hexamer – a review. *BioMetals*. **18**, 295-303.

[122] Fukunaka A & Fujitani Y. (2018). Role of zinc homeostasis in the pathogenesis of diabetes and obesity. *Int J Mol Sci*. **19**, 476.

[123] Ahmad A, Millett I S, Doniach S, *et al.* (2003). Partially folded intermediates in insulin fibrillation. *Biochemistry*. **42**, 11404-16.

[124] Smith G D, Swenson D C, Dodson E J, *et al.* (1984). Structural stability in the 4-zinc human insulin hexamer. *Proc Natl Acad Sci USA*. **81**, 7093-97.

[125] Goldman J & Carpenter F H. (1974). Zinc binding, circular dichroism, and equilibrium sedimentation studies on insulin (bovine) and several of its derivatives. *Biochemistry*. **22**, 4566-74.

[126] Nousiainen M, Derrick P J, Kaartinen M T, *et al.* (2002). A mass spectrometric study of metal binding to osteocalcin. *Chem Biol*. **9**, 195-202.

[127] Leitner A, Joachimiak L A, Unverdorben P, *et al.* (2014). Chemical cross-linking/mass spectrometry targeting acidic residues in proteins and protein complexes. *Proc Natl Acad Sci U S A*. **111**, 9455-60.

[128] Sun J, Kitova E N, Wang W & Klassen J S. (2006). Method for distinguishing specific from nonspecific protein-ligand complexes in nanoelectrospray ionization mass spectrometry. *Anal Chem*. **78**, 3010-18.

[129] Daubenfeld T, Bouin A P & van der Rest G. (2006). A deconvolution method for the separation of specific versus nonspecific interactions in noncovalent protein-ligand complexes analyzed by ESI-FT-ICR mass spectrometry. *J Am Soc Mass Spectrom*. **17**, 1239-48.

[130] Shimon L, Sharon M & Horovitz A. (2010). A method for removing effects of nonspecific binding on the distribution of binding stoichiometries: Application to mass spectroscopy data. *Biophys J*. **99**, 1645-49.

[131] Wang W, Kitova E N & Klassen J S. (2005). Nonspecific protein-carbohydrate complexes produced by nanoelectrospray ionization. Factors influencing their formation and stability. *Anal Chem*. **77**, 3060-71.

[132] Von Euw S, Wang Y, Laurent G, *et al.* (2019). Bone mineral: New insights into its chemical composition. *Sci Rep*. **9**, 8456.

[133] Marsh J A & Teichmann S A. (2015). Structure, dynamics, assembly, and evolution of protein complexes. *Annu Rev Biochem*. **84**, 551-75.

[134] Bergendahl L T & Marsh J A. (2017). Function determinants of protein assembly into homomeric complexes. *Sci Rep*. **7**, 4932.

[135] Haley D A, Horwitz J & Stewart P L. (1998). The small heat-shock protein, αB-crystallin, has a variable quaternary structure. *J Mol Biol*. **277**, 27-35.

[136] Aquilina J A, Benesch J L P, Bateman O A, *et al.* (2003). Polydispersity of a mammalian chaperone: Mass spectrometry reveals the population of oligomers in αB-crystallin. *Proc Natl Acad Sci U S A*. **100**, 10611-16.

[137] Stengel F, Baldwin A J, Painter A J, *et al.* (2010). Quaternary dynamics and plasticity underlie small heat shock protein chaperone function. *Proc Natl Acad Sci U S A*. **107**, 2007-12.

[138] Hannon M J & Verbalis J G. (2014). Sodium homeostasis and bone. *Curr Opin Nephrol Hypertens*. **23**, 370-6.

[139] Medeiros D M & Wildman R E C. (1964). Advanced human nutrition, 2nd Ed. Sudbury, MA: Jones & Bartlett Learning c2012.

[140] Heaney R P. (2007). Effects of protein on the calcium economy. *Internat Cong Series*. **1297**, 191-7.

[141] Liu Y L, Hsu J T, Shih T Y, *et al*. (2018). Quantification of volumetric bone mineral density of proximal femurs using a two-compartment model and computed tomography images. *Biomed Res Int*. **doi**: 10.1155/2018/6284269.

[142] Weiler H A, Janzen L, Green K, *et al*. (2000). Percent body fat and bone mass in healthy Canadian females 10 to 19 years of age. *Bone*. **27**, 203-7.

[143] Vagenende V, Yap M G S & Trout B L. (2009). Mechanisms of protein stabilization and prevention of protein aggregation by glycerol. *Biochemistry*. **48**, 11084-96.

[144] Hirai M, Ajito S, Sugiyama M, *et al*. (2018). Direct evidence for the effect of glycerol on protein hydration and thermal structural transition. *Biophys J*. **115**, 313-27.

[145] Bergstrom W H. (1955). The participation of bone in total body sodium metabolism in the rat. *J Clin Invest*. **34**, 997-1004.

[146] Aylett C H, Lowe J & Amos L A. (2011). New insights into the mechanisms of cytomotive actin and tubulin filaments. *Int Rev Cell Mol Biol*. **292**, 1-71.

[147] Wegner A. (1976). Head to tail polymerization of actin. *J Mol Biol*. **108**, 139-50.

[148] Mitchison T & Kirschner M. (1984). Dynamic instability of microtubule growth. *Nature*. **312**, 237-42.

[149] Gunning B E. (1965). The fine structure of chloroplast stroma following aldehyde osmium-tetroxide fixation. *J Cell Biol*. **24**, 79-93.

[150] Kleinschmidt A K, Moss J & Lane D M. (1969). Acetyl coenzyme A carboxylase: filamentous nature of the animal enzymes. *Science*. **166**, 1276-8.

[151] Olsen B R, Svenneby G, Kvamme E, *et al*. (1970). Formation and ultrastructure of enzymatically active polymers of pig renal glutaminase. *J Mol Biol*. **52**, 239-45.

[152] Eisenberg H & Reisler E. (1971). Angular dependence of scattered light, rotary frictional coefficients, and distribution of sizes of associated oligomers in solutions of bovine liver glutamate dehydrogenase. *Biopolymers*. **10**, 2363-76.

[153] Kemp R G. (1971). Rabbit liver phosphofructokinase. Comparison of some properties with those of muscle phosphofructokinase. *J Biol Chem*. **246**, 245-52.

[154] Huang C Y & Frieden C. (1972). The mechanism of ligand-induced structural changes in glutamate dehydrogenase. Studies of the rate of depolymerization and isomerization effected by coenzymes and guanine nucleotides. *J Biol Chem*. **247**, 3638-46.

[155] Josephs R & Borisy G. (1972). Self-assembly of glutamic dehydrogenase into ordered superstructures: multichain tubes formed by association of single molecules. *J Mol Biol*. **65**, 127-55.

[156] Miller R E, Shelton E & Stadtman E R. (1974). Zinc-induced paracrystalline aggregation of glutamine synthetase. *Arch Biochem Biophys*. **163**, 155-71.

[157] Frey T G, Eisenberg D & Eiserling F A. (1975). Glutamine synthetase forms three- and seven-stranded helical cables. *Proc Natl Acad Sci U S A*. **72**, 3402-6.

[158] Harper D B. (1977). Fungal degradation of aromatic nitriles. Enzymology of C-N cleavage by Fusarium solani. *Biochem J*. **167**, 685-92.

[159] Harper D B. (1977). Microbial metabolism of aromatic nitriles. Enzymology of C-N cleavage by Nocardia sp. (Rhodochrous group) N. C. I. B. 11216. *Biochem J*. **165**, 309-19.

[160] Trujillo J L & Deal W C Jr. (1977). Pig liver phosphofructokinase: asymmetry properties, proof of rapid association-dissociation equilibria, and effect of temperature and protein concentration on the equilibria. *Biochemistry*. **16**, 3098-104.

[161] Meredith M J & Lane M D. (1978). Acetyl-CoA carboxylase. Evidence for polymeric filament to promoter transition in the intact avian liver cell. *J Biol Chem*. **253**, 3381-3.

[162] Zeiri L & Reisler E. (1978). Uncoupling of the catalytic activity and the polymerization of beef liver glutamate dehydrogenase. *J Mol Biol*. **124**, 291-5.

[163] Reinhart G D & Lardy H A. (1980). Rat liver phosphofructokinase: kinetic and physiological ramifications of the aggregation behavior. *Biochemistry*. **19**, 1491-5.

[164] Beaty N B & Lane M D. (1983). Kinetics of activation of acetyl-CoA carboxylase by citrate. Relationship to the rate of polymerization of the enzyme. *J Biol Chem*. **258**, 13043-50.

[165] Park C K & Horton N C. (2019). Structures, functions, and mechanisms of filament forming enzymes: a renaissance of enzyme filamentation. *Biophys Rev*. **11**, 927-94.

[166] Narayanaswamy R, Levy M, Tsechansky M, *et al*. (2009). Widespread reorganization of metabolic enzymes into reversible assemblies upon nutrient starvation. *Proc Natl Acad Sci U S A*. **106**, 10147-52.

[167] Werner J N, Chen E Y, Guberman J M, *et al*. (2009). Quantitative genome-scale analysis of protein localization in an asymmetric bacterium. *Proc Natl Acad Sci U S A*. **106**, 7858-63.

[168] Liu J L. (2010). Intracellular compartmentation of CTP synthase in Drosophila. *J Genet Genomics*. **37**, 281-96.

[169] Noree C, Sato B K, Broyer R M, *et al*. (2010). Identification of novel filament-forming proteins in Saccharomyces cerevisiae and Drosophila melanogaster. *J Cell Biol*. **190**, 541-51.

[170] Ibstedt S, Sideri T C, Grant C M, *et al*. (2014). Global analysis of protein aggregation in yeast during physiological conditions and arsenite stress. *Biol Open*. **3**, 913-23.

[171] Lowe N, Rees J S, Roote J, *et al.* (2014). Analysis of the expression patterns, subcellular localisations and interaction partners of Drosophila proteins using a pigP protein trap library. *Development*. **141**, 3994-4005.

[172] Suresh H G, da Silveira Dos Santos A X, Kukulski W, *et al.* (2015). Prolonged starvation drives reversible sequestration of lipid biosynthetic enzymes and organelle reorganization in Saccharomyces cerevisiae. *Mol Biol Cell*. **26**, 1601-15.

[173] Shen Q-J, Kassim H, Huang Y, *et al.* (2016). Filamentation of metabolic enzymes in *Saccharomyces cerevisiae*. *J Genet Genomics*. **43**, 393-404.

[174] Stoddard P R, Lynch E M, Farrell D P, *et al.* (2020). Independent evolution of polymerization in the actin ATPase clan regulates hexokinase activity. *Science*. **367**, 1039-42.

[175] Glover C V. (1986). A filamentous form of Drosophila casein kinase II. *J Biol Chem*. **261**, 14349-54.

[176] Valero E, De Bonis S, Filhol O, *et al.* (1995). Quaternary structure of casein kinase 2. Characterization of multiple oligomeric states and relation with its catalytic activity. *J Biol Chem*. **270**, 8345-52.

[177] Seetoh W G, Chan D S, Matak-Vinkovic D, *et al.* (2016). Mass spectrometry reveals protein kinase CK2 high-order oligomerization via the circular and linear assembly. *ACS Chem Biol*. **11**, 1511-7.

[178] Hunkeler M, Hagmann A, Stuttfeld E, *et al.* (2018). Structural basis for regulation of human acetyl-CoA carboxylase. *Nature*. **558**, 470-4.

[179] Schuchmann K, Vonck J & Muller V. (2016). A bacterial hydrogen-dependent CO2 reductase forms filamentous structures. *FEBS J*. **283**, 1311-22.

[180] Radding C M. (1981). Recombination activities of E. coli recA protein. *Cell*. **25**, 3-4.

[181] Chen Y, Zhu J, Lum P Y, *et al.* (2008). Variations in DNA elucidate molecular networks that cause disease. *Nature*. **452**, 429-35.

[182] Viguet-Carrin S, Garnero P & Delmas P D. (2006). The role of collagen in bone strength. *Osteoporos Int*. **17**, 319-36.

[183] Moreno E C, Gregory T M & Brown W E. (1968). Preparation and solubility of hydroxyapatite. *J Res Natl Bur Stand A Phys Chem*. **72A**, 773-82.

[184] Marenzana M, Shipley A M, Squitiero P, *et al.* (2005). Bone as an ion exchange organ: evidence for instantaneous cell-dependent calcium efflux from bone not due to resorption. *Bone*. **37**, 545-54.

[185] Neuman W F, Neuman M W, Diamond A G, *et al.* (1982). Blood:bone disequilibrium. *Calcif Tissue Int*. **34**, 149-57.

[186] Dedic C, Hung T S, Shipley A M, *et al.* (2018). Calcium fluxes at the bone/plasma interface: acute effects of parathyroid hormone (PTH) and targeted deletion of PTH/PTH-related peptide (PTHrP) receptor in the osteocytes. *Bone*. **116**, 135-43.

[187] Sorenson J R J, Campbell I R, Tepper L B, *et al.* (1974). Aluminum in the environment and human health. *Environ Health Perspect*. **8**, 3-95.

[188] Kawahara M & Kato-Negishi M. (2011). Link between aluminum and the pathogenesis of alzheimer's disease: The integration of the aluminum and amyloid cascade hypotheses. *Int J Alzheimers Dis*. **2011**, 276393.

[189] Huat T J, Camats-Perna J, Newcombe E A, *et al.* (2019). Metal toxicity links to alzheimer's disease and neuroinflammation. *J Mol Biol*. **431**, 1843-68.

[190] Russo L S, Beale G, Sandroni S, *et al.* (1992). Aluminum intoxication in undialysed adults with chronic renal failure. *J Neurol Neurosurg Psychiatry*. **55**, 697-700.

[191] Cannata Andía J B. (2000). Adynamic bone and chronic renal failure: an overview. *Am J Med Sci*. **320**, 81-4.

[192] Li X, Hu C, Zhu Y, *et al.* (2011). Effects of aluminum exposure on bone mineral density, mineral, and trace elements in rats. *Biol Trace Elem Res*. **143**, 378-85.

[193] Li X, Han Y, Guan Y, *et al.* (2012). Aluminum induces osteoblast apoptosis through the oxidative stress-mediated JNK signaling pathway. *Biol Trace Elem Res*. **150**, 502-8.

[194] Nelson R W, Krone J R, Bieber A L, *et al.* (1995). Mass spectrometric immunoassay. *Anal Chem*. **67**, 1153-8.

[195] Hedengren J D, Asgharzadeh Shishavan R, Powell K M, *et al.* (2014). Nonlinear modeling, estimation and predictive control in APMonitor. *Comput Chem Eng*. **70**, 133-48.

[196] Van Berkel G J & Kertesz V. (2007). Using the electrochemistry of the electrospray ion source. *Anal Chem*. **79**, 5510-20.

[197] Izawa C, Wagner S, Martin M, *et al.* (2013). SIMS analysis on austenitic stainless steel: The influence of type of oxide surface layer on hydrogen embrittlement. *J Alloy Compd*. **580**, S13-7.

[198] Skoog D A, West D M, Holler F J, *et al.* (2014). Fundamentals of analytical chemistry, 9th ed. Cengage Learning, Belmont, CA.

[199] Martell A E, Hancock R D, Smith R M, *et al.* (1996). Coordination of Al(III) in the environment and in biological systems. *Coord Chem Rev*. **149**, 311-28.

[200] Martin R B. (1986). Citrate binding of Al^{3+} and Fe^{3+}. *J Inorg Biochem*. **28**, 181-7.

[201] Jarmoskaite I, AlSadhan I, Vaidyanathan P P, *et al.* (2020). How to measure and evaluate binding affinities. *eLife*. **9**, e57264.

[202] Kennedy J A & Powell H K J. (1986) Colorimetric determination of aluminum(III) with chrome azurol S and the reactivity of hydrolysed aluminum species. *Anal Chim Acta*. **184**, 329-33.

[203] Rowatt E & Williams R J P. (1994). The affinity of phosvitin for aluminum measured with the dye, chrome azurol S. *J Inorg Biochem*. **55**, 249-54.

[204] Hem J D & Roberson C E. (1967). Form and stability of aluminum hydroxide complexes in dilute solution. *U S G P O*. Washington, DC.

[205] Wehr J B, Blamey F P C, Hanna J V, *et al.* (2010). Hydrolysis and speciation of Al bound to pectin and plant cell wall material and its reaction with the dye chrome azurol S. *J Agric Food Chem*. **58**, 5553-60.

[206] Glowacki J, Rey C, Cox K, *et al.* (1989). Effects of bone matrix components on osteoclast differentiation. *Connect Tissue Res*. **20**, 121-9.

[207] Lian J B, Tassinari M & Glowacki J. (1984). Resorption of implanted bone prepared from normal and warfarin-treated rats. *J Clin Invest*. **73**, 1223-6.

[208] Mundy G R & Poser J W. (1983). Chemotactic activity of the gamma-carboxyglutamic acid containing protein in bone. *Calcified Tissue Int*. **35**, 164-8.

[209] Malone J D, Teitelbaum S L, Griffin G L, *et al.* (1982). Recruitment of osteoclast precursors by purified bone matrix constituents. *J Cell Biol*. **92**, 227-30.

APPENDIX A

PERMISSION TO REPRINT FIGURE 1.1B

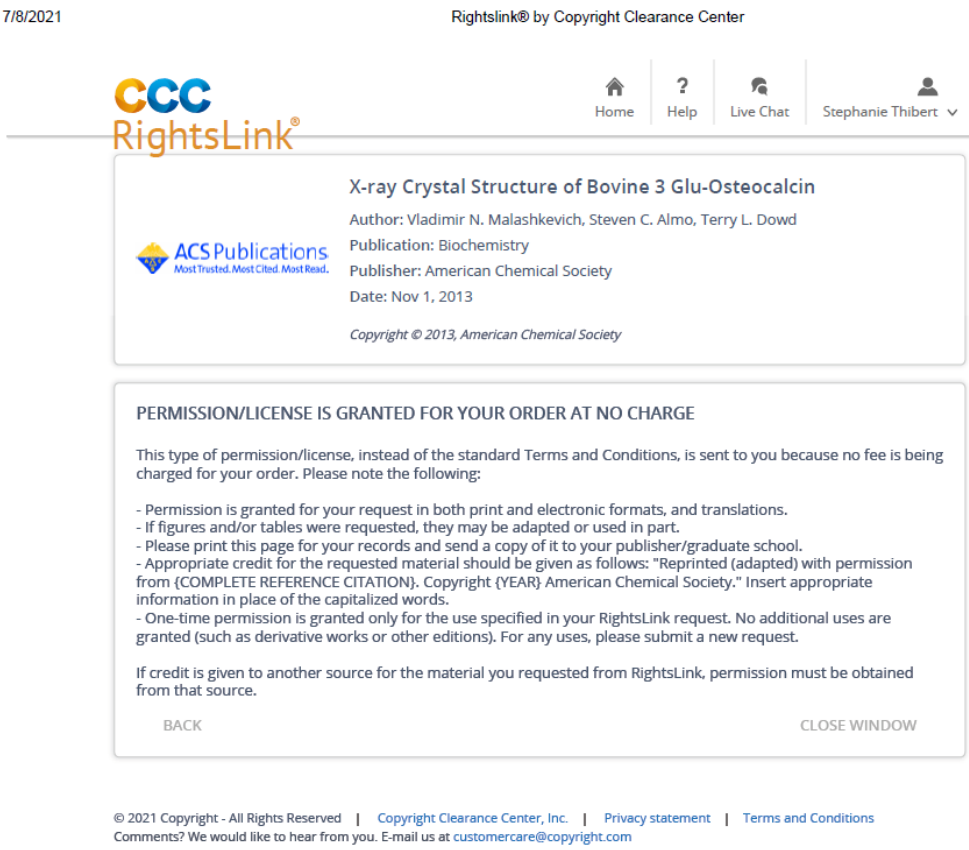


Figure 1b and c from Malashkevich V N, Almo S C & Dowd T L. (2013). X-ray crystal structure of bovine 3 Glu-Osteocalcin. *Biochemistry*. **52**, 8387-92 (reference [34]).

APPENDIX B

MATLAB SCRIPT FOR PLOTTING EXPERIMENTAL DATA WITH AN EQUATION

Below is the Matlab script that was used to generate **Figure 2.12f**.

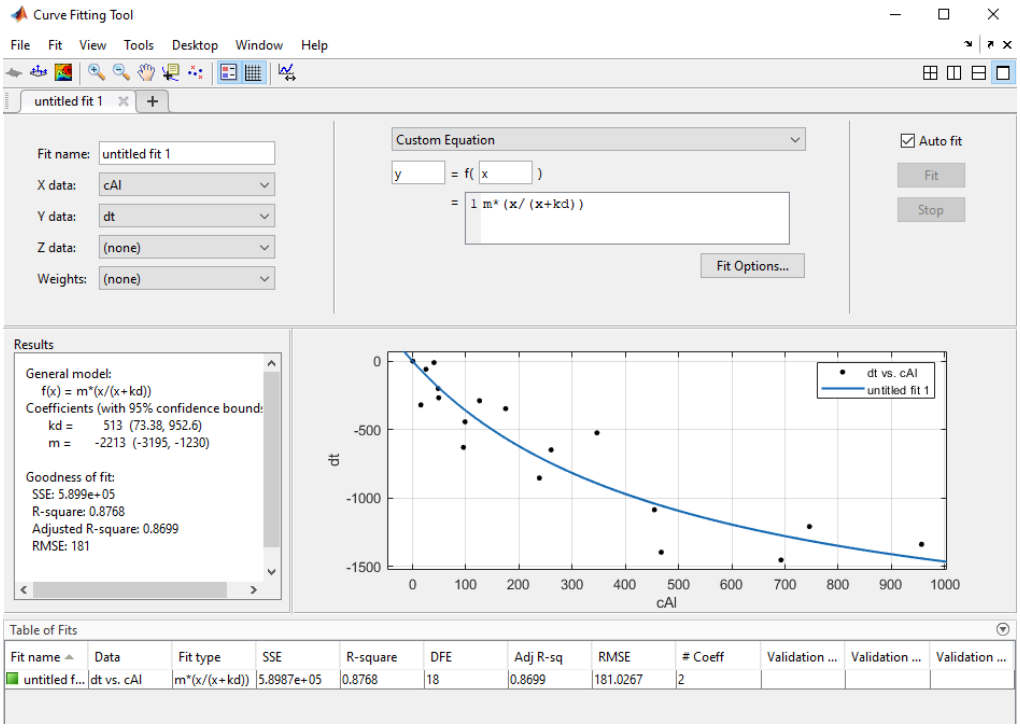
```
%load your data file
load Insdimer.csv %this data file includes free insulin concentrations
in the first column, and the fraction of total insulin that is Ins-Ins
dimer in the second column

%define each column of data present in the data file
Ins = Insdimer(:,1); %free insulin concentrations
FB = Insdimer(:,2); %fraction of total insulin as Ins-Ins dimer

%define the Kd value, as determined when fitting the data using cftool
Kd = 100; %From reference [107]
%generate a set of x values (free insulin concentrations) within the
range of interest
x = linspace(0,500);
%define the Kd equation
y = @(x) ((2*x^2)/((2*x^2)+(Kd*x)));

%generate a figure
figure
%fplot will plot the Kd equation over the specified x range
fplot(y,[0,500])
hold on
%plot your data points as a scatterplot overlayed on the Kd equation
plot(Ins,FB,'ko')
%label your axes
xlabel('Insulin monomer (\muM)')
ylabel('Fraction Bound')
%tada! :)
```

APPENDIX C

MATLAB SCRIPT USED FOR NONLINEAR REGRESSION CURVE FITTING USING CFTOOL TO CALCULATE THE DELTA THETA MAXIMUM FOR CD DATA

```
%load your data file
load dtmax_estimate.csv
%define each column of data present in the data file
cAl = dtmax_estimate(:,1);
dt = dtmax_estimate(:,2);

cftool
```

This will result in the data being opened in Matlab's curve fitting tool, cftool:



Select X data: cAl
Select Y data: dt
Select Custom Equation and type 'm*(x/(Kd+x))' (equation described by Dowd, *et al* [29])

APPENDIX D

MATLAB SCRIPT USED TO SIMULATE DATA DURING PROTOCOL OPTIMIZATION

Below is the Matlab script used to generate single-site binding simulated data in **Figure 5.8**.

```
%Aluminum, Citrate and Osteocalcin Equilibrium simulation: Equilibrium
constants from
%Martell 1996 [199] (all but KAlCit23m) and Martin 1986 [200] for
KAlCit23m).

%This simulation will run a user-designated series of total aluminum
%concentrations and output a plot of Fraction of Oc Bound vs.
Concentration
%of Al that is not bound to Oc. The main parameters adjusted by users
are
%KdAlOc (which is the equilibrium dissociation constant for Al binding
to
%Oc), OrigcAl (which is the starting concentration of total Al),
%cCit (which is the total concentration of citric acid), MaxxcOc (which
is
%the maximum level of total aluminum that will be simulated based on
this
%designated multiple of the total osteocalcin concentration, cOc), and
%increments (which is the number of data points that are calculated and
%show in the final plot).

%IMPORTANT NOTE: For Al to stay soluble, the final "AlOH3
%Solubility Product" value must be < 3e-34.

function BindingSim =
AlCitrateOc2Arrays_Activities(FractionAlBound,ConcAlUnbound)
syms Cit HCit H2Cit H3Cit Al AlCit AlCitHp AlCitHm AlCit23m AlOH12p
AlOH21p AlOH3 AlOH41m Oc AlOc real
assumeAlso(Cit, 'positive')
assumeAlso(HCit, 'positive')
assumeAlso(H2Cit, 'positive')
assumeAlso(H3Cit, 'positive')
assumeAlso(Al, 'positive')
assumeAlso(AlCit, 'positive')
assumeAlso(AlCitHp, 'positive')
assumeAlso(AlCitHm, 'positive')
assumeAlso(AlCit23m, 'positive')
assumeAlso(AlOH12p, 'positive')
assumeAlso(AlOH21p, 'positive')
assumeAlso(AlOH3, 'positive')
assumeAlso(AlOH41m, 'positive')
assumeAlso(Oc, 'positive')
assumeAlso(AlOc, 'positive')
KCit3 = 5.012e5; %All equilibrium constant values from the literature
are for ionic strength u = 0.1M and 25 C
KCit2 = 2.239e4;
KCit1 = 7.943e2;
KAlCit = 9.55e7;
KAlCitHp = 8.71e2;
KAlCitHm = 2.042e3;
KAlCit23m = 1e5;
```

```
KAlOH12p = 2.951e8;
KAlOH21p = 2.138e8; %Note that Q values in Martell 1996 [199] for
ligands 2 and 3 are Beta values. Individual K values and their
equations are used here.
KAlOH3 = 7.943e7;
KAlOH41m = 6.309e6;
KdAlOc = 1e-9; %User specified value for Kd of Al binding to Oc.

OrigcAl = 0.8e-6; %Initial total Al concentration.
cAl = OrigcAl;
cCit = 5e-3; %Total citrate concentration.
cOc = 50e-6; %Total osteocalcin concentration.

%Define the pH for the system
pH = 4.5;
H = 10^-pH; %H+ concentration at specified pH.
OH = 10^-14/H; %OH- concentration at specified pH.

MaxxcOc = 20.0; %This value helps to define the range of cAl values
used for the simulation.
increments = 10; %How many cAl titrations you want to test.
FractionAlBound = zeros(increments,1);
ConcAlUnbound = zeros(increments,1);
n = 1;

while cAl <= MaxxcOc*cOc
    eq1 = (HCit/(Cit*H)) - KCit3;
    eq2 = (H2Cit/(HCit*H)) - KCit2;
    eq3 = (H3Cit/(H2Cit*H)) - KCit1;
    eq4 = (AlCit/(Al*Cit)) - KAlCit;
    eq5 = (AlCitHp/(AlCit*H)) - KAlCitHp;
    eq6 = (AlCit/(AlCitHm*H)) - KAlCitHm;
    eq7 = (AlCit23m/(AlCit*Cit)) - KAlCit23m;
    eq8 = (AlOH12p/(Al*OH)) - KAlOH12p;
    eq9 = (AlOH21p/(AlOH12p*OH)) - KAlOH21p;
    eq10 = (AlOH3/(AlOH21p*OH)) - KAlOH3;
    eq11 = (AlOH41m/(AlOH3*OH)) - KAlOH41m;
    eq12 = Cit + HCit + H2Cit + H3Cit + AlCit + AlCitHp + AlCitHm +
2*AlCit23m - cCit;
    eq13 = Al + AlCit + AlCitHp + AlCitHm + AlOc + AlCit23m + AlOH12p +
AlOH21p + AlOH3 + AlOH41m - cAl;
    eq14 = Oc + AlOc - cOc;
    eq15 = ((Al*Oc)/AlOc) - KdAlOc;
    sol =
vpasolve(eq1,eq2,eq3,eq4,eq5,eq6,eq7,eq8,eq9,eq10,eq11,eq12,eq13,eq14,e
q15);
    ConcAlUnbound(n) = double(sol.Al) + double(sol.AlCit) +
double(sol.AlCitHp) + double(sol.AlCitHm) + double(sol.AlCit23m) +
double(sol.AlOH12p) + double(sol.AlOH21p) + double(sol.AlOH3) +
double(sol.AlOH41m);
    %ConcAlUnbound(n) = double(sol.Al);
    FractionAlBound(n) = double(sol.AlOc)/cOc;
    cAl = cAl + ((MaxxcOc*cOc - OrigcAl))/increments;
    n = n + 1;
```

```
    disp(['AlOH3 Solubility Product = '
num2str(double(sol.Al)*double(OH^3))]);
    disp(['Alfree = ' num2str(double(sol.Al))]);
    disp(['Alicp = ' num2str(double(sol.Al) + double(sol.AlCit) +
double(sol.AlCitHp) + double(sol.AlCitHm) + double(sol.AlCit23m) +
double(sol.AlOH12p) + double(sol.AlOH21p) + double(sol.AlOH3) +
double(sol.AlOH41m))]);
    disp(['Fraction Oc Bound = ' num2str(double(sol.AlOc)/cOc)]);
    disp(['cAl = ' num2str(double(sol.AlOc) + double(sol.Al) +
double(sol.AlCit) + double(sol.AlCitHp) + double(sol.AlCitHm) +
double(sol.AlCit23m) + double(sol.AlOH12p) + double(sol.AlOH21p) +
double(sol.AlOH3) + double(sol.AlOH41m))]);
end

format long g
ConcAlUnbound
format long g
FractionAlBound

figure
plot(10^9*ConcAlUnbound,FractionAlBound,'ko')
xlabel('Al_{ICP} (nM)')
ylabel('Fraction Gla-Oc Bound')
```

OUTPUT

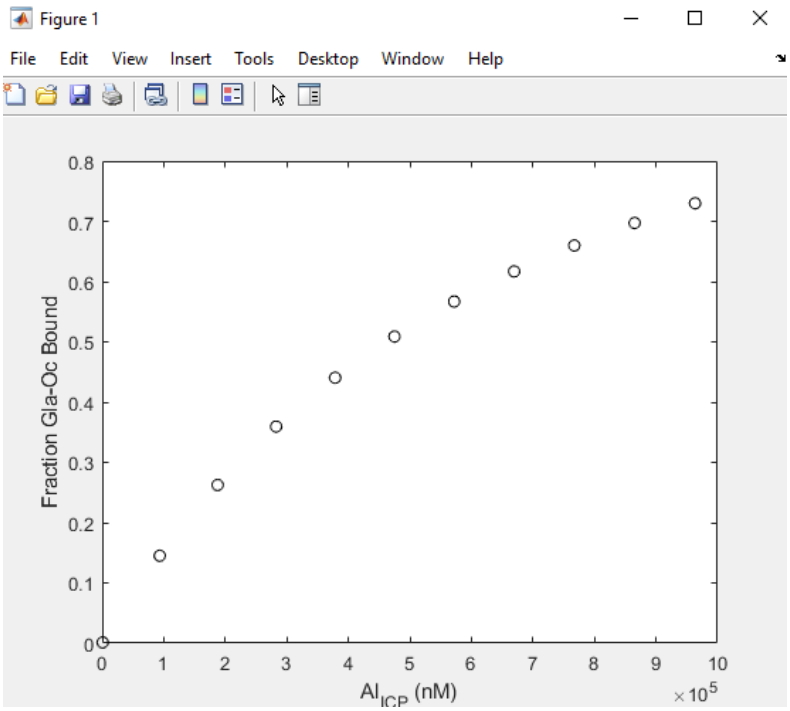
```
>> AlCitrateOc2Arrays_Activities
Warning: Solution '[Al = 0.000000000001280318805085261716281844838863,
AlCit =
0.000000021557301145262946361843708288161, AlCit23m =
0.000000380073229717504592793863113982237, AlCitHm =
0.000000333840214618938441749454669092 6,
AlCitHp = 0.000000000593762196597380845937143790453 07, AlOH12p =
0.000000000000119477832114382737614263826 13723, AlOH21p =
0.000000000000000807783605715852478978653625693 25, AlOH3 =
0.000000000000000202898855499595055785050254947 66, AlOH41m =
0.000000000000000000404799646618882030900926464983 11, AlOc =
0.000000063934084243919690098198201844268, Cit =
0.000176308354722327013607330912758 02,
H2Cit = 0.001978509083991130647226642439944 1, H3Cit =
0.000049696134694588366786064352039408, HCit =
0.002794370288854558164425417333509 3, Oc =
0.000049936065915756080309901801798156]' seems to be affected by some
numeric instability.
Inserting this solution into equation 'AlCit - 95500000*Al*Cit = 0'
produces the residue
'8.665973103396807355756217638918 2e-37'.
> In symengine
In mupadengine/evalin_internal
In mupadengine/feval_internal
In sym/vpasolve (line 172)
In AlCitrateOc2Arrays_Activities (line 79)
```

```
AlOH3 Solubility Product = 4.0487e-41
Alfree = 1.2803e-12
Alicp = 7.3607e-07
Fraction Oc Bound = 0.0012787
cAl = 8e-07
AlOH3 Solubility Product = 5.3667e-39
Alfree = 1.6971e-10
Alicp = 9.3466e-05
Fraction Oc Bound = 0.14509
cAl = 0.00010072
AlOH3 Solubility Product = 1.1254e-38
Alfree = 3.5587e-10
Alicp = 0.00018752
Fraction Oc Bound = 0.26247
cAl = 0.00020064
AlOH3 Solubility Product = 1.7746e-38
Alfree = 5.6117e-10
Alicp = 0.00028259
Fraction Oc Bound = 0.35946
cAl = 0.00030056
AlOH3 Solubility Product = 2.4895e-38
Alfree = 7.8725e-10
Alicp = 0.00037846
Fraction Oc Bound = 0.44048
cAl = 0.00040048
AlOH3 Solubility Product = 3.276e-38
Alfree = 1.036e-09
Alicp = 0.00047496
Fraction Oc Bound = 0.50883
cAl = 0.0005004
AlOH3 Solubility Product = 4.1407e-38
Alfree = 1.3094e-09
Alicp = 0.00057197
Fraction Oc Bound = 0.56699
cAl = 0.00060032
Warning: Solution '[Al = 0.000000001609974030387382593646834271877l,
AlCit =
0.000021806735269972297380913881849326, AlCit23m =
0.000309285144359109401711308458918864,
AlCitHm = 0.000337702996010966033437144035442274, AlCitHp =
0.000000600632470051176064876312179703 5, AlOH12p =
0.000000000150240866686582827608838903801 94, AlOH21p =
0.000000000010157709331533226879726167563 34, AlOH3 =
0.000000000000255141053034768505291178621643 82, AlOH41m =
0.000000000000000509027061055304263487104633 82027, AlOc =
0.000030842721261644582848542548333 72, Cit =
0.000141830099980620486633155416474 85, H2Cit =
0.001591598660409325858346412212365, H3Cit =
0.000039977729719527961190098969902361, HCit =
0.002247912857421317383524782253948 7, Oc =
0.000019157278738355417151457451666 28]' seems to
be affected by some numeric instability. Inserting this solution into
equation 'AlCit -
```

```
95500000*Al*Cit = 0' produces the residue
'3.838521765351824882189113268988 5e-33'.
> In symengine
In mupadengine/evalin_internal
In mupadengine/feval_internal
In sym/vpasolve (line 172)
In AlCitrateOc2Arrays_Activities (line 79)
AlOH3 Solubility Product = 5.0912e-38
Alfree = 1.61e-09
Alicp = 0.0006694
Fraction Oc Bound = 0.61685
cAl = 0.00070024
Warning: Solution '[Al = 0.000000001940285254833527647026597838 3565,
AlCit =
0.000025389499087655917850052535328329, AlCit23m =
0.000347887889497309906115354682702 06,
AlCitHm = 0.000393186316198627584008255688853 58, AlCitHp =
0.000000699314104637187533168205689409 83, AlOH12p =
0.000000000181065118320725298284162299500 67, AlOH21p =
0.000000000012241721460635373773040971094 48, AlOH3 =
0.000000000000307487210204841135675755696339 91, AlOH41m =
0.000000000000000613461883381564989538265679 07571, AlOc =
0.000032994847211574073006953376469231, Cit =
0.000137020383228611624242118764615 19, H2Cit =
0.001537624653900893663006989940002 4, H3Cit =
0.000038622012164752428061261048999423, HCit =
0.002171682042320201783067444447107 6, Oc =
0.000017005152788425926993046623530769]' seems to
be affected by some numeric instability. Inserting this solution into
equation 'AlCit -
95500000*Al*Cit = 0' produces the residue
'0.000000000000000000000000000001789804439132986785752731431774 2'.
> In symengine
In mupadengine/evalin_internal
In mupadengine/feval_internal
In sym/vpasolve (line 172)
In AlCitrateOc2Arrays_Activities (line 79)
AlOH3 Solubility Product = 6.1357e-38
Alfree = 1.9403e-09
Alicp = 0.00076717
Fraction Oc Bound = 0.6599
cAl = 0.00080016
AlOH3 Solubility Product = 7.2837e-38
Alfree = 2.3033e-09
Alicp = 0.00086522
Fraction Oc Bound = 0.69727
cAl = 0.00090008
Warning: Solution '[Al = 0.000000002702315543402415933203830321 7707,
AlCit =
0.000032918504511955687244724604602 46, AlCit23m =
0.000419895057943734364528771795791 17,
AlCitHm = 0.000509781838512778951851361146082 78, AlCitHp =
0.000000906688801905735527250718948716 18, AlOH12p =
0.000000000252176880892739660526878245095 39, AlOH21p =
```

```
0.00000000001704955191442516837166947521530 2, AlOH3 =
0.00000000000042825015830225436565459966410 18, AlOH41m =
0.00000000000000085439374371226474328971498640025, AlOc =
0.00003649493825854449922116299777895 1, Cit =
0.00012755593371236912647759586703127, H2Cit =
0.00143141585073695648380686266503 03, H3Cit =
0.00003595426247864095081614045801 7949, HCit =
0.00202167680535792433521852094468 42, Oc =
0.00001350506174145550077883700222 1049]' seems to
be affected by some numeric instability. Inserting this solution into
equation 'AlCit -
95500000*Al*Cit = 0' produces the residue
'8.381252504325067059857760306464 2e-33'.
> In symengine
In mupadengine/evalin_internal
In mupadengine/feval_internal
In sym/vpasolve (line 172)
In AlCitrateOc2Arrays_Activities (line 79)
AlOH3 Solubility Product = 8.5455e-38
Alfree = 2.7023e-09
Alicp = 0.00096351
Fraction Oc Bound = 0.7299
cAl = 0.001

ConcAlUnbound =

       7.3606591575608e-07
      9.34656146796667e-05
      0.000187516658054692
      0.000282587235699669
      0.000378456014162454
      0.000474958469163085
      0.000571970525348299
      0.000669397278738355
      0.000767165152788426
      0.000865216378774048
      0.000963505061741456


FractionAlBound =

       0.00127868168487839
         0.145087706406666
          0.26246683890617
         0.359455286006616
         0.440479716750929
         0.508830616738306
         0.566989493034015
         0.616854425232892
         0.659896944231481
         0.697272424519033
          0.72989876517089
```

And generates the following figure:



APPENDIX E

MATLAB SCRIPT USED TO CALCULATE FREE AL CONCENTRATIONS FROM EXPERIMENTALLY DETERMINED PARAMETERS

Below is the Matlab script used to solve for free Al and KdAlOc for each experimental set of data.

```
%Aluminum and Citrate Equilibrium simulation: Equilibrium constants
from
%Martell 1996 [199] (all but KAlCit23m) and Martin 1986 [200] for
KAlCit23m)
%Aluminum and Citrate Equilibrium simulation:
syms Cit HCit H2Cit H3Cit cAl Al AlCit AlCitHp AlCitHm AlCit23m AlOH12p
AlOH21p AlOH3 AlOH41m Oc KdAlOc Al2OH24p Al3OH45p AlOHCitHm2m
Al3OH4Cit34m Al3OH4CitHm37m real
assumeAlso(Cit, 'positive')
assumeAlso(HCit, 'positive')
assumeAlso(H2Cit, 'positive')
assumeAlso(H3Cit, 'positive')
assumeAlso(Al, 'positive')
assumeAlso(AlCit, 'positive')
assumeAlso(AlCitHp, 'positive')
assumeAlso(AlCitHm, 'positive')
assumeAlso(AlCit23m, 'positive')
assumeAlso(AlOH12p, 'positive')
assumeAlso(AlOH21p, 'positive')
assumeAlso(AlOH3, 'positive')
assumeAlso(AlOH41m, 'positive')
assumeAlso(Oc, 'positive')
assumeAlso(KdAlOc, 'positive')
assumeAlso(Al2OH24p, 'positive')
assumeAlso(Al3OH45p, 'positive')
assumeAlso(AlOHCitHm2m, 'positive')
assumeAlso(Al3OH4Cit34m, 'positive')
assumeAlso(Al3OH4CitHm37m, 'positive')
assumeAlso(cAl, 'positive')
KCit3 = 5.012e5; %All equilibrium constant values from the literature
are for ionic strength u = 0.1M and 25 C
KCit2 = 2.239e4;
KCit1 = 7.943e2;
KAlCit = 9.55e7;
KAlCitHp = 8.71e2;
KAlCitHm = 2.042e3;
KAlCit23m = 1e5;
KAlOH12p = 2.951e8;
KAlOH21p = 2.138e8; %Note that Q values in Martell 1996 [199] for
ligands 2 and 3 are Beta values. Individual K values and their
equations are used here.
KAlOH3 = 7.943e7;
KAlOH41m = 6.309e6;
KAl2OH24p = 1e20;
KAl3OH45p = 1e42;
KAlOHCitHm2m = 1.698e6;
KAl3OH4Cit34m = 2.692e14;
KAl3OH4CitHm37m = 1.23e-11;
cAl = 346.455e-6; %Total Al concentrations measured by ICP-MS.
cCit = 5e-3; %Total citrate concentration in the system.
cOc = 50e-6; %Total osteocalcin in the system.
```

```
AlOc = 11.812e-6; %Concentration of Al bound Oc, determined by
multiplying the total Oc concentration by the fraction of Oc bound,
which is determined from the change in delta theta at 222 nm in CD
data, divided by the maximum delta theta value at 222 nm.

pH = 4.5;
H = 10^-pH;
OH = 10^-14/H;

    eq1 = ((HCit)/(Cit*H)) - KCit3;
    eq2 = ((H2Cit)/(HCit*H)) - KCit2;
    eq3 = (H3Cit/(H2Cit*H)) - KCit1;
    eq4 = (AlCit/(Al*Cit)) - KAlCit;
    eq5 = ((AlCitHp)/(AlCit*H)) - KAlCitHp;
    eq6 = (AlCit/(AlCitHm*H)) - KAlCitHm;
    eq7 = ((AlCit23m)/(AlCit*Cit)) - KAlCit23m;
    eq8 = ((AlOH12p)/(Al*OH)) - KAlOH12p;
    eq9 = ((AlOH21p)/(AlOH12p*OH)) - KAlOH21p;
    eq10 = (AlOH3/(AlOH21p*OH)) - KAlOH3;
    eq11 = ((AlOH41m)/(AlOH3*OH)) - KAlOH41m;
    eq12 = (Al2OH24p/((Al^2)*(OH^2))) - KAl2OH24p;
    eq13 = (Al3OH45p/((Al^3)*(OH^4))) - KAl3OH45p;
    eq14 = (AlCitHm/(AlOHCitHm2m*H)) - KAlOHCitHm2m;
    eq15 = ((Al3OH4Cit34m*(H^4))/((Al^3)*(Cit^3))) - KAl3OH4Cit34m;
    eq16 = ((Al3OH4CitHm37m*(H^7))/((Al^3)*(Cit^3))) - KAl3OH4CitHm37m;
    eq17 = Cit + HCit + H2Cit + H3Cit + AlCit + AlCitHp + AlCitHm +
2*AlCit23m + AlOHCitHm2m + 3*Al3OH4Cit34m + 3*Al3OH4CitHm37m - cCit;
    eq18 = Al + AlCit + AlCitHp + AlCitHm + AlOc + AlCit23m + AlOH12p +
AlOH21p + AlOH3 + AlOH41m + 2*Al2OH24p + 3*Al3OH45p + AlOHCitHm2m +
3*Al3OH4Cit34m + 3*Al3OH4CitHm37m - cAl;
    eq19 = Oc + AlOc - cOc;
    eq20 = ((Al*Oc)/AlOc) - KdAlOc;
    sol =
vpasolve(eq1,eq2,eq3,eq4,eq5,eq6,eq7,eq8,eq9,eq10,eq11,eq12,eq13,eq14,e
q15,eq16,eq17,eq18,eq19,eq20);

disp(['Cit3- = ' num2str(double(sol.Cit)) ' out of ' num2str(cCit) '
total Citrate']);
disp(['HCit2- = ' num2str(double(sol.HCit)) ' out of ' num2str(cCit) '
total Citrate']);
disp(['H2Cit- = ' num2str(double(sol.H2Cit)) ' out of ' num2str(cCit) '
total Citrate']);
disp(['H3Cit = ' num2str(double(sol.H3Cit)) ' out of ' num2str(cCit) '
total Citrate']);
disp(['Al3+ = ' num2str(double(sol.Al)) ' out of ' num2str(cAl) ' total
Al']);
disp(['AlCit = ' num2str(double(sol.AlCit)) ' out of ' num2str(cAl) '
total Al and ' num2str(cCit) ' total Citrate']);
disp(['AlHCit+ = ' num2str(double(sol.AlCitHp)) ' out of ' num2str(cAl)
' total Al and ' num2str(cCit) ' total Citrate']);
disp(['AlCit-H- = ' num2str(double(sol.AlCitHm)) ' out of '
num2str(cAl) ' total Al and ' num2str(cCit) ' total Citrate']);
disp(['Al(Cit)2 3- = ' num2str(double(sol.AlCit23m)) ' out of '
num2str(cAl) ' total Al and ' num2str(cCit) ' total Citrate']);
```

```
disp(['Al(OH)Cit-H 2- = ' num2str(double(sol.AlOHCitHm2m)) ' out of '
num2str(cAl) ' total Al and ' num2str(cCit) ' total Citrate']);
disp(['Al3(OH)4Cit3 4- = ' num2str(double(sol.Al3OH4Cit34m)) ' out of '
num2str(cAl) ' total Al and ' num2str(cCit) ' total Citrate']);
disp(['Al3(OH)4CitHm3 7- = ' num2str(double(sol.Al3OH4CitHm37m)) ' out
of ' num2str(cAl) ' total Al and ' num2str(cCit) ' total Citrate']);
disp(['AlOH 2+ = ' num2str(double(sol.AlOH12p)) ' out of ' num2str(cAl)
' total Al']);
disp(['Al(OH)2 + = ' num2str(double(sol.AlOH21p)) ' out of '
num2str(cAl) ' total Al']);
disp(['Al(OH)3 = ' num2str(double(sol.AlOH3)) ' out of ' num2str(cAl) '
total Al']);
disp(['Al(OH)4 - = ' num2str(double(sol.AlOH41m)) ' out of '
num2str(cAl) ' total Al']);
disp(['Al2OH2_4+ = ' num2str(double(sol.Al2OH24p)) ' out of '
num2str(cAl) ' total Al']);
disp(['Al3OH4_5+ = ' num2str(double(sol.Al3OH45p)) ' out of '
num2str(cAl) ' total Al']);
disp(['Oc = ' num2str(double(sol.Oc)) ' out of ' num2str(cOc) ' total
Oc']);
disp(['KdAlOc = ' num2str(double(sol.KdAlOc))]);
disp(['cAl = ' num2str(double(sol.cAl))]);
disp(['Solubility Product = ' num2str(double(sol.Al)*double(OH^3))]);
```

OUTPUT

```
>> AlCitOc_SolveForKd_Verbose_210514
Cit3- = 0.00015883 out of 0.005 total Citrate
HCit2- = 0.0025173 out of 0.005 total Citrate
H2Cit- = 0.0017823 out of 0.005 total Citrate
H3Cit = 4.4769e-05 out of 0.005 total Citrate
Al3+ = 6.7302e-10 out of 0.00034646 total Al
AlCit = 1.0208e-05 out of 0.00034646 total Al and 0.005 total Citrate
AlHCit+ = 2.8117e-07 out of 0.00034646 total Al and 0.005 total Citrate
AlCit-H- = 0.00015809 out of 0.00034646 total Al and 0.005 total
Citrate
Al(Cit)2 3- = 0.00016213 out of 0.00034646 total Al and 0.005 total
Citrate
Al(OH)Cit-H 2- = 2.9441e-06 out of 0.00034646 total Al and 0.005 total
Citrate
Al3(OH)4Cit3 4- = 3.2879e-07 out of 0.00034646 total Al and 0.005 total
Citrate
Al3(OH)4CitHm3 7- = 4.7507e-19 out of 0.00034646 total Al and 0.005
total Citrate
AlOH 2+ = 6.2805e-11 out of 0.00034646 total Al
Al(OH)2 + = 4.2462e-12 out of 0.00034646 total Al
Al(OH)3 = 1.0666e-13 out of 0.00034646 total Al
Al(OH)4 - = 2.1279e-16 out of 0.00034646 total Al
Al2OH2_4+ = 4.5295e-18 out of 0.00034646 total Al
Al3OH4_5+ = 3.0485e-24 out of 0.00034646 total Al
Oc = 3.8188e-05 out of 5e-05 total Oc
KdAlOc = 2.1759e-09
Solubility Product = 2.1283e-38
```

APPENDIX F

MATLAB SCRIPT USED FOR NONLINEAR REGRESSION CURVE FITTING USING APMONITOR

```
%This model uses free Al values determined by substituting experimental
%values into 'AlCitOc_SolveForKd.m' script, and fraction Oc bound
values
%(ym)determined experimentally, to minimize the difference between
experimental
%ym and predicted values, y, to generate an estimated Kd value.

%Parameters with experimentally determined values
%-----------------------------------------------
%ym
%Alm

%Nonlinear regression equation
%------------------------------
%y = Alm/(Kd + Alm)

%Type clear all; close all; clc before starting.
addpath('apm') %Must download prior to use from apmonitor.com

%select the server
s = 'http://byu.apmonitor.com';
%name the application
a = 'data_regression';

%Clear any prior applications
apm(s,a,'clear all');

%Load the model file
apm_load(s,a,'model18.apm.txt');
%Load the data file
csv_load(s,a,'GlaAlrep1.csv');

%Specify parameters with fixed values, ie, one solution
apm_info(s,a,'FV','Kd');
apm_option(s,a,'Kd.status',1);
apm_option(s,a,'nlc.imode',2);

%Solve the optimization problem
output = apm(s,a,'solve');
disp(output)

%Retrieve the results into Matlab
y = apm_sol(s,a);
z = y.x;

%Display the results
disp('Solution')

for i=1:length(z.y)
    disp(['y = ' num2str(z.y(i))])
end
```

```
for i=1:length(z.alm)
    disp(['Alm = ' num2str(z.alm(i))])
end

disp(['Kd = ' num2str(z.kd(1))])

%Generate a figure
figure(1)
plot(z.alm,z.ym,'o')
hold on
plot(z.alm,z.y,'-');
xlabel('Al^{3+} (nM)')
ylabel('Fraction Gla-Oc Bound')
legend('Measured','Predicted')
```

The model must be specified in a notepad document, as shown below:

```
Model

    Parameters
        Alm
        ym
        Kd >= 0
    End Parameters

    Variables
        y >= 0
    End Variables

    Equations
        y = Alm/(Kd + Alm)

        minimize ((ym-y)/ym)^2
    End Equations
End Model
```

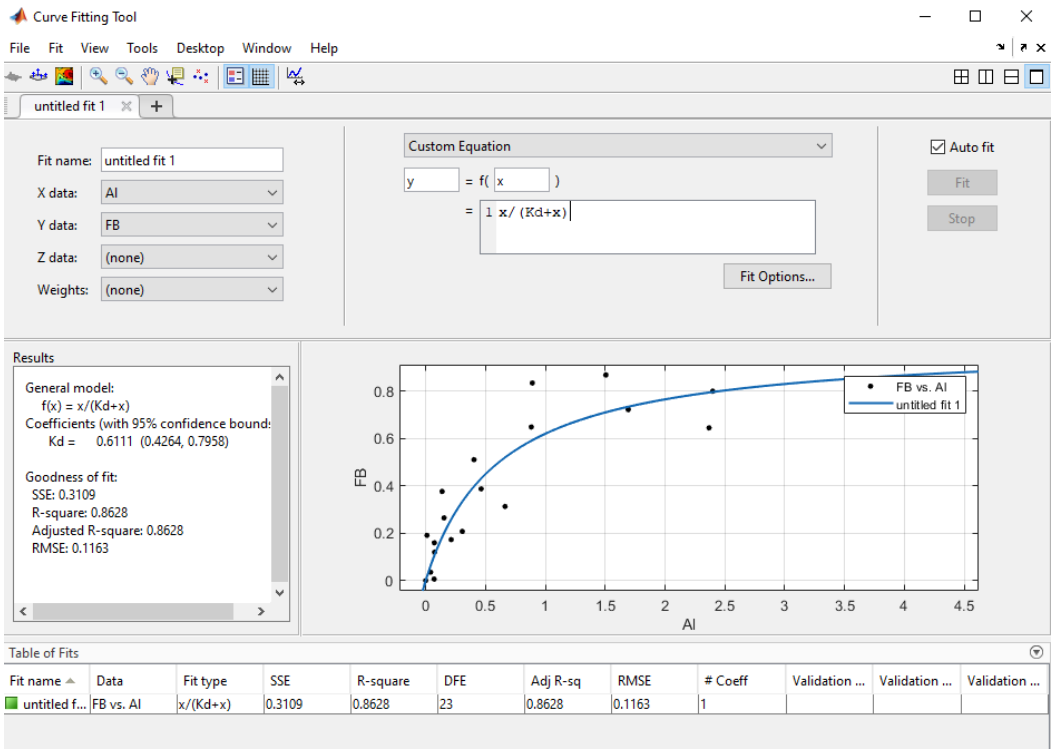
APMonitor reference [195].

APPENDIX G

MATLAB SCRIPT USED FOR NONLINEAR REGRESSION CURVE FITTING USING CFTOOL TO CALCULATE KD OF AL BINDING TO GLA-OC

```
%load your data file
load GlaAlKd.csv
%define each column of data present in the data file
Al = GlaAlKd(:,1);
FB = GlaAlKd(:,2);

cftool
```

This will result in the data being opened in Matlab's curve fitting tool, cftool:



Select X data: Al
Select Y data: FB
Select Custom Equation and type 'x/(Kd+x)'