2009

DEVELOPMENT OF THE REAGENTS AND PROTOCOLS NECESSARY TO ELUCIDATE THE BSP-BINDING REGION OF TYPE I COLLAGEN

Aaron Langdon

DEVELOPMENT OF THE REAGENTS AND PROTOCOLS NECESSARY TO ELUCIDATE THE BSP-BINDING REGION OF TYPE I COLLAGEN

(Spine title: Mapping the BSP-Collagen Interaction)

(Thesis format: Monograph)

by

Aaron Langdon

Graduate Program in Biochemistry

A thesis submitted in partial fulfillment
of the requirements for the degree of
Master of Science

School of Graduate and Postdoctoral Studies
The University of Western Ontario
London, Ontario, Canada

ABSTRACT

Bone sialoprotein, because of its affinity for the triple-helical type I collagen molecule and increased HA-nucleation potency upon binding collagen *in vitro*, is postulated to interact with type I collagen *in vivo* to initiate bone formation. Although the collagen-binding region of BSP has been located, the area of collagen involved in this interaction remains unknown. Through chemically cross-linking BSP to type I collagen, this region can be precisely mapped. The purpose of this study is to develop the reagents and protocols necessary for the covalent attachment of BSP to collagen, thereby creating the possibility of locating the BSP-binding region of collagen in the future. Five prokaryotically expressed, single-cysteine mutants of the rat BSP (1-100) peptide were expressed, purified, and conjugated to the sulfhydryl reactive, photoactivatable cross-linking reagents BPM, APB and APDP. These conjugates were incubated with type I collagen extracted from rat tail tendons, and irradiated by UV light to induce cross-link formation. SDS-PAGE and Western blot were used to detect and characterize cross-linked complexes, including covalently linked aggregates of rBSP (1-100). Using APDP, successful cross-links to collagen were detected with four of the five rBSP (1-100) mutants. Using the protocols and reagents developed in this study, it will be possible to better characterize the BSP-collagen interaction and its role in early mineral formation.

Key Words: bone sialoprotein, phosphoprotein, bone formation, hydroxyapatite nucleation, biomineralization, SIBLING proteins, chemical cross-linking, protein-collagen interactions, APDP, type I collagen

ACKNOWLEDGMENTS

I would like to thank my family for their constant support and encouragement in everything I do. Mom, Dad, Andrea, Nan, Rita and Grandad, Max, Maggie and Holly, I couldn't have made it this far without your great influence and I love you all.

Jordan, no matter what, you are always there to support me and words can't describe how much I appreciate everything you do. You have had a tremendous impact on my life in every aspect, and I love you very much.

I would also like to acknowledge the great opportunities, guidance, support, and education provided to me by my supervisors Drs. Harvey Goldberg and Graeme Hunter in both my undergraduate and graduate education. The many valuable things I have learned from you and the education I've received under your supervision will undoubtedly remain with me for the rest of my life.

I thank Drs. Stan Dunn and Eric Ball of my advisory committee and Dr. Derek McLachlin for their technical expertise, advice, and encouragement regarding my research. I also owe thanks to Kristina Jurcic, Lee-Ann Briere, and Paige Harden of the department of biochemistry for technical assistance in studies not presented here.

Within the Hunter and Goldberg labs, I would like to thank Hong, Bernd, and Heidi, with whom I have become good friends, for their valuable technical assistance, advice, and supervision. Gurpreet, I cannot begin to thank you enough for the tremendous amount of help and support I've received from you throughout my graduate education. Gurpreet, Paul, Krista, Erik, Ron, Kamal, Jason, Vasek, Brian, Dave, Kyle, and all past members of the Hunter and Goldberg labs, I thank you for your great advice, troubleshooting, and most of all, the great friendships we have established both in and outside of the lab. These recent years have been full of memorable experiences, and I have made many lifelong friends.

There are simply too many people to name that have influenced me, including all of the friends I have met in London since beginning school at Western in 2003, as well as from my hometown of Burlington.

Thank you.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

APB	ρ-azidophenacyl bromide
APDP	N-[4-(ρ-azidosalicylamido)butyl]-3'-(2'-pyridyldithio) propionamide
BCA	bicinchoninic acid
BPM	benzophenone-4-maleimide
BSA	bovine serum albumin
BSP	bone sialoprotein
CID	collision-induced dissociation
CK1	casein kinase 1
CK2	casein kinase 2
CNBr	cyanogen bromide
DMF	dimethylformamide
DMP-1	dentin matrix protein 1
DPP	dentin phosphophoryn
DSP	dentin sialoprotein
DSPP	dentin sialophosphoprotein
ECM	extracellular matrix
EDTA	ethylenediaminetetraacetic acid
ER	endoplasmic reticulum
ESI	electrospray ionization
FM	fluorescein-5-maleimide
FPLC	fast protein liquid chromatrography
FTICR	fourier transform ion cyclotron resonance
GPDH	glyceraldehyde-3-phosphate dehydrogenase
GuHCl	guanidine hydrochloride
HA	hydroxyapatite
IPTG	isopropyl-β-D-thiogalactopyranoside
LC	liquid chromatography
MALDI-TOF	matrix-assisted laser desorption/ionization, time-of-flight (mass spectrometry)
MEPE	matrix extracellular phosphoglycoprotein
MMP-1	matrix metalloproteinase-1
MS	mass spectrometry
MS/MS	tandem mass spectrometry
NCPs	non-collagenous proteins
NMR	nuclear magnetic resonance
OD	optical density
OPN	osteopontin
PBS	phosphate-buffered saline
PCR	polymerase chain reaction
PEDF	pigment epithelium-derived factor

PVDF	polyvinylidene fluoride
RGD	arginine, glycine, aspartic acid
SIBLING	small integrin-binding ligand, N-linked glycoprotein
TBST	Tris-buffered saline with Tween
TCEP	Tris(2-carboxyethyl)phosphine
THPs	triple-helical peptides
TNAP	tissue non-specific alkaline phosphatase
TNF-α	tumour necrosis factor α
VEGF	vascular endothelial growth factor

Chapter One – Literature Review

1.1 General introduction to biomineralization

Biomineralization is defined as the process where living organisms use dissolved ions to form minerals on their interior or exterior for various functions. It occurs within all five biological kingdoms [1], and possible remains of mineralized exoskeletons of prokaryotic organisms dated to roughly 3.5 billion years ago have been discovered [2, 3]. Since then, skeletal biomineralization has expanded to a diverse array of biological lineages [4]. These included bloodworms and other small organisms that developed teeth composed of elements and minerals like copper, zinc, and magnetite [5], mollusks that developed calcium carbonate-based exoskeletons, and mammals with calcium phosphate-based endoskeletons. Both of the latter types of minerals are commonly found in living organisms today. The mechanisms that organisms use to initiate and control biomineralization are poorly understood, and this is an active area of research. These control mechanisms include a complex interplay between specialized cells, as well as the proteins and ions secreted into the extracellular matrix. In an effort to develop a better understanding of mammalian biomineralization, this thesis will focus on the minerals and proteins found within mammalian bone.

1.2 General introduction to bone

Bone is a highly complex connective tissue serving many functions within the body. The most obvious functions are to provide physical support for the body and to protect internal organs. Other functions include acting as a calcium and phosphate source for the body, as well as housing the bone marrow that produces red blood cells. Bones function in tandem with tendons, ligaments, and muscles to allow the complex body movements of which mammals are capable. Within bone, there are three phases: a

cellular phase, an organic phase, and an inorganic phase. A brief description of each follows, however they will be discussed in greater detail throughout this section.

1.2.1 Cellular phase of bone

The extracellular matrix (ECM) is constantly undergoing an intricate, dynamic resorption and remodeling process that is mediated by the cellular phase within bone. This phase consists of osteoblasts, osteocytes, and osteoclasts. Osteoblasts are associated with bone deposition, osteocytes are matured osteoblasts that have been surrounded by mineral, and osteoclasts are associated with the resorption of bone.

1.2.2 Inorganic phase of bone

The inorganic phase of bone is a crystalline form of calcium phosphate, hydroxyapatite (HA). HA imparts the rigid properties of bone. The unit-cell formula of HA is $Ca_{10}(PO_4)_6(OH)_2$. Hydroxyapatite forms a very stable three-dimensional crystal lattice structure with low solubility, making it ideal for skeletal functions. The mechanism of HA deposition is poorly understood. While some research suggests that HA is directly precipitated from solution, others suggest that the calcium phosphate in mammals matures through various forms prior to forming HA [6]. The hydroxyapatite of bone has a plate-like morphology, with approximate dimensions of 45 nm by 25 nm by 5 nm [7-10]. These crystals are organized parallel to type I collagen fibrils [11].

1.2.3 Organic phase of bone

The organic phase of bone is predominantly type I collagen, while 10% is made up of non-collagenous proteins (NCPs) including bone sialoprotein (BSP) and many other often acidic proteins. These non-collagenous proteins are thought to play a key role

in the control of the mineralization process through interactions with type I collagen, hydroxyapatite, and dissolved ions.

The biological mechanisms leading to biomineralization are poorly understood. Type I collagen provides a scaffold in the extracellular matrix, upon which hydroxyapatite crystals are formed. Collagen itself does not promote HA nucleation [12-15]. However, increasing evidence supports the theory that acidic phosphoproteins within the extracellular matrix do promote nucleation of HA [15-17]. *In vitro*, bone sialoprotein has the highest nucleation potency among the non-collagenous proteins [18]. Bone sialoprotein also has an expression pattern *in vivo* that suggests a direct role in the biomineralization process [19].

1.3 Bone development

Bone is created through two developmental processes within the body, endochondral and intramembranous ossification [20-22]. Although quite distinct, both of these involve the deposition of HA onto a collagen scaffold. Long bones are developed through endochondral ossification, and flat bones through intramembranous ossification. In both types of bone development, woven bone is initially formed. Woven bone consists of mineralized type I collagen that appears unorganized and not well oriented, forming irregular bundles. Calcification of this woven bone is delayed and also appears irregular. As the bone develops, this woven bone is replaced by lamellar bone that is much more highly organized, and collagen fibers are oriented in a parallel, head-to-tail fashion. The main distinction between the two processes of bone formation is the presence of a cartilaginous precursor in endochondral ossification.

1.3.1 Intramembranous ossification

Flat bones such as the skull, mandible, pelvis, and clavicle are formed through intramembranous ossification. Mesenchymal stem cells form a bone nodule through condensation within vascularized embryonic tissue and the deposition of type I collagen [20-22]. These cells differentiate into osteoblasts that become osteocytes once surrounded by mineralized matrix. Type I collagen and the non-collagenous proteins are secreted by osteoblasts and osteocytes. These collagen fibers aggregate in irregular bundles, and woven bone is formed. Calcification is also irregular within these bundles. Vascularization occurs throughout trabeculae, leading to the formation of hematopoietic bone marrow. As previously mentioned, this woven bone is replaced with lamellar bone over time.

1.3.2 Endochondral ossification

The long bones of mammalian appendages, including hands and feet, are formed from a cartilaginous precursor through endochondral ossification. To initiate cartilage deposition, mesenchymal stem cells differentiate into prechondroblasts, and subsequently chondroblasts that become associated with the matrix to form chondrocytes [20-22]. These chondrocytes deposit a hyaline cartilaginous matrix consisting primarily of type II collagen and proteoglycans. This is considered a growth plate, and there is one present at each end of the forming tissue. The growth plate is responsible for the elongation of bones until puberty. At the mineralization front, chondrocytes become hypertrophic, undergo apoptosis, and calcification of the cartilage occurs. Once the matrix is mineralized, osteoclasts resorb areas to allow for vascularization, and osteoblasts deposit woven bone. As development continues, the woven bone is replaced by lamellar bone.

1.4 Bone repair

There are three main phases to bone repair following fracture. They are the reactive stage, repair stage, and remodeling stage. During the reactive stage, excess red blood cells are observed at the site of injury. This is followed by clotting and constriction of local blood vessels to curb additional bleeding [23, 24]. Within and around the clotting, most cells undergo apoptosis; however, fibroblasts survive and replicate forming a granulation tissue vascularized with small blood vessels [25]. Days later, cells of the periosteum nearest the fracture, as well as fibroblasts, replicate and differentiate into chondroblasts that form cartilage, initiating the reparative stage of bone healing. Periosteum cells that are further from the fracture site differentiate into osteoblasts and form woven bone [23]. These processes form the fracture callus and eventually the woven bone and hyaline cartilage bridge the fracture gap [26]. The cartilage begins to mineralize and lamellar bone replaces the woven bone present. This tissue is further innervated with blood vessels, and is then termed trabecular bone, restoring strength to the fractured bone [25]. In the remodeling phase, a complex interplay between osteoclast resorption of trabecular bone and osteoblast deposition of compact bone reshapes the bone to closely resemble its form and strength prior to fracture [25].

In some cases, if fractures are too severe or the fractured pieces are displaced, natural healing can be inhibited. In these cases, the insertion of metal scaffolds or pins to hold the bones in place can facilitate bone repair. Metal implants are also used for procedures such as dental implants and joint replacements. Additionally, implants constructed of or coated in ECM materials such as HA, collagen, or demineralized bone matrix have increased rates of osseointegration and bone repair [27]. Proteins such as

bone morphogenetic protein and bone sialoprotein have also been shown to increase mineralization rates *in vivo* when incorporated with metal or biologically derived implants [27-29].

1.5 The organic-inorganic phase interaction

Although the types of bone formation have been well-characterized, the underlying molecular events leading to hydroxyapatite nucleation and growth within collagen fibrils are not well understood. Biological fluids such as blood and milk have concentrations of calcium and phosphate ions well above the solubility products of most forms of calcium phosphate minerals. Because spontaneous calcification does not occur in healthy individuals, the body appears to have evolved mechanisms to control the nucleation and/or growth of these crystals. It is therefore likely that the body, through the use of secreted proteins, controls crystal dimensions, growth habits, and nucleation rates. There are currently three theories as to how HA is nucleated on collagen fibrils: The theory of matrix vesicles, an inhibition-based mechanism, and heterogeneous matrix-mediated nucleation. All three theories are proposed to involve the modulation of nucleation and growth of crystals by the surface chemistry of the extracellular matrix. With regards to the proteins present, this surface chemistry depends on the primary and secondary structure of the proteins as well as post-translational modifications. There have been many ECM proteins discovered to have a positive or negative effect on crystal nucleation and/or growth, and many contain highly acidic regions. The nucleation activities of these proteins can be further altered by phosphorylation [30].

1.5.1 Matrix vesicles

One theory of the mechanism of biomineralization involves matrix vesicles. These are thought to be micelles, 100 nm in diameter [31], composed of a phospholipid bilayer and derived by budding from the bodies of osteoblast, odontoblast, and chondrocyte membranes [32]. Matrix vesicles from mineralized tissues were first isolated in 1970 [33]. The vesicle contents are reportedly very high in calcium and phosphates due to various phosphatases, phosphate transporters, and calcium binding factors [33-36]. HA nucleation occurs within the vesicles due to the high local supersaturation of these ions, and crystals are subsequently deposited onto mineralized tissue. From here, the crystal grows and supports secondary nucleation, resulting in bone growth. In this postulated mechanism, the cells play a more prominent role, as they direct matrix-vesicle creation and deposition.

1.5.2 Inhibition mechanisms

Another postulated mechanism of biomineralization is that it is controlled through inhibition mechanisms. Molecules such as osteopontin and matrix extracellular phosphoglycoprotein (MEPE) inhibit HA nucleation and growth [37]. There are two modes of action associated with HA inhibitors. The first of these is the binding of a protein to the surface of a pre-existing crystal, inhibiting further deposition of ions onto the surface. A recent study indicates that a potent inhibitor of HA growth, fetuin (46 kDa), is too large to diffuse into the collagen fibril, and that the exclusion of this inhibitor leads to increased crystal growth within the collagen fibers [38]. Collagen may therefore use exclusion of these types of inhibitors of crystal growth to direct proper physiological biomineralization.

Inhibition of HA mineralization also occurs through the covalent linkage of phosphate ions. For instance, free orthophosphate can be sequestered into polyphosphates, thereby inhibiting HA mineralization by reducing local phosphate concentrations [6]. Additionally, pyrophosphate consists of two phosphate ions covalently linked by an ester bond. This molecule appears to inhibit crystal growth through the same mechanism used by inhibitory ECM proteins, by binding to crystal surfaces and preventing further ion deposition [39]. Tissue non-specific alkaline phosphatase (TNAP) cleaves pyrophosphate into its constituent phosphate ions, inactivating the inhibitor [40]. Loss of this gene causes hypophosphatasia resulting in inhibited mineralization [41]. Conversely, induction of TNAP expression initiates mineral formation [41]. Therefore, through the control of phosphate ions, mineral formation can be inhibited *in vivo*.

1.5.3 Heterogeneous nucleation

Heterogeneous nucleation is defined as the process by which nucleation occurs upon a pre-existing interface, lowering the free energy barrier to formation by homogenous nucleation. The most widely accepted theory of the mechanism of biomineralization is that of heterogeneous nucleation by proteins. According to this theory, the surface chemistry of proteins in the ECM promotes the formation of critical nuclei resulting in increased crystal nucleation [42]. Crystal formation occurs through a multi-stage process. First, ions are sequestered at the surface of the organic matrix. The ions then condense into nanoparticles on their way to forming a critical nucleus. In the third step of the process, ions deposit onto the nucleated crystal in a highly organized fashion resulting in distinct crystal faces. Finally, growth of these nucleated crystals into

a specific morphology occurs through differential ion deposition onto each crystal face [43]. With respect to bone, the nucleation and growth of hydroxyapatite crystals during this process is sensitive to the structures and charges of components of the extracellular matrix, and these components play a vital role in each step of the mineralization process. The promotion of a critical nucleus likely occurs through epitactic nucleation where a protein mimics the crystal lattice or surface ions of the crystal [42]. This is thought to be the underlying mechanism behind experiments in which bone sialoprotein and dentin phosphophoryn have been shown to increase nucleation of hydroxyapatite in agarose or gelatin matrices [18, 44, 45].

Although it has not yet been determined exactly which mechanisms lead to calcium phosphate mineralization *in vivo*, it is likely that all three of the above theories occur within the body. Matrix vesicles are too large to diffuse into the hole zones of collagen where initial HA crystals are found, and therefore may be involved in the calcification of areas where no type I collagen has been deposited, such as in cartilage calcification. The formation of crystals within the collagen fibril likely occurs through heterogeneous nucleation, while inhibition mechanisms help control crystal growth and prevent nucleation from occurring in other areas of the ECM.

1.6 A closer look at the organic phase of bone

Regardless of which mechanism or combination of mechanisms is involved in the biomineralization process, each of them relies upon components of the organic matrix. As mentioned previously, this includes mainly type I collagen as well as many non-collagenous proteins such as those of the SIBLING protein family. This protein family will be discussed shortly. Collagen has a highly organized structure and orientation that

provides a scaffold for new bone formation. Non-collagenous proteins, through interactions with collagen, HA, and the cellular phase of bone, modulate crystal formation and deposition upon the collagen scaffold.

1.6.1 Type I collagen

Although 28 types of collagen are present in the human proteome, type I collagen is the most abundant in the body. Not only is it found in bone, but also in skin, ligaments, and the cornea of the eye among other areas. It is part of a subgroup of collagens termed fibril-forming collagens along with types II, III, V, XI, XXIV, and XXVII. It is expressed by osteoblasts [46], fibroblasts [46], and chondroblasts [47]. Although the rigidity of bone is due to hydroxyapatite, its flexibility and tensile strength is the result of type I collagen fibers.

1.6.1.1 Type I collagen structure

Type I collagen fibers are highly organized and their structure is extremely important. Mutations in the genes for collagen I, COL1A1 and COL1A2, are associated with severe phenotypes such as osteogenesis imperfecta, osteoporosis, and Ehlers-Danlos syndrome [48]. The structure of type I collagen is complicated and has not been solved to the atomistic level. The microfibrillar structure of collagen, however, has been elucidated through x-ray crystallography [49].

Type I collagen molecules exist as a right-handed super-helix with non-helical telopeptide ends. The monomers involved are two α1 chains and one α2 chain that are each in a left-handed helix [50-52]. In order to facilitate this helix conformation, every third residue in the helical region of each α chain is glycine in the form Gly-X-Y, where glycine faces the centre of the helix and X and Y represent outer positions. X most

commonly represents proline, and Y hydroxyproline. Hydroxyproline forms important intramolecular hydrogen bonds with other collagen molecules [53, 54], resulting in a very stable structure resistant to proteolytic degradation [55].

Both the collagen alpha 1 and alpha 2 chains have an approximate molecular weight of 100 kDa, and are roughly 1000 amino acids long spanning 300 nm. Triple-helical collagen monomers, through their specific surface chemistry, interact with other monomers in a staggered array to form fibrils. These interactions are primarily hydrogen bonds, with very minor effects from electrostatics or hydrophobics [56]. Right-handed triple-helical monomers are packed together into left-handed fibrils [57]. This staggered array results in overlap and hole zones with a D-periodicity of 67 nm [58]. By electron microscopy, it is possible to visualize this characteristic 67 nm banding pattern of type I collagen fibrils [50-52]. Interestingly, initial mineralization of the collagen fibrils appears to originate within the hole zones [7, 9, 59-61].

1.6.1.2 Gene regulation and biosynthesis

The COL1A1 (~18 kb) and COL1A2 (~38 kb) genes are at unlinked loci 17q21.3–q22 and 7q21.3–q22.1, respectively [62]. An interesting feature is that the exons encoding the triple-helical regions are mostly 54 bp in length. Exceptions have a multiple of this number, or a combination of 54 and 45 bp segments [62, 63]. These multiples of 9 bp are essential for the repeating Gly-X-Y motif. Although the genes contain a similar number of exons and encode similar-sized protein products, introns vary greatly in size leading to the discrepancy in gene size [64]. Although transcriptional control of collagen is still largely uncharacterized, *cis*-acting elements play a role, including a promoter, enhancer, and silencer. The identification of these elements and their control mechanisms

is currently a major task for understanding the underlying transcription factors and the role they play in signaling pathways [63, 65, 66].

The α-chains are synthesized with N and C-terminal propeptides, as well as with a 22-amino acid N-terminal signal sequence, and are termed pre-pro-α-chains. With the help of signal-recognition particles and receptors, nascent translated pre-pro-α-chains protrude into the rough endoplasmic reticulum. The signal peptide is cleaved and nascent procollagen chains are post-translationally modified with hydroxylations and glycosylations prior to triple-helix formation [67, 68]. The formation of the triple-helix within the ER is aided by chaperone proteins such as Hsp47 [69]. The collagen complex is transported to the golgi apparatus where translation and triple-helix formation is completed. Procollagen molecules are then transported to the extracellular matrix through secretory vesicles.

1.6.1.3 Post-translational modification and protein folding

Roughly half of the proline residues in the Y position are hydroxylated by prolyl-4-hydroxylase [70-72]. The extent of hydroxylation depends on cofactors, as well as tissue type, and may change during development or aging [68]. Also, some proline residues in the X position may be converted to 3-hydroxyproline by prolyl-3-hydroxylase [70, 71]. Lysine hydroxylation also varies with the same factors and is facilitated by lysyl hydroxylase [73]. Glucosyl and galactosyl residues are then transferred to hydroxylysines by glycosyl transferases, forming glucosyl-galactosyl-disaccharides. Also, a mannose-rich oligosaccharide is coupled to one or both of the propeptides.

The folding of type I collagen is a complex process involving enzymes, chaperones, and propeptides [74-77]. To initiate the folding process, three C-terminal

propeptides associate and are covalently linked by disulfide bonds [78]. This trimer provides a start-site for triple-helix formation. A repeat Gly-X-Hyp region at the end of the future helical segments of the procollagen molecules nucleates triple-helix formation by establishing stable hydrogen bonds. The helix then propagate in the C- to N-terminal direction [77]. Thus triple-helix formation requires a number of factors including prolyl hydroxylase and protein disulfide isomerase [79, 80].

Following secretion into the extracellular matrix, the propeptide sequences are cleaved by procollagen N-proteinase and procollagen C-proteinase, which belong to the M12 family of Zn^{2+} -dependent metalloproteinases [81, 82]. Evidence shows that these cleaved propeptides induce feedback inhibition of type I collagen synthesis [83-85]. Collagen molecules then aggregate into fibrils, providing a scaffold for mineral formation in bone tissues.

1.6.2 The SIBLING protein family

A large category of the non-collagenous proteins is the SIBLING (Small Integrin-Binding LIgand, N-linked Glycoprotein) protein family. There are 5 members of this group: bone sialoprotein, osteopontin (OPN), dentin matrix protein 1 (DMP-1), dentin sialophosphoprotein (DSPP), and matrix extracellular phosphoglycoprotein (MEPE). Physiologically, DSPP is cleaved into dentin sialoprotein (DSP) and dentin phosphophoryn (DPP) by the proteinase PHEX. Of note, DMP-1 exists in 37-kDa and 57-kDa forms after proteolysis. OPN and MEPE are thought to be inhibitors of HA formation, whereas the other SIBLINGs are thought to act as nucleators. Although one would not group these proteins into a family based on primary sequence alone, these proteins share many common features [86]. All are expressed and secreted into the ECM

during the formation of mineralized tissues. Also, each has an Arg-Gly-Asp (RGD) sequence that is recognized by cell-surface integrins, as well as a high level of acidic residues (with the exception of MEPE where phosphorylation may make up for the lack of acidity [87]). There are many similarities in post-translational modifications as well, including phosphorylation and glycosylation. The genes for the SIBLING proteins involved in regulation of mineralization are all of similar structure, within a 600-kb region at chromosome 4q21-23 [88]. Mutations within this region are linked to mineralization disorders such as dentinogenesis imperfecta type II (likely by mutations in DMP1)[89, 90]. Because of these similarities, it has been suggested that these proteins all play similar roles in the control of the mineralization process. It should be noted that because BSP is the main focus of this thesis, information regarding this protein will be presented in greater detail separately from the other SIBLING proteins.

1.6.2.1 In vivo functional relevance

In vivo, BSP-null mice, as will be further discussed, have delayed mineral formation among other poor bone phenotypes [91], whereas OPN and MEPE knockouts have an increased bone mass and mineral content [37, 92]. When DMP-1 expression is abolished, mice develop skeletal deformations including osteomalacia, enlarged growth plates, and dentin wall hypomineralization [93-95]. DSPP knockout mice display extensive tooth abnormalities [96]. When BSP is knocked out in mice, there is a pronounced bone phenotype including reduced bone length and decreased mineral density [91]. Because this phenotype is much more pronounced in younger mice, it appears as though BSP plays an important role in early mineral formation. Taking these findings into account, along with studies of SIBLING proteins and their effect on

mineralization, it appears as though the body utilizes the SIBLING proteins to maintain tight control over the biomineralization process.

1.6.2.2 SIBLING structure

The SIBLING proteins all have a flexible secondary structure. OPN is found to possess a random coil structure through NMR and other studies [86]. Analysis of DMP-1 by computational analysis [97], circular dichroism [98, 99], and small-angle x-ray scattering [100] suggests a random coil secondary structure, and NMR analysis of DPP has indicated a lack of structure as well [101]. BSP is also unstructured, a fact that will be further discussed in this chapter. Although it is possible that one or more of these proteins may have a region that adopts secondary structure upon binding hydroxyapatite like statherin appears to [102, 103], their general flexible structure is thought to be necessary for interaction with a wide variety of binding partners including collagen, integrins, and hydroxyapatite.

1.6.2.3 Collagen-binding properties

Many of the SIBLING proteins have been found to bind to type I collagen. As will be discussed later, BSP has a high affinity for collagen [104, 105]. OPN also binds collagen, albeit more weakly than BSP, and does so through a highly conserved region containing both hydrophobic and positively-charged amino acids [106]. DMP-1 binds to the N-terminal telopeptide of the α1 chain of collagen through two distinct regions [107]. DPP has been found to bind collagen within the N-terminal quarter of the triple-helical region of the molecule [108, 109]. Some of these proteins may also be covalently linked to collagen through transglutaminase 2, which can cross-link the γ-carboxamide of glutamine residues to lysine ε-amino groups [110, 111]. *In vitro*, this enzyme appears to

catalyze the covalent attachment and polymerization of BSP, OPN, DMP-1, and DPP [111], although the physiological relevance of this observation is unclear.

1.6.2.4 Post-translational modifications

Each SIBLING protein has extensive post-translational modifications. Although many have multiple sites of glycosylations, no functional relevance of this type of modification has yet been discovered. Phosphorylations are a major post-translational modification of the SIBLING protein family. These occur mainly on serine residues, with some on threonines, and fewer on tyrosine residues. According to consensus sequence analysis, the protein kinases CK1, CK2, and golgi-casein kinase are thought to be responsible for most of these phosphorylations. Rat bone OPN, for instance, has over 30 potential phosphorylation sites with an average of 10 phosphates per molecule *in vivo* [112]. The DMP-1 37-kDa and 57-kDa fragments have an average of 12 and 41 phosphates, respectively [113], while DSP has an average of 6–10 phosphates. DPP has an astounding average of 209 phosphorylations [114, 115]. Phosphates appear to be very significant with regards to protein function, as phosphorylation of DMP-1 and DPP have been positively associated with the proteins' abilities to nucleate hydroxyapatite [113, 116]. Corroborating this data, Torres-Quintana *et al.* determined that dentin mineralization was negatively affected in the presence of kinase inhibitors [117]. Conversely, phosphorylation of OPN appears to increase its ability to inhibit mineral formation [112, 118].

1.6.3 BSP

BSP, mentioned only briefly thus far, is one of the most extensively studied SIBLING proteins and appears to have the most potent mineralization capabilities of the

ECM proteins. This acidic protein was first isolated as a 23-kDa breakdown product from bovine bone [119], and purified in its intact form in 1983 [120]. BSP has subsequently been isolated from mineralized tissues of humans [121], pigs [122], rats [123], rabbits [124], amphibians [125], and reptiles [125] and has been found to contain highly conserved functional regions indicating that BSP is a highly evolved protein involved in vertebrate mineralization. BSP is essentially unique to mineralized tissues and is expressed at high levels in areas of *de novo* bone formation [126, 127]. BSP has a high affinity for type I collagen [104] and is found to nucleate hydroxyapatite *in vitro* [44], suggesting that this protein binds to type I collagen to elicit mineralization in areas of new bone formation.

Mammalian BSPs are roughly 327 amino acids in length including a 16-residue signal sequence. Without post-translational modifications, the protein has a molecular weight of 33-34 kDa. Glu and Gly residues constitute about one-third of the total amino acids, with the high Glu content concentrated in two poly-glutamic acid segments in the N-terminal half of the molecule [128]. Like the other SIBLING proteins, BSP has an RGD sequence which is located near the C-terminus of the protein. A hydrophobic region near the N-terminus (between residues 19 and 46) has been found to bind to type I collagen [129] (Figure 1.1). Sequence identity in mammalian BSP is high at 45%, with up to an additional 23% in conservative replacements. The majority of this conservation is at both termini of the molecule, in the collagen-binding domain, RGD region, and poly-Glu regions, as well as potential sites of post-translational modifications [88]. The functional regions of BSP display 85% - 90% identity between the mammalian species (Figure 1.1).

```
Rat             FSMKNFHRRIKAEDSEENGVFKYRPRYFLYKHAYFYPPLKRFPVQGGSDSSEENGDGDSS 60
Mouse           FSMKNFHRRIKAEDSEENGVFKYRPRYFLYKHAYFYPPLKRFPVQGGSDSSEENGDGDSS 60
Bovine          LSMKNLNRRAKLEDSEENGVFKYRPQYYVYKHGYFYPALKRFAVQSSSDSSEENGNGDSS 60
Human           FSMKNLHRRVKIEDSEENGVFKYRPRYYLYKHAYFYPHLKRFPVQGSSDSSEENGD-DSS 59
Porcine         FSMKNFHRRAKLEDPEENGVFKYRPRYYLYKHAYFYPPLKRFPVQSSSDSSEENGNGDSS 60
                :****::** * **.**********:*::***.**** ****.**..********: ***

Rat             EEEGEEEETSNEEENNEDSE---GNEDQEAEAEN---ATLSGVTASYGVETTADAGKLEL 114
Mouse           EEEGEEEETSNEEENNEDSE---GNEDQEAEAENSTLSTLSGVTASYGAETTPQAQTFEL 117
Bovine          EEEEEEEETSNEEGNNGGNEDSDENEDEESEAEN---TTLSTTTLGYG-EITPGTGDIGL 116
Human           EEEEEEEETSNEGENNEES-----NEDEDSEAEN---TTLSATTLGYGEDATPGTGYTGL 111
Porcine         EEEEEEEENSNEEENNEENEDSDGNEDEDSEAEN---ITLSTTTLGYGGDVTPGTASIGL 117
                *** ****.***  **  .     ***:::****    *** .* .** : *. :    *

Rat             AALQLPKKAGDAEGKAPKMKESDEEEEEEEEEEEENENEEAEVDENEQVVNGTSTNSTEV 174
Mouse           AALQLPKKAGDAESRAPKVKESDEEEEEEEEEEE-NENEEAEVDENELAVNGTSTNSTEV 176
Bovine          AAIWLPRKAGATGKKATKEDESDEEEEEEEE-----EENEAEVDDNEQGINGTSSNSTEV 171
Human           AAIQLPKKAGDITNKATKEKESDEEEEEEEEGNE-NEESEAEVDENEQGINGTSTNSTEA 170
Porcine         AALQLPKKAGDIGKKSAKEEESDEDEEEEEEN----EENEAEVDDNEQGTNGTSTNSTEV 173
                **: **:***    ::.* .****:******     *:.*****:**   ****:****.

Rat             DGGNGPSGGDNGEEAE--EASVTEAGAEGTTAGARELTSYGTTTAVLLNGFQQTTPPPEA 232
Mouse           DGGNGSSGGDNGEEAEAEEASVTEAGAEGTTGG-RELTSVGTQTAVLLNGFQQTTPPPEA 235
Bovine          DNGHGSSGGDNGEED-GEEESVTEANTEGITVAG-------ETTTSPNGGFKPTTPHQEV 223
Human           ENGNGSSGGDNGEE--GEEESVTGANAEGTTETGGQGKGTSKTTTSPNGGFEPTTPPQ-V 227
Porcine         DSGNGHSGGDNGEE--GDQESVTEA--QGTTVAGEQDNGGAKTTTSPNGGLEPTPPPQDI 229
                :.*:* *******    :*: ** *  :* *            *:   .*:: *.*

Rat             YGTTSPPARKSSTVEYGEEYEQIG-NEYNTAYETYDENNGEPRGDTYRAYEDEYSYYKGH 291
Mouse           YGTTSPPIRKSSTVEYGGEYEQTG-NEYNNEYEVYDNENGEPRGDTYRAYEDEYSYYKGH 294
Bovine          YGTTPPPFGKITTP---GEYEQTGTNEYDNGYEIYESENGDPRGDNYRAYEDEYSYYKGR 280
Human           YRTTSPPFGKTTTVEYEGEYEYTGVNEYDNGYEIYESENGEPRGDNYRAYEDEYSYFKGQ 287
Porcine         SGTTLPPSGKTTTPEYEGEYEQTGAHEYDNGYEIYESENGEPRGDSYRAYEDEYSYYKGR 289
                  ** **  * :*     ***  * :**:. ** *:.:**:****.**********:**:

Rat             GYEGYEGQDYYYHQ 305
Mouse           GYEGYEGQNYYYHQ 308
Bovine          GYDSYDGQDYYSHQ 294
Human           GYDGYDGQNYYHHQ 301
Porcine         SYNSYGGHDYY--- 300
                .*:.* *::**
```

Figure 1.1 Alignment of known mammalian BSP amino acid sequences.

The functional domains of rat BSP have been highlighted. The collagen-binding region is in red, the poly-Glu regions in green, and the RGD sequence in blue.

1.6.3.1 Gene structure and regulation

The human BSP gene (IBSP) was initially mapped to the long arm of chromosome 4 [130, 131], and has since been more accurately located to human chromosome 4q21.2-4q21.3 in close proximity to the other SIBLING proteins [132]. IBSP is roughly 15 kb and consists of seven exons with six introns [131, 133]. Exon 2 encodes the 16-residue signal sequence and the first two amino acids of BSP, exon 3 encodes 15 amino acids, and exon 4 encodes 27 amino acids. Exons 5 and 7 encode the two poly-glutamic acid regions of BSP, while the RGD motif is also encoded in exon 7.

The first ~370 bp of the BSP promoter is highly conserved between mammalian species, with relatively poor conservation in the avian gene [134]. A TATA-like sequence (TTTATA) and an inverted CCAAT box (ATTGG) are required for basal activity [133, 135]. A vitamin D_3 response element overlaps the inverted TATA box, and the presence of D_3 results in suppression of BSP expression [136]. Expression of BSP is induced by glucocorticoids [137-139] and bone morphogenetic proteins (BMPs) [140, 141] that are associated with promotion of differentiation of cells into osteogenic lineages [142-145]. The fact that BSP expression is limited to mineralizing tissues suggests that transcription of the gene is normally inhibited; however, the mechanism of this inhibition is unknown. Tumour necrosis factor-α (TNF-α), a prominent mediator of inflammatory responses in periodontal disease, increases bone resorption and inhibits HA deposition. TNF-α, through a tyrosine-kinase dependent pathway, inhibits BSP gene transcription [146].

Mechanical forces on cells may also regulate gene regulation in bone [147], as the absence of force upon the skeleton during immobilization or space flight can lead to significantly increased bone resorption [148, 149]. The application of magnetic fields to

osteogenic cell lines has been shown to increase bone sialoprotein mRNA levels after 24 h of stimulation [150]. Additionally, chlorpromazine, a tranquilizing drug for the treatment of psychiatric disorders that mimics hypotonic stress, causes membrane deformation and imparts mechanical stress on cells, suppresses BSP expression in osteoblast-like cell lines [151]. Evidently BSP expression is mediated by many factors including hormones, pharmacological agents, vitamins, transcription factors, and mechanical forces.

1.6.3.2 Tissue distribution

BSP expression, under normal physiological conditions, is specific to mineralizing tissues [126, 127]. It should be noted, however, that expression levels in mineralized cartilage, dentin, and enamel are much lower than in newly forming bone and cementum [127]. Expression of BSP occurs along with the formation of mineral in both membranous and endochondral bones [127]. Although no BSP mRNA is detectible in pre-osteoblasts, it is easily detected in osteoblasts [126, 152] and osteocytes [153, 154] of embryonic bones, indicating that BSP is induced as cells differentiate into osteoblasts. Similar findings apply to cementoblasts [155]. BSP protein is present in almost all mineralized mammalian tissues [88], where the only non-mineral associated sites where BSP is found are the trophoblastic cells of the human placenta [152], and platelets where it is probably internalized by endocytosis from the serum [156]. Although some studies using cell lines have reported that BSP is expressed by osteoclasts as well [152, 154], other studies contradict these findings and suggest that the expression observed may be due to the osteosarcoma cell lines used for study [157, 158]. Serum levels in humans vary between 5.0 and 21.6 ng/mL with a linear increase with age in females [159].

Immunogold-labeling studies in human and porcine subjects show BSP concentrated at the mineralization front [19, 160, 161]. Additional human studies reveal a uniform distribution of BSP throughout trabecular and cortical bone, as well as associated with electron-dense "grey patches" between collagen fibers presumably associated with mineral crystals [19, 161, 162]. Kasugai and coworkers performed experiments in which radiolabeled BSP was expressed by cultures of rat bone cells. They found that soon after expression, most of the radiolabeled BSP was incorporated into the cell matrix, with small amounts secreted into the culture medium. Sequential extractions of the matrix layer with 4 M GuHCl and 0.5 M EDTA show that some BSP was associated with the collagenous matrix, while most was bound to pre-existing hydroxyapatite. Some of the radiolabeled BSP was detectable only after de-mineralization of the matrix, suggesting that secreted BSP deposited in the osteoid is entrapped by hydroxyapatite [163].

Pathologically, BSP is expressed in many cancers that metastasize to bone, including breast, lung, thyroid, and prostate cancers [164-167]. BSP expression in cancer cells is also associated with the presence of HA-mineral deposits [165]. BSP binds to factor-H (protecting cells from complement-mediated lysis) [168] and has a postulated role in the mediation of matrix metalloproteinases allowing cancerous cells to become more invasive [168-170]. BSP also increases the angiogenic capacity of tumours by mediating vascular endothelial cell attachment and migration [170]. Expression of BSP therefore enhances the ability of cancer cells to become more invasive, increase angiogenesis, metastasize to bone, and evade lysis. BSP expression in tumours is, unsurprisingly, associated with poor prognosis [171]. Other pathological expressions of

BSP occur in heterotopic ossification of pressure sores in paraplegics [172] and in calcification sites of atherosclerotic plaques in coronary arteries [173].

1.6.3.3 Secondary structure

Secondary structure prediction analysis suggests that BSP has an open, flexible structure with potential to form α-helices and β-sheets [153]. A rotary shadow-casting technique shows the protein to be in a generally extended structure roughly 40 nm in length, with the C-terminal portion appearing to fold over itself [174]. NMR, small angle x-ray scattering, and circular dichroism studies of BSP also indicate a random coil structure [86, 175]. As suggested for the SIBLING proteins, BSP most likely utilizes its highly flexible structure to be able to interact with a wide variety of binding partners.

1.6.3.4 Post-translational modifications

Up to 30% of the molecular weight of native BSP is attributed to post-translational modifications. The distribution and level of modification are likely tissue-dependent. The modifications found on BSP include phosphorylation, glycosylation, tyrosine sulfation, and transglutaminase cross-linking.

Although two tyrosines have been found to be sulfated near the RGD sequence [176], there is no apparent effect of these modifications on cell-attachment activity, or any other known function of BSP for that matter [177].

The BSP sequence has three potential sites for N-linked glycosylation near the middle of the protein [153], two of which are conserved across all mammalian species. There is a fourth potential N-linked glycosylation site in the human form of BSP [130]. In UMR 106-01 cells, a rat osteosarcoma cell line, there is evidence of 3 N-linked and between 21 and 24 O-linked glycosylations on the rat BSP molecule [178]. More recent

studies using bone-extracted human BSP have utilized chromatography and mass spectrometry to determine the specific sites of glycosylation. From these studies it appears as though human bone-extracted BSP has between 2 and 4 sites of N-linked glycosylation, and 8 – 11 O-linked glycosylations [174, 176]. These and other studies indicate a high level of heterogeneity with regards to glycosylations of BSP. These glycosylations have been postulated to affect cell attachment activity [105], however there is no apparent affect on the HA nucleation potency or collagen-binding activity of BSP [179, 180].

Phosphorylation is perhaps the most significant post-translational modification of ECM proteins, as it appears to affect many of the SIBLING protein abilities to enhance or inhibit mineral formation. Through observation of consensus sequences, BSP has 5 potential protein kinase C phosphorylation sites, 9-11 for protein kinase CK2, and 1-2 for tyrosine kinase [88]. Additionally, inhibiting CK2 in cell culture impedes mineral formation [181], a sign of its importance in mineralized tissue *in vivo*. A series of studies have been conducted *in vitro* and *in vivo* to determine levels and locations of phosphorylation of BSP, with varying results [176, 182-185]. Likely, phosphorylation at each site depends on the organism, tissue, and temporal factors, leading to the heterogeneous results in these studies. These phosphates are not found to alter BSP's ability to bind collagen [129]. However, as will be discussed shortly, they do affect mineralization properties [175].

1.6.3.5 Function

BSP has a number of binding partners. The RGD motif of BSP is highly conserved. This motif binds the cell-surface integrin $\alpha_v\beta_3$ [186] on the surface of

osteoblasts [187] and osteoclasts [188, 189] and is involved in focal adhesions of these cells [190-192]. An interaction with this integrin in osteoclasts results in an increase in bone resorption [193, 194], while an interaction with osteoblasts causes increased cell differentiation and mineral formation [195, 196].

Two poly-glutamic acid regions within BSP are responsible for tight (Kd ~ 2.6 nM) binding to hydroxyapatite crystals [177, 197]. Each of these regions contains a contiguous Glu sequence, with a high proportion of periodic glutamic acid residues in close proximity on either side of these regions as well (Figure 1.1). In 1993, it was discovered that BSP nucleates hydroxyapatite *in vitro* using a steady-state, double-diffusion agarose gel system [44]. Subsequently, using the same system, it was discovered that the two poly-glutamic acid regions of BSP were responsible for this nucleation effect [175, 180, 198]. An experimental screen of bone and tooth-matrix proteins revealed that BSP was the most potent nucleator of HA *in vitro* [18, 199]. As mentioned previously, post-translational modifications play a role in this nucleation effect: native BSP appears to nucleate HA at a concentration 100 times lower than prokaryotically-expressed recombinant BSP [175]. Published and unpublished data from our laboratory suggest that a contiguous sequence of 8 poly-glutamic acid residues is required for HA nucleation. For instance, both rat and porcine BSP have two Glu-rich domains: One contiguous Glu sequence that is equal to or greater than, and one which is less than, 8 residues in length. When peptides of these are examined for nucleation potency, only those containing the longer poly-Glu sequence appear to nucleate HA [175, 198]. Mutagenesis experiments in which the poly-Glu regions were interrupted with a proline residue or where the whole region was mutated to poly-alanine, indicate that the

continuous glutamic acids and the presence of both poly-Glu regions are very important and that the charge of this region plays a large role in nucleation ability [175].

BSP is able to bind to type I collagen and through doing so, alters collagen fibril-forming characteristics *in vitro* [104, 105, 200]. Through expression of recombinant peptides, our lab has mapped the collagen-binding region of BSP to residues 19-46, determined a Kd of 12 nM [129], and characterized the interaction as mainly hydrophobic in nature [45]. Interestingly, in addition to being highly conserved, this region has an increased abundance of both hydrophobic and basic amino acids. BSP's ability to nucleate HA was increased upon binding to type I collagen [45], providing additional evidence that BSP is used in the ECM to elicit mineral formation within collagen fibrils. Using a far-Western blotting technique, Fujisawa and coworkers observed that BSP bound to the α2 chain of collagen, and more specifically to the α2CB4 and α2CB3-5 cyanogen bromide peptides of this chain [105]. However, the observed binding was fairly non-specific, where the α1 chain and multiple CNBr peptides also bound BSP. More recently, our lab has established that the affinity of BSP for triple-helical collagen is significantly higher than to gelatin, suggesting that the triple-helix is important for proper binding [45]. Taking into consideration the non-specific binding of BSP to collagen in the study by Fujisawa, as well as the fact that the collagen used was not in a triple helical form, the region of collagen to which BSP binds is unclear.

1.7 Purpose of thesis

The development of therapeutics for non-union fractures, biomaterials for tissue engineering, and implant coatings for enhanced osseointegration require an in-depth understanding of the biological mechanisms behind controlling mineralization. Evidently,

the control of biomineralization *in vivo* is an extremely complex process involving many factors. These include the differentiation of cells, the presence of type I collagen and other proteins of the extracellular matrix including the SIBLING protein family, calcium, phosphate, and hydroxyapatite, as well as interactions between all of these contributors. A better understanding of these interactions is pivotal to elucidating the way in which the body controls mineralization. Bone sialoprotein is the major non-collagenous protein in bone capable of HA nucleation, and therefore it is of interest to focus on this protein's properties and interactions to understand its role in the bigger picture. It is hypothesized that BSP is secreted into the ECM by osteoblasts, binds to type I collagen, and from there elicits mineral formation. The expression and nucleation ability of BSP have been fairly well characterized. As BSP's nucleation potency is increased upon binding type I collagen, this interaction must be important *in vivo*. The region of collagen to which BSP binds is unclear, but important in fully understanding the mineralization process. Perhaps BSP binds to collagen within or adjacent to the hole zones, where initial mineral formation is seen to occur, and may be responsible for this nucleation. Our lab seeks to delineate this region through covalent attachment of BSP to collagen using chemical cross-linking, followed by digestion and characterization of products by Western blot and ultimately mass spectrometry. This method will be discussed in greater detail in the next chapter. The purpose of this thesis is to establish the reagents, as well as develop the protocols necessary to proceed with these experiments. Through this work, our lab will eventually be able to determine the exact region of collagen that is responsible for interaction with BSP. This information would provide us with a better understanding of the initial stages in biomineralization leading to bone formation.

Chapter Two – Establishing a Cross-link Between BSP and Type I Collagen

2.1 Introduction

2.1.1 Mapping ligand-binding regions of type I collagen

Type I collagen forms a stable, insoluble fibrillar structure *in vivo* with various binding partners. Many studies have been conducted to elucidate the binding region(s) of various proteins on the type I collagen molecule. Because of the propensity of collagen to form large insoluble aggregates primarily through intermolecular cross-links and hydrogen bonding between collagen chains, *in vitro* experiments are difficult to conduct. In order to undertake protein-collagen interaction studies, modifications of traditional biochemical techniques are likely required. However due to this complexity of collagen, many researchers limit their studies to attempting to visualize proteins on extracted insoluble collagen fibrils. The various ways in which researchers have located binding sites on type I collagen are summarized below.

2.1.1.1 Electron microscopy

Electron microscopy, in conjunction with either metal-labeled proteins or rotary shadowing, has been used extensively to visualize binding locations on collagen. For instance, San Antonio and coworkers mapped the binding location of heparin to the N-terminal region of type I collagen using gold labeling with electron microscopy [201]. Through rotary shadowing, Keene and coworkers mapped the decorin-binding site of type I collagen to a region near the C-terminus [202]. Methods similar to these have been extensively used to map binding sites on type I collagen for integrins [203], cartilage oligomeric matrix protein [204], dentin phosphophoryn [205, 206], and proteoglycans [207]. Although these techniques provide a striking visual image of a ligand on collagen, the observations are only qualitative in nature, and do not allow one to map exact binding

locations down to the amino acid sequence, or even determine which of the α-chains of collagen the ligand is bound to.

2.1.1.2 Cyanogen bromide cleavage

In an effort to more accurately map ligand-binding locations on collagen, some researchers have cleaved collagen into fragments and by dot-blot, far-Western blot, competitive-binding assays, or other methods, determined which fragments have high affinity for the ligand of interest. Most commonly, cyanogen bromide (CNBr) has been used to cleave collagen, as the triple-helix of collagen is largely resistant to enzymatic cleavage. Staatz and coworkers purified collagen CNBr fragments and determined binding locations for the α2β1 integrins through a platelet-adhesion assay [208]. Similarly, through a dot-blot binding assay, Beher and coworkers determined that CNBr fragment 6 of the α1 chain (α1CB6) contains the binding location for amyloid protein [209]. Similar experiments have been conducted to map binding locations for fibronectin [210], other integrins [211], interleukin-2 [212] , phosphophoryn [213], and BSP [105]. However, once cleaved with CNBr, collagen is no longer in its triple-helical form. Most of these proteins were found to bind to denatured type-I collagen to the same extent as triple-helical collagen, and therefore binding-locations are most likely physiological. However, some proteins may require collagen to be in its triple-helical form for physiological binding. BSP appears to require triple-helical collagen, as it has a higher affinity for this form over heat-denatured collagen [45]. This fact calls into question previous studies that have mapped BSP binding locations using collagen CNBr peptides. This may also be the case for other collagen-binding partners.

2.1.1.3 Synthetic triple-helical peptides

In some instances, particularly for collagens II and III, short synthetic triple-helical peptides have been used to mimic physiological collagen in order to determine collagen-binding characteristics and locations of cells or proteins. For collagen II and III, synthetic triple-helical peptides (THPs) have been created that span the entire length of the molecule, termed a THP "toolkit", and these peptides have been used extensively to map protein or peptide-binding locations [214]. This approach has been used less extensively with type I collagen as of yet. Emsley and coworkers recently characterized the triple-helical structure of a synthetic, mimetic type I collagen peptide complex that is 21 residues in length, and binds the α2β1 cell-surface integrin [215]. However, triple-helical peptides for the majority of type I collagen have not been synthesized, making these types of studies with type I collagen unfeasible. Another critique of this type of analysis is that if a protein binds to a region of collagen that overlaps two of the triple-helical peptides, proper binding may be abolished.

2.1.1.4 Site-directed mutagenesis

Another method that has been used to map specific regions of collagen responsible for binding ligands is through mutagenesis of suspected binding sites on collagen. This has been used to confirm fibronectin binding-sites [216] as well as the 6-residue sequence responsible for binding the α1β2 intergin [217]. However, for this approach to be economically viable, a fairly specific binding region must already be established.

2.1.2 Chemical cross-linking

The study of binding locations on interacting proteins is quite problematic. Because most protein-protein interactions are non-covalent in nature, these dynamic binding events are in an equilibrium state where proteins bind and are subsequently released over time. Many of the current biochemical techniques for studying proteins would disrupt non-covalent interactions, further complicating this type of study.

Some conventional techniques used to characterize protein-protein interactions are nuclear magnetic resonance, x-ray crystallography, and mass spectrometry. Unfortunately, the nature of BSP and collagen prohibit these types of studies. BSP has very low expression levels, and therefore nuclear magnetic resonance experiments are not economically viable. Additionally, the highly acidic nature of BSP prevents the formation of crystals for x-ray crystallography. Although BSP alone is compatible with mass spectrometry, the very large size of collagen and heterogeneity of post-translational modifications severely complicates MS analysis.

The process of chemical cross-linking averts most of these limitations by creating a covalent attachment between two interacting proteins. The covalent complex can then be further characterized by methods including chemical or enzymatic digestions, Western blots, and mass spectrometry that would not be possible if the proteins were not covalently attached. Although the covalent cross-linking of proteins has been used for decades, more recent advances in analysis techniques such as newer, high-resolution mass spectrometers, as well as an increase in the number of third-party companies inventing and producing various cross-linking compounds, have increased interest in this type of study over the past 10 years. Many studies are being conducted in which two

proteins are covalently cross-linked and subsequently cleaved into smaller fragments chemically or enzymatically. These fragments can be characterized by mass spectrometry to determine which fragments correspond to cross-linked peptides, thereby mapping the protein-protein interface. However, with conventional mass spectrometers such as MALDI-TOF MS, resolution limitations have made deconvolution of the data difficult when there are many cleavage products of similar sizes, as well as incomplete cleavage products. This would be especially difficult in the case of collagen because of post-translational modifications, modifications imparted during digestion, and inherently incomplete cleavage reactions. With the current refinement of high-resolution mass spectrometers, as well as advanced analysis software, protein-protein interaction studies utilizing cross-linking with subsequent MS analysis have become increasingly routine [218, 219].

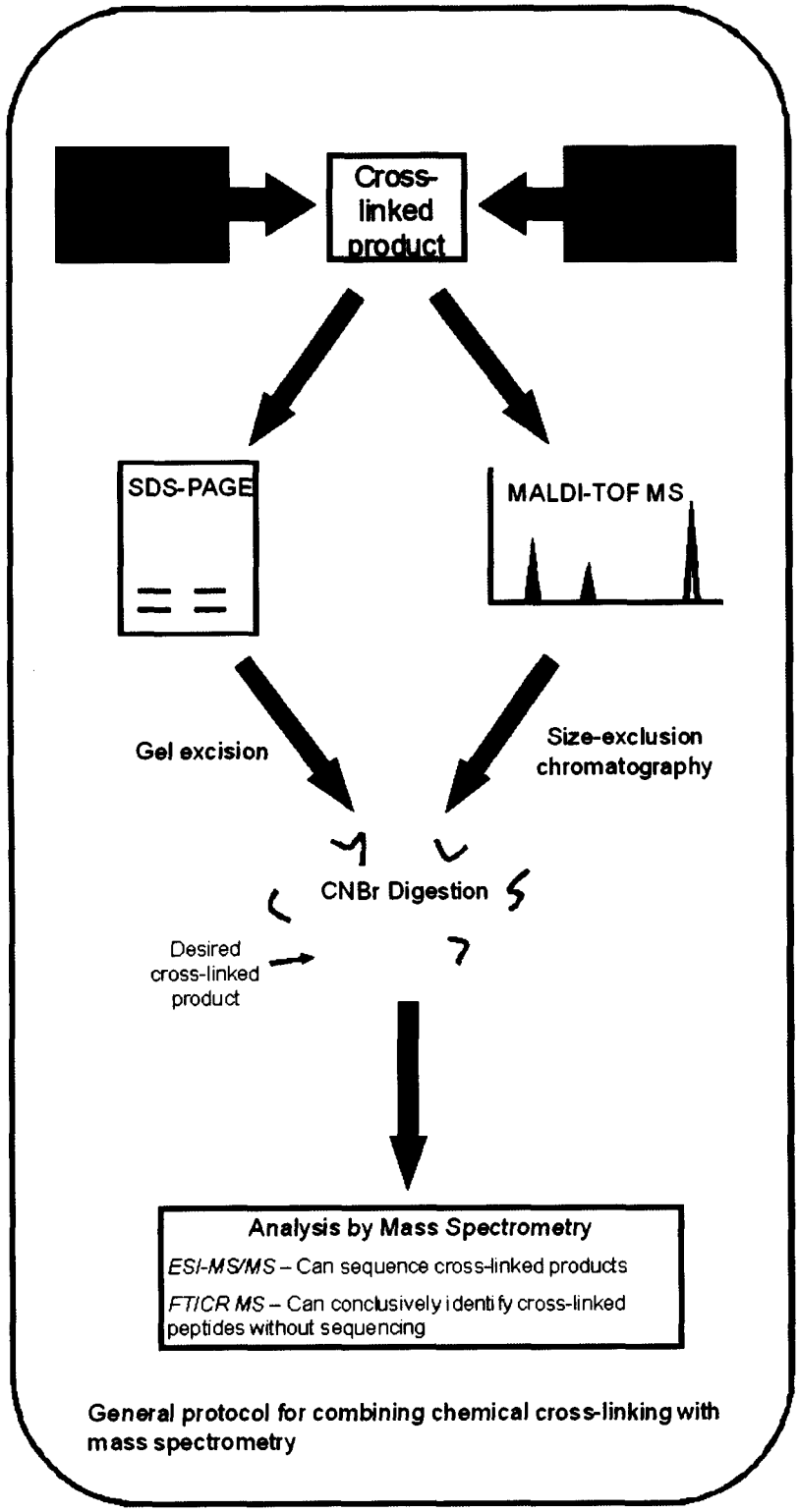
With the advent of these new techniques to study protein-protein interactions, the possibility of determining the binding site for BSP on type I collagen has become more promising. The purpose of this M.Sc. project is to create the reagents needed, as well as develop the protocols necessary to obtain a successful cross-link between BSP and collagen. Once this has been accomplished, it will eventually be possible to map the BSP-collagen interface through mass spectrometry. An outline of the proposed protocol is depicted in Figure 2.1.

2.1.2.1 Types of cross-linking compounds

There are many cross-linking compounds available that are applicable to protein chemistry. Typically, a cross-linker consists of two reactive groups that can react with specific amino acid side-chains, attached by a spacer-arm. If the two functional groups

Figure 2.1 General outline of a potential cross-linking and analysis protocol.

Either SDS-PAGE or MS could potentially be used to identify the presence of a BSP-collagen complex after cross-linking. This complex could then be isolated and further characterized by digestion and characterization through mass spectrometry.



General protocol for combining chemical cross-linking with mass spectrometry

are the same, the cross-linker is said to be homobifunctional. If the groups are different, it is a heterobifunctional cross-linker. There are also commercially available heterotrifunctional reagents that have a third functional group, that may either be for cross-linking a third protein, or may be a "tag" for identification or purification purposes. In what are referred to as "one-step" protocols, both proteins of interest are reacted with the cross-linking reagent at the same time; in "two-step" protocols, the cross-linker is first reacted with one protein, then in the second step the complex is cross-linked to the second protein. Most commonly, heterobifunctional reagents are used, and a two-step protocol is employed. This is because with a heterobifunctional reagent, one functional group can be conjugated to a specific site on the first protein, and the second functional group is usually one that can bind protein fairly non-specifically. This way, there is much more control over the cross-linking process.

2.1.2.2 Amino acid reactivity

Most cross-linking functional groups require amino acid side chains that act as nucleophiles. Acidic and basic amino acids act as stronger nucleophiles when the pH is above their pKa, and therefore the sidechain is deprotonated. This results in a negatively charged acidic residue, or a basic residue that possesses no net charge. Cysteine residues act as very strong nucleophiles in their unprotonated thiolate form. The theory of nucleophilicity states that the thiolate is actually the strongest nucleophile relative to the α-amines of the N-terminus of proteins, ε-amines of lysine residues, and the secondary amines of histidine residues [220-222]. The pKa of the cysteine sulfhydryl is 8.8 – 9.1, whereas the N-terminal α-amines 7.6 – 8.0, and ε-amines of lysine 9.3 – 9.5. Therefore, one would expect that the pH would have to be quite high in cross-linking reactions in

order to attach a cross-linking molecule. However, the surrounding microenvironment of amino acid residues in proteins significantly alters the pKa of individual residues [223]. For instance, ε-amines of lysine residues that have a pKa between 9 and 10 actually exist in large numbers in their deprotonated form even at pH 7.2 [223]. Therefore, only slightly alkaline conditions are required for cross-linking reactions. This does, however, complicate issues if it is desired to keep the reaction specific to a given amino acid side-chain. Cysteine residues are an attractive target for cross-linking reactions, as they are quite rare in proteins. Therefore, it is usually known exactly where on the protein the cross-linker is bound. The fact that the thiolate of cysteine residues are the most nucleophilic group is an added bonus.

2.1.2.3 Cysteine-reactive cross-linkers

A heterobifunctional cross-linker that is conjugated to a cysteine residue at one end must have another functional group that will covalently attach to a residue on another protein. This second functional group could be any number of chemical compounds. However, the less specific the functional group to a particular amino acid, the better, as it is usually unknown what part of the second protein will be in close proximity to the first protein-cross-linker complex. One type of functional group that has been used much more extensively as of recent times is a photoactivatable group. Once excited into a triplet state by UV light, these will generally either insert or add to carbon-hydrogen or nitrogen-hydrogen bonds fairly non-specifically in any protein in close proximity. Without a photoactivatable cross-linker, functional groups are limited to those that may be specific to, for instance, amine or sulfhydryl groups, which may not be in close proximity. This would significantly decrease the likelihood of a successful cross-link.

Therefore, the cross-linkers that have most successfully resulted in the mapping of protein interfaces have been those that are bifunctional, with a cysteine-reactive group at one end and a photoactivatable group on its other end.

The spacer-arm length can also affect the success of a cross-linker in a specific protein interaction. The longer the spacer-arm length, the further away the second protein can be for a successful cross-link to occur. Despite the non-selectivity of photoactivatable functional groups, some have been reported to have a slight preference for certain amino acid side-chains. A longer spacer-arm allows a greater possibility of reaching one of these amino acids in the second protein. However, this comes at the cost of a greater possibility of creating non-specific cross-links that may not be physiologically relevant. In some cross-linking experiments, cross-linkers that have the same functional groups with different lengths of spacer arms are used to ensure that a cross-linker of optimal length is utilized [224].

For use in the BSP-collagen interaction, it is important to select a cysteine-reactive cross-linker with a photoactivatable group at its other end. Cysteine is an especially attractive target in the BSP-collagen system, as there are no cysteines within BSP. Because we know the collagen-binding region of BSP, cysteine residues can be incorporated into this region via site-directed mutagenesis. Multiple sites can be mutated to cysteine (maintaining only one cysteine per BSP peptide) in order to maximize the chances of a successful cross-link occurring.

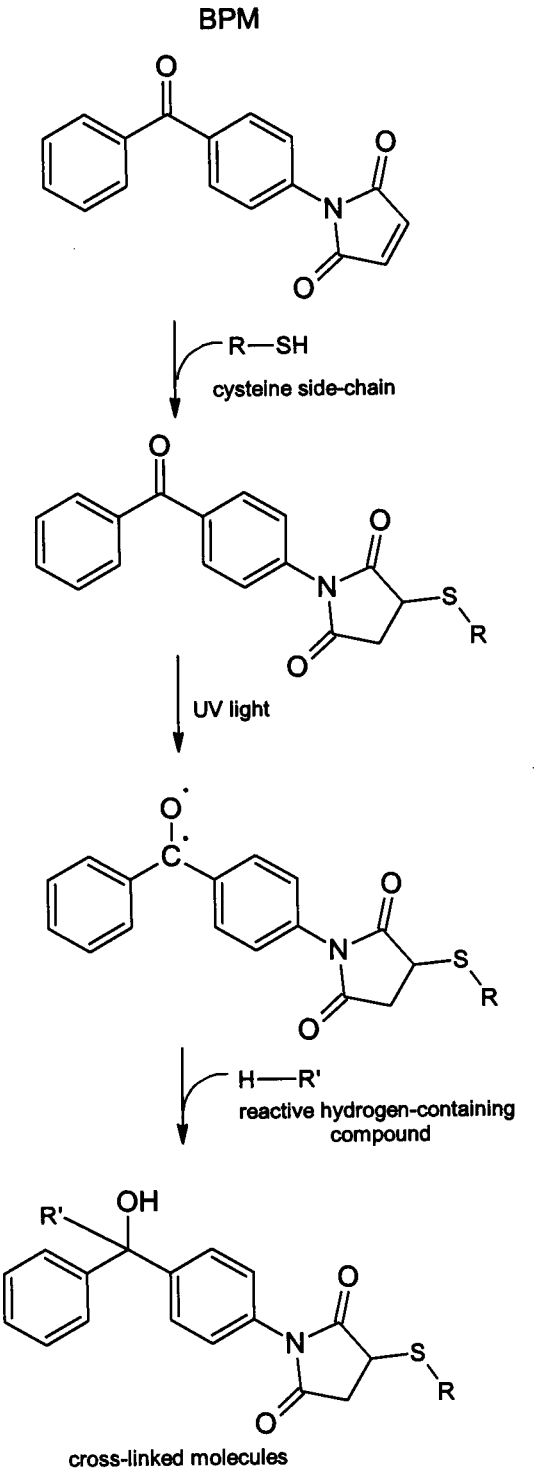
Out of the several hundred cross-linking molecules commercially available for use with proteins, there are less than 15 readily available that are heterobifunctional and contain sulfhydryl reactive and photoactivatable functional groups. Three of varying

lengths and functional groups have been chosen to be used in these studies: benzophenone-4-maleimide (BPM), ρ-azidophenacyl bromide (APB), and N-[4-(ρ-azidosalicylamido)butyl]-3'-(2'-pyridyldithio) propionamide (APDP). BPM and APB were chosen due to extensive use and characterization by others in the literature. APDP, created more recently, was chosen based on its inherent advantages over other cross-linking molecules (see below) and an increased spacer-arm length. Also, APDP has reportedly been used to successfully cross-link another protein to type I collagen [225]. Each of the three cross-linking reagents will be described in detail.

BPM

Benzophenone-4-maleimide is a heterobifunctional cross-linking reagent containing a maleimide moiety at one end, and a benzophenone at the other, spanning between 0.86 and 1.14 nm. The double bond of the maleimide group undergoes an alkylation reaction with the thiolate ion of cysteine by addition to the double bond (Figure 2.2). More specifically, one of the carbons next to the double bond undergoes nucleophilic attack by the cysteine side-chain to generate this cross-link. It is considered specific for thiols in the pH range of 6.5 – 7.5. Benzophenone groups are excited into a triplet-state ketone that is highly reactive upon UV exposure. The excited electron of the benzophenone can then insert into carbon-hydrogen bonds and other active groups yielding covalent linkages. One significant advantage to benzophenone groups over other photoactive groups is the fact that its decay does not result in an inactive product. Therefore, the benzophenone can be excited multiple times, resulting in higher cross-link yields over other photoactivatable groups. Benzophenone-4-maleimide is very insoluble in aqueous solutions; however, this does not usually pose a problem, as cysteine residues

Figure 2.2 Benzophenone-4-maleimide (BPM) and its mechanisms of action.


BPM
R—SH
cysteine side-chain
UV light
H—R'
reactive hydrogen-containing compound
cross-linked molecules

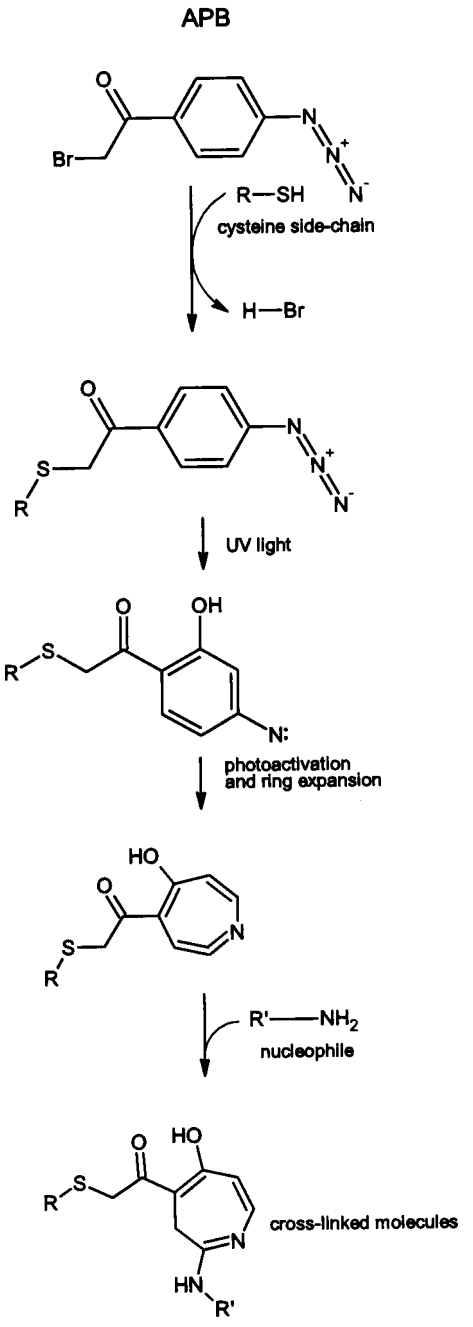
are quite rare and therefore the compound only needs to be present in very small quantities.

Benzophenone-4-maleimide has been used extensively in the literature for a range of proteins and their interacting partners. It has been used to create intermolecular cross-links to obtain information of protein structure, as well as intramolecular cross-links to map binding sites. It was first used to map protein-protein interactions in 1985 when Tao and colleagues used BPM to study the conformation of the C-terminal region of actin [226]. Guo and colleagues have analyzed the structure of rod photoreceptor cGMP phosphodiesterase by cross-linking to better understand how its inhibitory subunit functions [227]. Intramolecular cross-links using BPM have been used to map binding sites between DNA mismatch repair proteins MutL and MutH [228], between subunits of the *E. coli* ATP synthase complex [229], troponin I and other thin filament proteins [230], and many other complexes.

APB

ρ-azidophenacyl bromide (APB) consists of a bromoacetyl group on one end, where the bromine ion is replaced by an attacking nucleophile. The primary reaction is with cysteine residues, however it is not totally specific. Other potential nucleophiles include the imidazolyl nitrogens of histidines, the thioether of methionine, and primary amines [223]. The only reaction that results in a definitive product is the alkylation of cysteine residues, creating a carboxymethylcysteinyl derivative [231]. The photoactivatable group on APB is a phenyl azide group. Upon excitation by UV light, the phenyl azide forms an intermediate nitrene that reacts immediately with the surrounding environment [232] (Figure 2.3). Nitrenes react nonspecifically either through addition via

Figure 2.3 ρ-azidophenacyl bromide (APB) and its mechanisms of action.


APB
Br
R—SH
cysteine side-chain
H—Br
UV light
OH
N:
photoactivation
and ring expansion
HO
R'——NH2
nucleophile
HO
cross-linked molecules
HN
R'

double bonds or insertions into carbon-hydrogen or nitrogen-hydrogen bonds. Another possible mechanism by which nitrenes react is by ring expansion, creating a nucleophile-reactive dehydroazepine. This will primarily react with nucleophiles such as amines as opposed to inserting into carbon-hydrogen or nitrogen-hydrogen bonds. The length of this cross-linker is 0.9 nm [233]. The disadvantage to this type of reactive group is that the molecule can only be excited once, and subsequently reacts or decays into an inactive product, limiting the chances of a successful cross-link.

APB was first synthesized and used in 1975 to analyze its inhibitory effects on the cysteine-containing active site of glyceraldehyde-3-phosphate dehydrogenase (GPDH) [234]. Since then, it has been used mainly for mapping interactions between proteins and nucleic acids, such as those between ribosomal proteins and RNA [235], between histone proteins and DNA [236] and between domains of HIV reverse transcriptase and DNA [237]. In these instances, a modified nucleotide containing a thiol group is incorporated into the DNA or RNA, and is highly reactive with the halogen group of APB. When irradiated with UV light, the nucleic acid will covalently attach to nearby proteins through APB [238]. However this cross-linker has also been more recently used to map protein interactions such as those between yeast ATPase subunits [239], and to study the interaction between yeast cytochrome C and cytochrome C peroxidase [240], for instance.

APDP

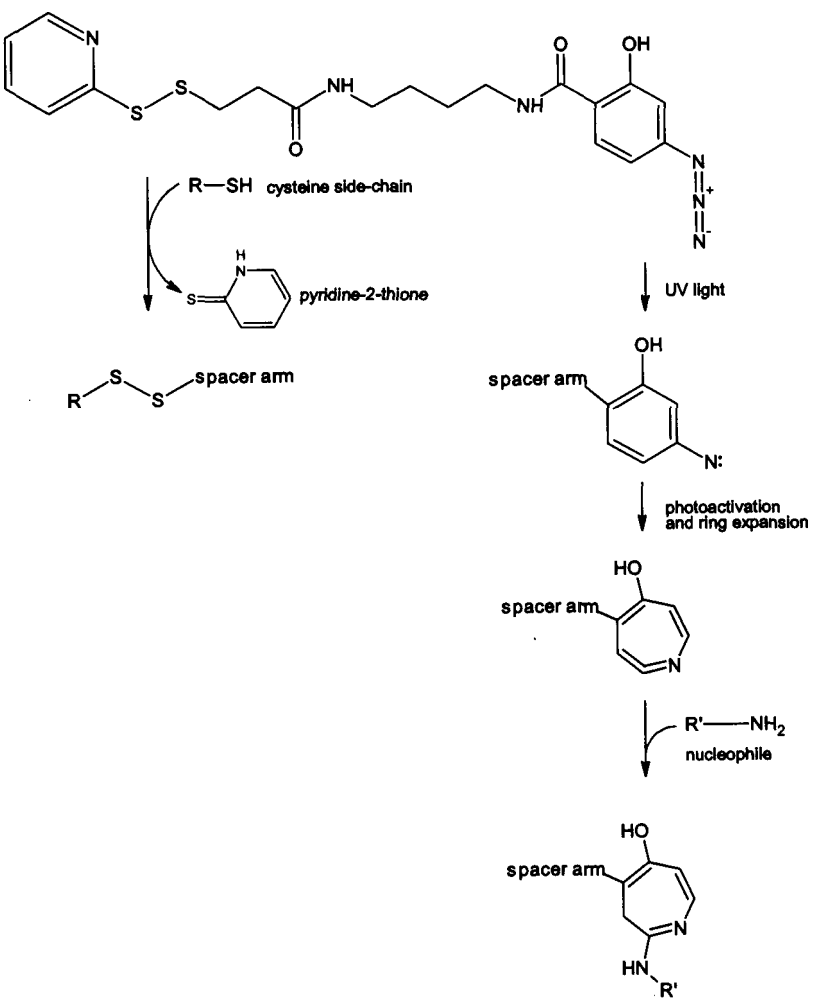
N-[4-(ρ-azidosalicylamido)butyl]-3'-(2'-pyridyldithio) propionamide is quite different from the other cross-linkers discussed thus far. At one end, it contains a pyridyl disulfide that will undergo disulfide exchange, where pyridine-2-thione functions as a

leaving group, and the rest of the molecule forms a disulfide bond with a cysteine residue. Pyridine-2-thione is unreactive, and therefore disulfide formation is confined to the rest of the cross-linking molecule, increasing efficiency. At the other end of APDP is a phenyl azide similar to that of APB. This molecule has a much longer spacer arm than the previous two cross-linkers, at roughly 2.1 nm [223] (Figure 2.4). There are some distinct characteristics of this cross-linker that can be exploited experimentally. Firstly, after a cross-link between two proteins has been created, the disulfide of the cross-linker can be broken. This labels the second protein with a thiol group that would not normally be present, but can be detected with thiol-reactive compounds [225]. Another distinct characteristic of this cross-linker is that the phenyl ring can be radioiodinated, and therefore detection methods for cross-links can be much more sensitive. A label-transfer protocol such as the one mentioned above can be used to attach a radioisotope to the second protein involved in the cross-link. Because of the ability to transfer thiols or a radioisotope to interacting proteins, this cross-linker is often commercially categorized as a "label transfer" reagent. This type of method could be used to probe complex protein mixtures to detect which proteins interact with a protein of interest. The limitation to this cross-linker is that the presence of reducing agents used to initially reduce cysteine residues will inactivate the pyridyl disulfide of the cross-linker. Therefore, care must be taken to ensure that all reducing agents are removed from solution prior to the addition of cross-linker. This may require the addition of a chelating agent such as 1 – 10 mM EDTA to prevent metal-catalyzed oxidation of cysteine residues.

APDP was developed more recently than the other cross-linkers discussed. Its first recorded use was in 1992 to characterize the interaction of *E.coli* ribosomal proteins

Figure 2.4 N-[4-(ρ-azidosalicylamido)butyl]-3'-(2'-pyridyldithio) propionamide (APDP) and its mechanisms of action.

APDP


R—SH cysteine side-chain
pyridine-2-thione
UV light
spacer arm
photoactivation
and ring expansion
R'——NH2
nucleophile

L7/L12 with the 50 S subunit [241]. This was accomplished through radiolabelling and 2D gel electrophoresis. Because of APDP's advantages over other cross-linkers such as the abilities to be radioactively labeled, and to introduce a reactive thiol onto interacting proteins, it is becoming increasingly popular for site-directed photocrosslinking experiments. It has been used to study protein-protein interactions in the *E.coli* multicomponent preprotein translocase complex [242], bacterial surface layer protein subunits that form the cell wall of archaea and many eubacteria [243], and mapping of *E.coli* ribosomal subunits [244]. Yasui and coworkers utilized APDP to conjugate pigment epithelium-derived factor (PEDF) to type I collagen, and cleaved the cross-link using a reducing agent, thereby introducing a thiol group to the interacting region of collagen [225]. The introduced thiol was labeled with a fluorescent probe, and the α-1 and α-2 collagen monomers separated by 2D gel electrophoresis. It was determined through this experiment that PEDF interacted with the α2 chain of collagen [225]. Although the region of collagen involved in the interaction wasn't resolved to the amino acid level, the previous use of this cross-linker in mapping a protein-collagen interaction is promising for the purposes of this thesis.

It is hoped that between the three cross-linking molecules discussed, a covalent linkage between BSP and collagen can be created at the physiological interface. This protein complex could then be studied through digestion and gel electrophoresis to obtain low-resolution data regarding binding locations, or through mass spectrometry to determine the exact amino acid sequences involved in the interaction. Extensive reviews have been written by Sinz and others regarding the analysis of cross-linked complexes by mass spectrometry, initially using the relatively low-resolution MALDI-TOF MS, and

subsequently high-resolution ESI-FTICR-MS along with various software analysis techniques [218, 224, 245, 246]. Once cross-links are characterized by mass spectrometry, it becomes possible to use molecular modeling to simulate the interaction, and obtain information about the mechanisms of interaction down to the atomistic level [247].

2.1.3 Mapping the BSP-binding region on type I collagen

In 2002, the lab of James D. San Antonio mapped most of the known ligand-binding sites on a two-dimensional model of type I collagen [248]. However, without the context of the three-dimensional collagen structure, it was hard to extrapolate these binding sites to the physiological level. With the more recent resolution of the microfibrillar structure of collagen [49], San Antonio and colleagues discovered that there appears to be a region of the 3D collagen molecule responsible for binding cells through integrins and other cell-surface proteins, and another region, overlapping the hole zone, responsible for binding extracellular matrix proteins [249]. A cross-link obtained between BSP and collagen could be mapped onto this 3D model to determine if BSP does, in fact, bind to collagen adjacent to the hole zone. Further studies could then be conducted, such as molecular dynamics, to determine how this interaction promotes the nucleation of hydroxyapatite.

The work conducted in our lab thus far towards locating the BSP-binding region of type I collagen will be presented in this chapter. This includes the creation, expression, and purification of various single-cysteine mutants of the BSP (1-100) peptide, their conjugation to heterobifunctional cross-linking reagents, and the subsequent formation of cross-links between BSP and collagen.

2.2 Materials and methods

2.2.1 Materials

Dpn I, Platinum *Pfx* polymerase, $MgSO_4$, and fluorescein-5-maleimide were purchased from Invitrogen/Molecular Probes (Carlsbad, CA). Acrylamide/bis-acrylamide, sodium dodecyl sulfate, and Coomassie blue were purchased from Bio-Rad Laboratories (Hercules, CA). β-mercaptoethanol, glycine, Tween®-20, and APDP were purchased from Thermo Fisher Scientific (Rockford, IL). Methanol and isopropanol were purchased from VWR International (West Chester, PA). Urea was purchased from BioShop Canada (Burlington, ON). Medical grade N_2 was purchased from Praxair Inc. (Mississauga, ON). BL21 (DE3) cells were purchased from EMD Biosciences/Novagen (Gibbstown, NJ), while XL-1 cells were purchased from Stratagene (Cedar Creek, TX). All other reagents were purchased from Sigma-Aldrich (St. Louis, MO) unless otherwise specified. The type I collagen used in these studies was extracted from rat tail tendons as previously described [250, 251].

2.2.2 Methods

2.2.2.1 Site-directed mutagenesis

A pET28 plasmid containing the rat BSP (1-100) sequence with a hexa-His tag, and conferring kanamycin resistance was already present in our laboratory as described [129]. However, upon sequencing this particular construct, a base-pair mutation was found that would create a glutamic acid-to-lysine substitution at amino acid 54 (E54K). Therefore, site-directed mutagenesis was first required to correct this improper mutation, and subsequently to create each of the five selected single-cysteine mutants using the corrected BSP (1-100) sequence as a template. The location of the BSP (1-100) peptide

with respect to the functional domains of BSP, as well as the specific residues selected for mutation, are outlined in Figure 2.5.

The Quickchange XL site-directed mutagenesis kit (Stratagene) was used following the manufacturer's instructions. Primers for each PCR reaction were selected based on a length between 25 and 45 bases, a melting temperature greater than or equal to 78°C, a GC content of about 40%, and the desired mutation near the centre of the primer. The most highly used codon requiring only one base substitution was selected for each mutation. The selected primers were purchased from Sigma Aldrich and are listed in Table 2.1.

Through trial and error, an optimum PCR mixture was found to consist of: 36.4 µl distilled autoclaved water, 5 µl Invitrogen 10x Pfx buffer, 2.5 µl Invitrogen Pfx enhancer solution, 1.5 µl 50 mM $MgSO_4$, 1 µl 10 mM dNTP solution, 1 µl of 0.1 µg/µl of each primer, 1 µl template DNA (0.2 µg), and 0.6 µl Pfx polymerase (2.5 U/ µl).

Also through trial and error, the optimum temperature cycle was found to be 18 cycles of: 95°C melting temperature for 30 seconds, 52°C annealing temperature for 30 seconds, and 68°C elongation temperature for 6 min. This cycle was preceded by a one-min incubation at 95°C to activate the enzyme and ended with an extra four minutes of 68°C elongation, followed by 4°C storage of product. The thermocycler used was the Applied Biosystems GeneAmp PCR System 2700. These data are summarized in Table 2.2. Parental DNA was digested with *Dpn* I and the PCR product analyzed by agarose gel electrophoresis with ethidium bromide staining.

The PCR product for each sample was transformed into XL-1 cells, the cells plated on LB-agar containing kanamycin, and the dishes incubated overnight at 37°C. A

Table 2.1 Primers used for site-directed mutagenesis PCR.

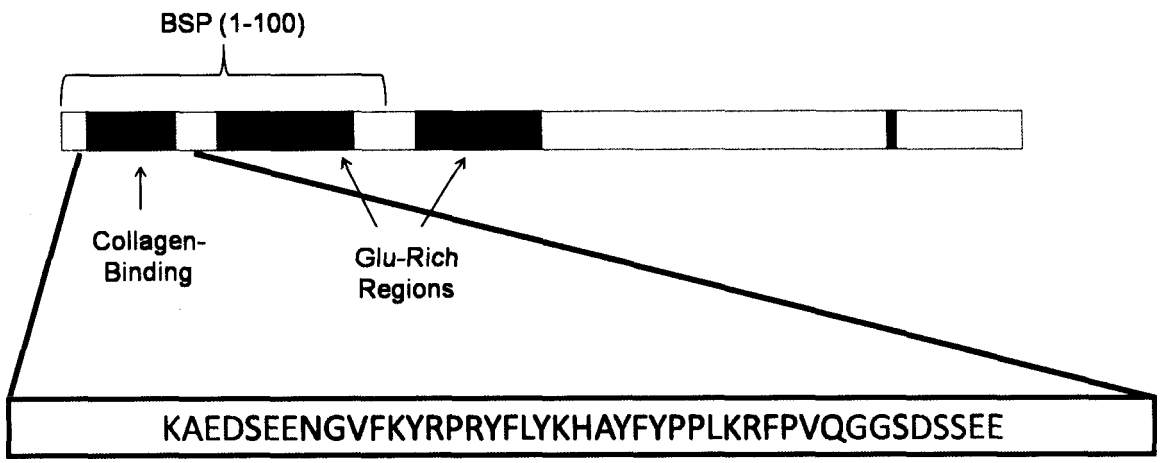
Primers were designed based on the following criteria: 1) a length between 25 and 45 bases, 2) a melting temperature above 77°C, 3) a GC content of approximately 40%, and 4) the desired mutation near the centre of the primer. Each PCR reaction only required a 1-base substitution to incorporate a high-usage *E.coli* cysteine codon.

Primer	Sequence (5'-3')
5' K54E	CAGCGACTCTTCGGAAGAAAATG(**GAG**)ATGGC
3' K54E	GCCAT(**CTC**)CATTTTCTTCCGAAGAGTCGCTG
5' S16C	CATCAAAGCAGAGGAT(**TGT**)GAAGAAAACGGGGTC
3' S16C	GACCCCGTTTTCTTC(**ACA**)ATCCTCTGCTTTGATG
5' Y24C	GTCTTTAAG(**TGC**)CGGCCACGCTACTTTCTTTATAAGCACGC
3' Y24C	GCGTGCTTATAAAGAAAGTAGCGTGGCCG(**GCA**)CTTAAAGAC
5' Y31C	CCACGCTACTTTCTT(**TGT**)AAGCACGCCTACTTTTATCCTCCTCTG
3' Y31C	CAGAGGAGGATAAAAGTAGGCGTGCTT(**ACA**)AAGAAAGTAGCGTGG
5' F36C	CTTTATAAGCACGCCTAC(**TGT**)TATCCTCCTCTGAAACGGTTTCC
3' F36C	GGAAACCGTTTCAGAGGAGGATA(**ACA**)GTAGGCGTGCTTATAAAG
5' S49C	CAGGGAGGC(**TGC**)GACTCTTCGGAAGAAAATG
3' S49C	CATTTTCTTCCGAAGAGTC(**GCA**)GCCTCCCTG

Note: Cysteine codons are in brackets with the mutated bases underlined
Codon usage: Cys: TGT: 47.6%, TGC: 52.4%, Glu: GAG: 27.8%

Figure 2.5 The BSP (1-100) peptide and sites of single-cysteine mutagenesis.

This schematic depicts the location of the BSP (1-100) peptide within the full-length BSP protein and its functional domains. The amino acids selected for mutation within and surrounding the collagen-binding region are highlighted.



■ Residues within the collagen-binding domain

■ Residues mutated to cysteine

Note: Each BSP mutant will only have **one** cysteine residue

colony was selected to inoculate a 3-ml starter culture over a period of 8-12 h at 37°C with shaking. The starter culture was then added to 247 ml LB broth and shaken at 37°C until an OD of ~ 0.8 was reached. The cells were then pelleted, lysed, and plasmids purified by midi-prep using the Qiagen Midi-prep kit. Outsourced sequencing using the T7 promoter and T7 terminator sequences was used to assess the success of the PCR reactions (Robarts Research Institute).

2.2.2.2 Protein expression and purification

The 6 rBSP (1-100) constructs were transformed into BL21 (DE3) cells and each grown in 6 L LB broth containing 15 μg/ml kanamycin. Once the cells reached an optical density of 0.6 to 0.8, isopropyl-β-D-thiogalactopyranoside (IPTG) was added to a concentration of 2 mM to induce expression. Expression was carried out for 4 h at 37°C with shaking. The bacterial cells were sonicated in denaturing His-column binding buffer (5 mM imidazole, 0.5 M NaCl, 0.02 Tris-HCl, 6 M urea, pH 7.9). The protein extract was then passed through a Pharmacia KK-16 column containing a nickel-binding resin that had previously been charged with 50 mM $NiSO_4$. The column was then washed with 10 column volumes of binding buffer and proteins were eluted by the addition of a 0.5-M imidazole elution buffer. Elution fractions were analyzed by 12.5% polyacrylamide gels using the Phastgel system (Amersham Biosciences). Staining of the gels was carried out using Stains-All and silver nitrate as described previously [252]. Fractions containing the BSP (1-100) peptide, indicated by cyan-positive Stains-All staining at approximately 18 kDa, were pooled for further FPLC purification. The resulting protein solution was diluted 3 to 5-fold with Mono-Q buffer A (50 mM Tris-HCl, 7 M urea, pH 7.4) and added to the FPLC sample loop for addition to a Q Sepharose Fast Flow 1 x 10 cm column

(FastQ). The proteins were eluted with an increasing NaCl concentration and once again analyzed using the Phastgel system followed by Stains-All and silver staining. Fractions containing the desired peptide were pooled and concentrated 10 – 15x using a pressurized Amicon ultrafiltration apparatus with a YM-3 (3-kDa cutoff) membrane (Millipore). The concentrated protein solution was then subjected to further FPLC with a Superdex 200PG column (1.6 cm x 60 cm) (Amersham Biosciences) in 4 M urea, 200 mM NaCl, pH 7.4. The fractions containing the purified peptide were pooled and dialyzed against 10 mM ammonium bicarbonate for 72 h at 4°C, using a 3-kDa cutoff Spectra/Por dialysis membrane (Spectrum Laboratories), aliquoted into microfuge tubes, and lyophilized overnight. Samples were then subjected to amino acid analysis, BCA assays, and MALDI-TOF MS to determine relative purity and quantity.

2.2.2.3 Fluorescein maleimide reaction with single-cysteine mutants

Mutant proteins were dissolved in PBS, 1 mM EDTA, pH 7.5 at a concentration of 1.136 mg/ml. A 3-µl aliquot of a 9-mM TCEP solution in 50 mM TRIS pH 7.5 was added to 26.4 µl of protein solution, and the mixture incubated at room temperature for 30 min. Then, 0.6 µl 50 mM fluorescein maleimide in DMF was added and the reaction incubated at room temperature for 90 min while stirring in the dark. Final concentrations of each compound were 82 µM protein, 0.9 mM TCEP, and 1mM fluorescein maleimide, resulting in a thiol:reducer:dye ratio of 1:11:12.2. This ratio maintains the optimum 9:10 reducer:dye ratio as determined by Tyagarajan and coworkers [253]. The solution was quenched by the addition of 2 µl 0.2 M β-mercaptoethanol and incubated for 10 min at room temperature in the dark. An aliquot of 15 µl of this solution was mixed with 5 µl 4x SDS buffer lacking reducing agent and heated to 95°C for 5 min prior to loading on the

gel. Gel electrophoresis was carried out using the Laemmli method [254] with a 12.5% polyacrylamide gel. Samples were run at 80 V until they entered the separating gel, at which time voltage was increased to 180V. The gel was run until the bromophenol blue from the SDS sample-buffer reached the bottom of the gel. The gel was analyzed within its glass plates directly on a UV-transilluminator at 362 nm to excite the fluorescein, and an image taken. Subsequently, the gel was stained with Stains-All and silver to visualize all protein present.

2.2.2.4 Qualitative determination of reaction between BPM or APB and BSP (1-100) single-cysteine mutants using fluorescein maleimide

Fluorescein maleimide was used to determine if the cross-linkers benzophenone-4-maleimide (BPM) and ρ-azidophenacyl bromide (APB) were being successfully conjugated to BSP peptides. All BSP (1-100) single-cysteine mutants, as well as wild-type BSP (1-100), were first reacted with BPM or APB following the same protocol used to conjugate fluorescein maleimide to the proteins, replacing FM with either of the above cross-linkers. APB was dissolved in methanol instead of DMF. Controls in which only DMF or methanol was added were performed. Subsequently, all samples (including controls) were diluted with 300 µl PBS, 1 mM EDTA, 0.9 mM TCEP, pH 7.5, and added to a YM-3 microfuge filter (Microcon), in which a buffer exchange was performed to eliminate excess cross-linker molecules. The samples were subsequently reacted with FM as done previously (including samples that had not been first reacted with cross-linkers) and the results analyzed by SDS-PAGE followed by UV illumination of the gel. This qualitative measurement of the reaction between cross-linkers and the BSP (1-100)

peptides was repeated in the presence of 2x PBS, which is the solution used to conjugate cross-linkers to the BSP peptides for cross-linking experiments.

2.2.2.5 Characterization of the BSP-BSP interaction

To characterize the BSP-BSP interaction, BSP aggregation was measured through cross-linking experiments in the presence of different NaCl concentrations. BSP (1-100) Y24C was again conjugated to BPM as previously described in 2x PBS, pH 7.4. The mixture was then diluted to twice its volume with either H_2O or an NaCl solution, resulting in final NaCl concentrations of 150 mM or 500 mM, respectively. The solutions were incubated in the dark at room temperature with stirring for a period of 2 h, followed by irradiation with UV light (362 nm) for a period of 5 min. The reactions were then analyzed by SDS-PAGE in a 12.5% polyacrylamide gel with Stains-All and silver nitrate staining.

2.2.2.6 Type I collagen transfer assay

The following amounts of collagen in 30µl volume were loaded onto polyacrylamide gels of 12.5% or 6%: 20 µg, 15 µg, 10 µg, 5 µg, 4 µg, 3 µg, 2 µg, 1 µg, 0.5 µg, 0.2 µg. Gels were run using the Laemmli method at 80 V until the protein reached the separating gel, and then 120 V until the dye front reached the bottom of the gel. Using the Mini-protean 3 system (Bio-rad), the proteins were transferred to Hybond-P PVDF membrane (Amersham Biosciences) at either 30 or 100 V for time periods ranging from 1 h to 14 h to determine optimum transfer efficiency. The transfer buffers used were a Tris-glycine buffer with 40% methanol (pH 8.3) and a bicarbonate buffer (3 mM Na_2CO_3 and 10 mM $NaHCO_3$, pH 9.9; 20% (v/v) methanol) formulated by Dunn for use with basic proteins [255]. Following the transfer, the PVDF membrane was incubated in Ponceau S

stain (0.01% in 5% acetic acid) for 10 min. The membrane was destained by briefly rinsing in Milli-Q water for 2-3 washes. The gel was stained overnight in Coomassie blue (0.05% in 10% acetic acid, 40% methanol), followed by destaining in 10% acetic acid, 40% methanol until bands were clearly visible or the gel was completely destained. The gel and PVDF membrane were then imaged for analysis.

2.2.2.7 Cross-linking BSP (1-100) peptides to type I collagen

The method previously described to conjugate cross-linkers to the BSP (1-100) mutants was scaled up to obtain a final 100-µl solution of 1 mg/ml BSP, 2x PBS, 10 mM EDTA, 0.9 mM TCEP. After a 30-min incubation at room temperature, TCEP was removed using a Microcon YM-3 centrifugal-filter through buffer exchange. 4 µl of 50 mM APDP or BPM in DMF, or 50 mM APB in methanol was then added and the mixture incubated in the dark for 60 - 120 min. The solution was placed in the cap of a microfuge tube with a stir bar, and 100 µl of 1 mg/ml collagen type I dissolved in 5 mM acetic acid was added, resulting in a 200-µl solution at neutral pH. The mixture was incubated with slight stirring for 3 h in the dark at room temperature, followed by inversion over a TM-36 UV transillumator (VWR) for irradiation at 362 nm for 5 to 20 min. The mixture was then spun down for 2 min at 13,000 x g, and supernatant removed. Non-reducing SDS-PAGE sample buffer was added, and the reactions analyzed by SDS-PAGE in a 6% gel, followed by Western blot. In subsequent reactions, β-mercaptoethanol was added to the cross-linking reactions prior to SDS-PAGE to characterize the nature of the observed potential cross-link. Also, reactions were performed without the addition of either collagen or APDP.

After separating cross-linking reactions by SDS-PAGE, the proteins were transferred to Hybond-P PVDF membrane (Amersham Biosciences) for 3 h at 100 V in the Mini-Protean 3 system (Biorad), which was found to be the optimal condition for the transfer of collagen. Once the protein had been transferred to the membrane, the membrane was stored in TBST (25 mM Tris, 0.15 M NaCl, pH 7.6, 0.05% Tween®-20) overnight. To detect the presence of BSP, an anti-His antibody (Qiagen) was used, as well an N-terminal polyclonal antibody to BSP (courtesy of the late J. Sodek, University of Toronto). However, the Supersignal West HisProbe kit (Pierce) was found to have the highest sensitivity and specificity for the BSP (1-100) peptides. The membrane was blocked for 1 h at room temperature using 25 mg/ml BSA in TBST, followed by two 10-min washes. HisProb-HRP (4 mg/ml; Pierce) was diluted 1:5000 into TBST and incubated with the membrane for 1 h at room temperature. This was followed by 4 washes for 10 min each. 7.5 ml of SuperSignal West Pico Substrate Working Solution (Pierce) was added to the membrane for 5 min prior to analysis. The membrane was subsequently analyzed using the ChemiImager 5500 gel doc system and corresponding software (Alpha Innotech).

2.2.2.8 Cyanogen bromide cleavage of collagen

Cyanogen bromide cleavage of collagen was carried out by an adaptation of methods previously published [256, 257]. Briefly, 1 mg of lyophilized type I collagen was dissolved in 100 µl 70% formic acid in a 2-ml glass vial, and the solution saturated with N_2. Approximately 1 mg CNBr was then added and the solution gently agitated until dissolved. The solution was again saturated with N_2 and incubated for 4 h at 30°C. To quench the reaction, 100 µl 0.1 M acetic acid was added and the solution evaporated

under N_2. Then, the cleaved collagen was washed twice by adding 100 µl H_2O and evaporating under N_2. The fragments were then resuspended in PBS pH 7.5, transferred to a microfuge tube, and centrifuged for 2 min at 13,000 x g. The supernatant was used for subsequent analysis. Alternatively, this protocol was adapted to cleave the collagen in cross-linking reactions by diluting the cross-linking solution with high purity formic acid (to achieve the required 70% formic acid content) instead of dissolving lyophilized collagen directly.

2.3 Results

2.3.1 Site-directed mutagenesis

For cross-linking experiments, it was decided that the BSP (1-100) peptide would be used as it contains the collagen-binding region of BSP and expresses in much higher quantities than the full-length protein. Site-directed mutagenesis PCR was used to alter an existing construct in order to create BSP (1-100) and 5 single-cysteine mutants of BSP (1-100). The construct contained BSP (1-100) with a hexa-histidine tag within the lac operon, and conferred kanamycin resistance. The existing construct was found to contain a glutamic acid-to-lysine mutation at amino acid 54. This was first mutated back to glutamic acid to generate a template from which to generate single-cysteine mutants. Residues mutated to cysteine were serine 16, tyrosine 24, tyrosine 31, phenylalanine 36, or serine 49 (Figure 2.5). The original and mutated peptides will be referred to as wild-type, S16C, Y24C, Y31C, F36C, and S49C, respectively. Three of the mutations are therefore within the collagen-binding sequence encompassing residues 19-46 of rat BSP. The PCR reaction to incorporate the cysteine substitutions proved to be highly temperamental, and therefore a large variety of variables were experimented with to develop a consistently successful PCR reaction. These included altering the relative concentrations of each component of the reaction, as well as the temperature cycle followed by the thermocycler. The optimal conditions and primers used are detailed in Table 2.2.

For each of the six site-directed mutagenesis procedures, including the correction in the wild-type BSP (1-100) construct, a small aliquot of the PCR product was separated by agarose gel electrophoresis in the presence of ethidium bromide. A strong band at

Table 2.2 Optimized PCR reaction variables.

Through a process of trial-and-error, the optimal quantities of each reaction component (A), as well as the optimal temperature cycle (B), were elucidated. These conditions were replicated for each of the BSP (1-100) mutagenesis PCR reactions.

A)

Ingredient	Volume (μL)
H_20	36.4
10x Pfx buffer[1]	5
10x Pfx enhancer[1]	2.5
50 mM $MgSO_4$	1.5
10 mM dNTP mixture	1
5' primer[2] (0.1 μg/μL)	1
3' primer[2] (0.1 μg/μL)	1
100 – 200 ng plasmid	1
Platinum Pfx[1]	0.6

[1] Pfx buffer, enhancer, and polymerase are proprietary solutions from Invitrogen
[2] Primers varied depending on the reaction, and are listed in the Materials and Methods section

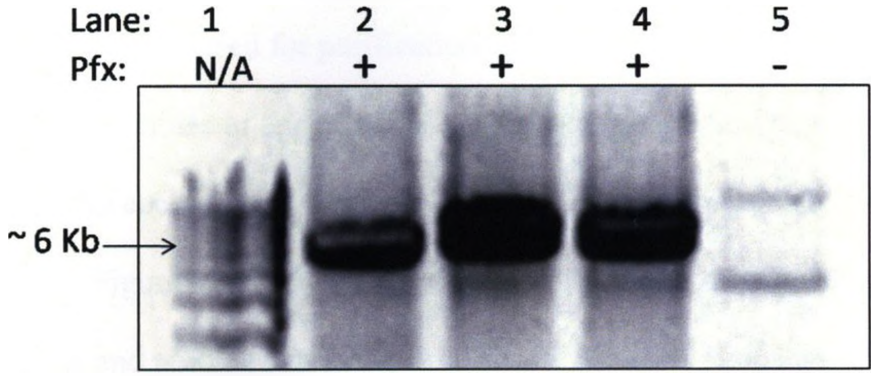
B)

Cycle stage	Temperature (°C)	Time (min)
Initiation	95	2
Melting	95	0.5
Annealing	52	0.5
Elongation	68	6
Extended elongation	68	3
Final hold	4	∞

Note: Within the dotted lines is the cycle that was repeated 18x

Figure 2.6 Site-directed mutagenesis PCR products analyzed by agarose gel electrophoresis.

A 2% agarose gel containing ethidium bromide was used to analyze PCR products. Band intensities corresponding to the BSP (1-100) construct (~6 Kb) (lanes 2-4) were compared with a control PCR reaction in which no polymerase had been added (lane 5). The gel was run at 80 V for a period of 45 minutes and imaged on a UV transilluminator for analysis.


Lane: 1 2 3 4 5
Pfx: N/A + + + -
~ 6 Kb

approximately 6 kb, corresponding to the desired construct, represented a successful PCR reaction (Figure 2.6). The solutions containing successful PCR products were digested with *Dpn* I to ensure that parental DNA had been degraded. XL-1 cells were then transformed with the amplified construct and a midi-prep performed to obtain large amounts of the purified plasmid.

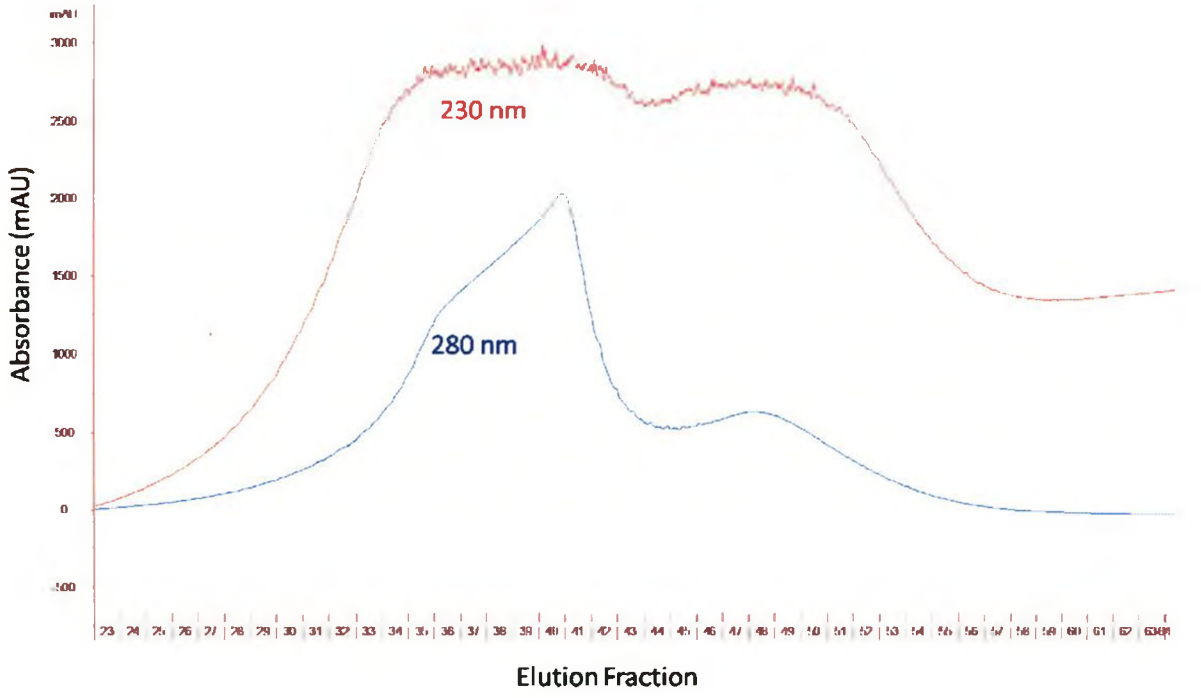
2.3.2 Expression and purification of BSP (1-100) peptides

Once the successful PCR product was sequenced to determine that the mutations had been incorporated, the 6 proteins were expressed in BL21 (DE3) cells and purified by FPLC. Three columns were used for purification: nickel-affinity (based on the hexa-histidine tag on BSP peptides in our construct), cation-exchange, and size-exclusion. Elution fractions from each column were analyzed spectrophotometrically for the presence of protein (Figure 2.7) and by SDS-PAGE with Stains-All and silver staining to identify the proteins and purity of each fraction. Fractions containing significant amounts of the protein of interest and lacking significant levels of contaminating proteins were pooled for the next FPLC purification step. Figure 2.8 represents the purified peptides separated by SDS-PAGE, with the gel stained by Stains-All and silver nitrate. Although each BSP (1-100) peptide is approximately 12.2 kDa, when analyzed by SDS-PAGE the protein migrates with an apparent mass of 18 kDa. This slow migration is characteristic of full-length BSP as well. The purity and identity of each peptide was further verified through MALDI-TOF MS (data not shown). Once purified, the proteins were dialyzed against ammonium bicarbonate, lyophilized, and subjected to amino acid analysis. The amino acid analysis results are summarized in Table 2.3, where the relative abundance of each amino acid is expressed as residues per 1000 amino acids. Certain amino acids, such

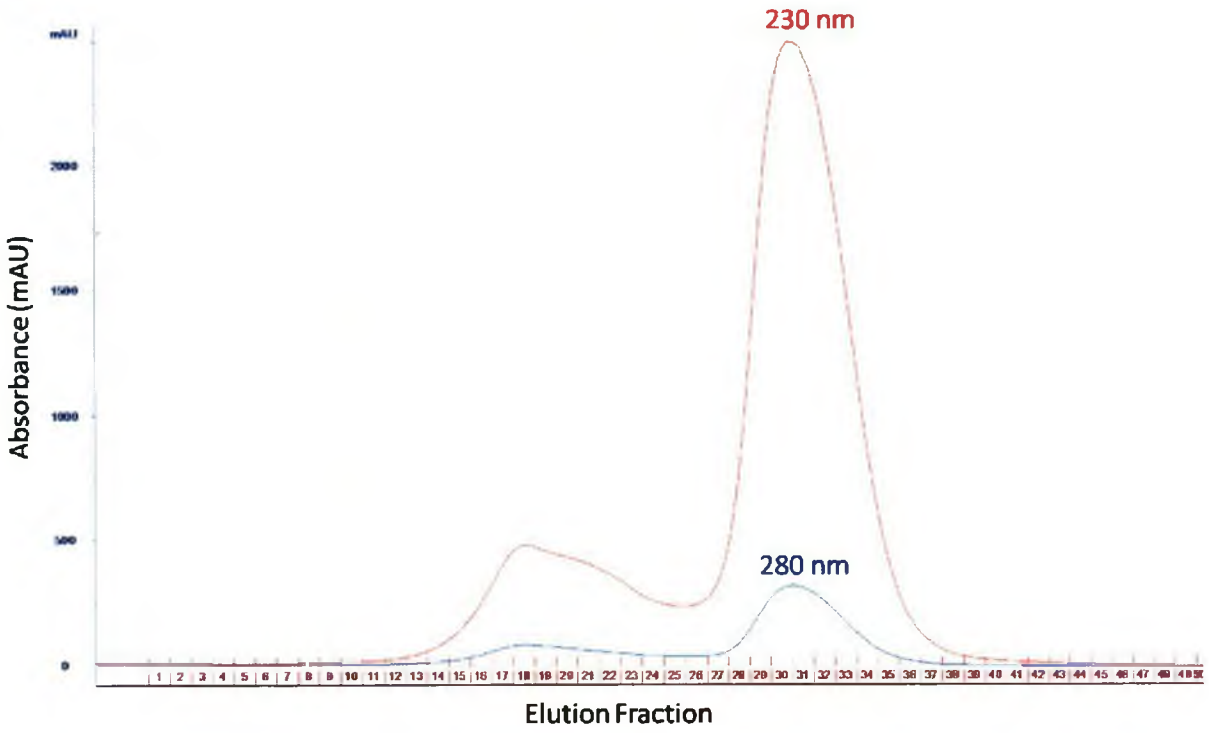
Figure 2.7 FPLC purification chromatograms.

BSP (1-100) Y24C was purified by nickel-affinity **(A)**, FastQ Sepharose ion-exchange **(B)**, and Superdex 200PG size-exclusion **(C)** chromatography. Absorbance of elution fractions was read at 230 nm and 280 nm. Peaks in absorbance indicate the presence of protein in elution fractions. For nickel-affinity chromatography, elutions were carried out with a stepwise increase in imidazole concentrations, while the peptides were eluted from the cation-exchange column with an increasing gradient of NaCl (0-1.0 M). Fractions were analyzed by SDS-PAGE and eluent containing the protein of interest was pooled either for further purification, or dialysis and lyophilization.

A)


Absorbance (mAU)
3000
2500
2000
1500
1000
500
0
230 nm
280 nm
Elution Fraction

B)



C)

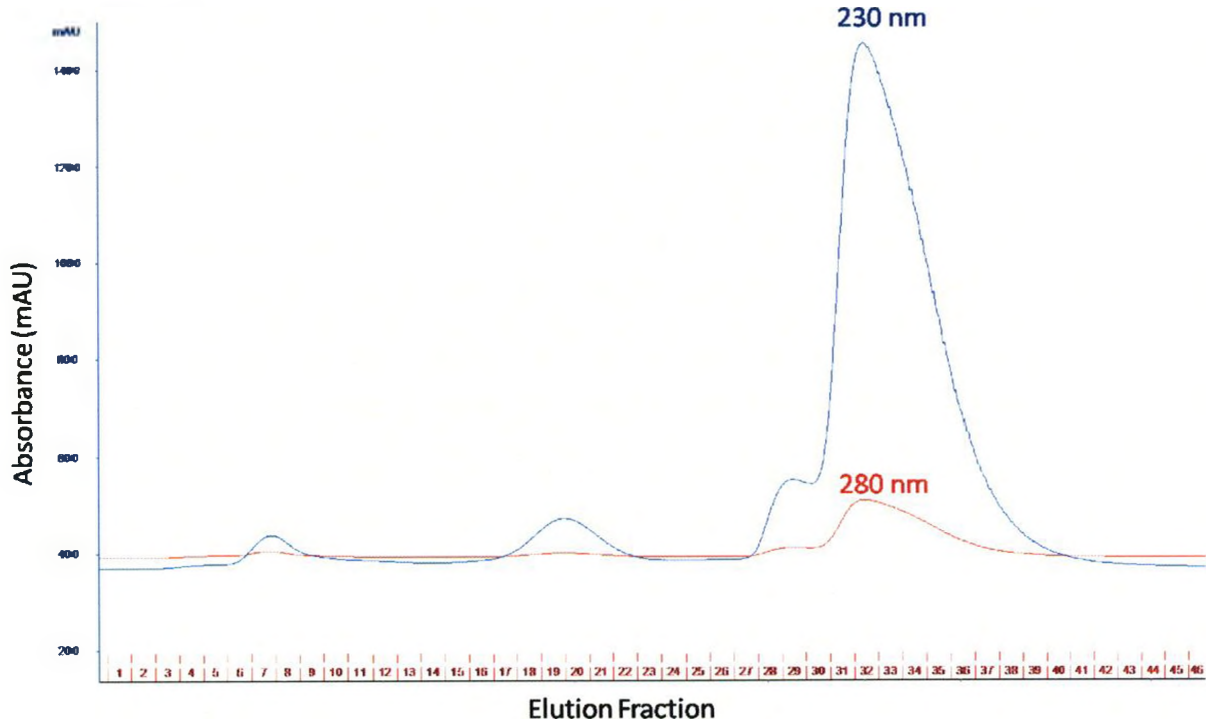


Figure 2.8 Demonstration of the purity of each BSP (1-100) peptide using SDS-PAGE.

Roughly 20 μg of each peptide was separated in a 12% polyacrylamide gel. The gel was initially stained with Stains-All (with an affinity for acidic proteins) and then with silver nitrate according to the method of Goldberg and Warner, 1997 [252].

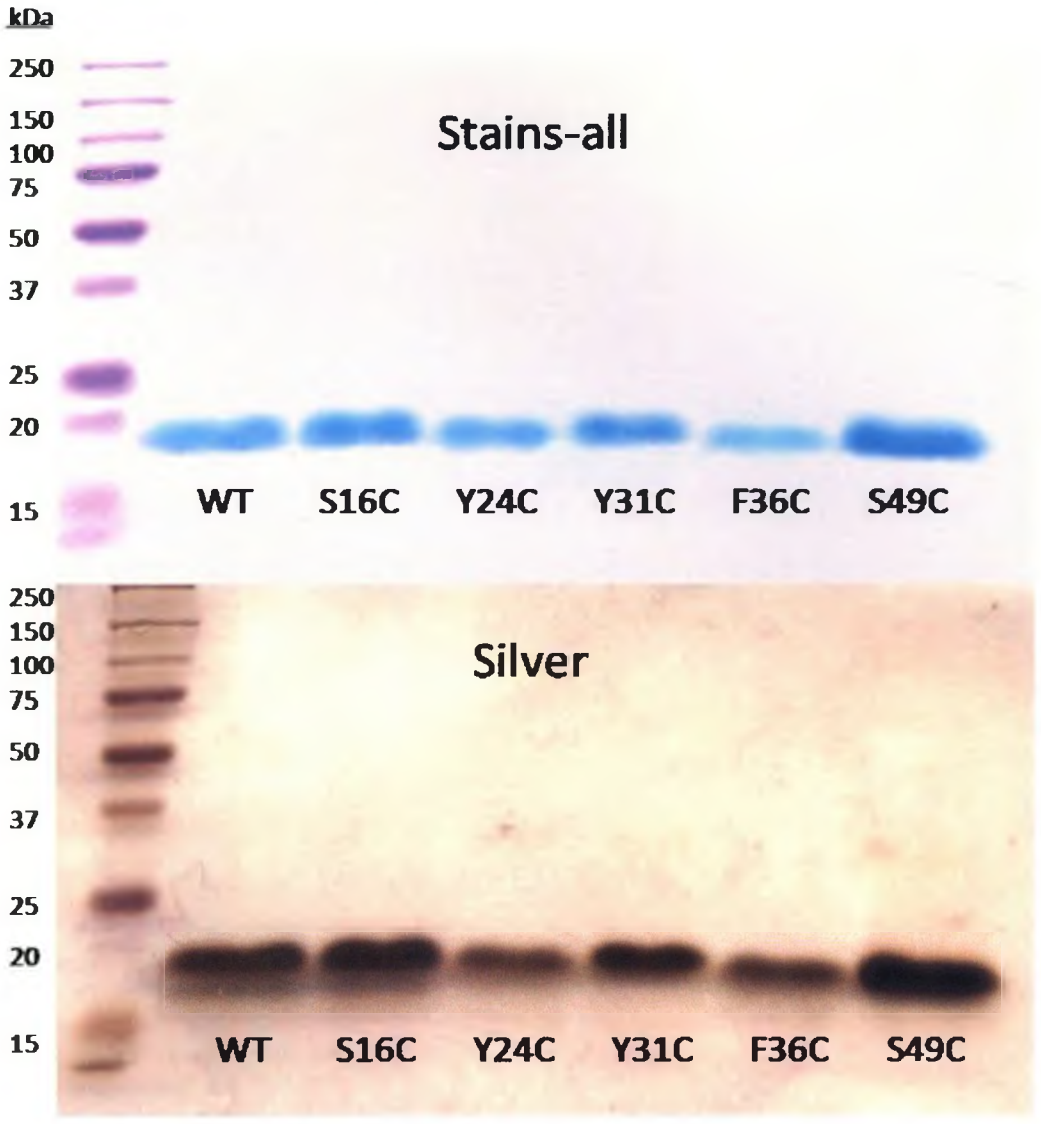

kDa
250
150
100
75
50
37
25
20
15
Stains-all
WT
S16C
Y24C
Y31C
F36C
S49C
250
150
100
75
50
37
25
20
15
Silver
WT
S16C
Y24C
Y31C
F36C
S49C

Table 2.3 Summary of amino acid analyses of the purified BSP (1-100) peptides.

The relative abundance of each amino acid was converted to residues per 1000. The expected values were calculated based on the 107-residue sequence of the BSP (1-100) peptide used, including a C-terminal His-tag. The discrepancies especially apparent with regards to cysteine and serine residues are due to poor recovery of these residues in amino acid analysis.

	Expected*	WT	S16C	Y24C	Y31C	F36C	S49C
Asp/Asn	132	140	136	137	136	139	135
Glu/Gln	217	227	225	230	223	230	230
Ser	104 (94)	88	79	89	88	88	79
Gly	75	77	85	85	87	84	81
His	66	75	74	75	69	71	71
Arg	47	48	50	46	47	47	50
Thr	28	24	26	25	27	26	25
Ala	57	56	58	57	61	59	58
Pro	38	52	40	40	42	42	41
Tyr	47 (38)	47	47	38	36	43	47
Val	38	33	36	35	39	38	38
Met	19	8	10	10	11	11	10
Cys	0 (9)	0	2	2	1	2	2
Ile	9	9	11	10	13	11	10
Leu	28	30	30	30	31	31	31
Phe	47 (38)	46	47	47	45	37	48
Lys	47	40	44	44	43	40	44

Note: In brackets are the values (per 1000 residues) expected upon successful mutation of those residues in the corresponding peptides. Underlined are the values that were expected to change upon successful mutation.

* Values expected based on how many of each residue are contained within the BSP (1-100) peptide, rounded to the nearest whole number.

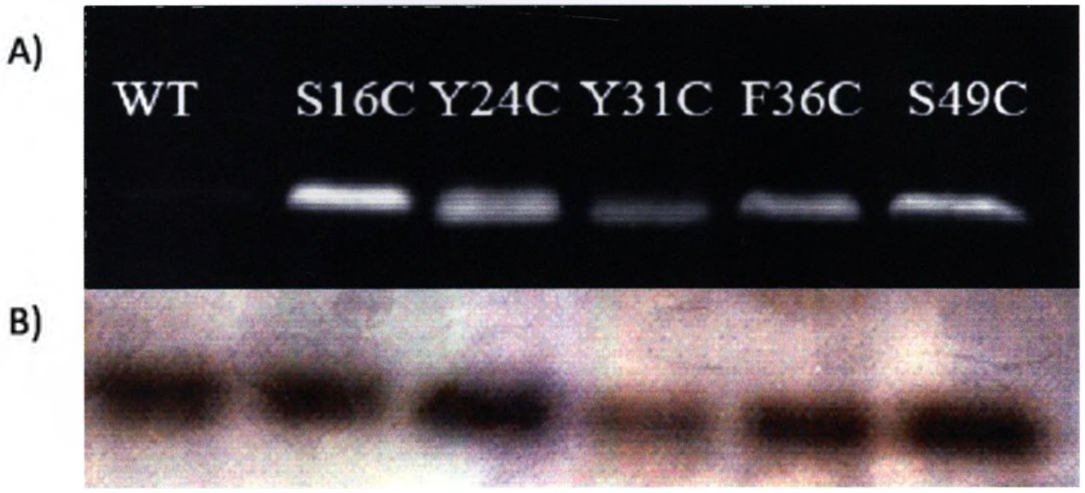
as cysteine, have very poor recovery in amino acid analysis and therefore were reduced in number from what was expected. Additionally, 10 – 15% of serine residues are also lost during this process. Yields of the purified peptides were relatively low, ranging from 1.3 to 3.5 mg per liter of expression media.

2.3.3 Reaction between BSP (1-100) peptides and fluorescein-5-maleimide

Fluorescein-5-maleimide was used to ensure that cysteine residues were both present and reactive. This compound will react with cysteine residues, yielding a protein-fluorophore complex that fluoresces upon UV irradiation. The proteins were first reacted with fluorescein maleimide, and subsequently analyzed by SDS-PAGE followed by UV illumination of the gel. When the proteins were initially reacted with fluorescein maleimide, no observable reaction took place, indicating that the cysteine residues were most likely oxidized (data not shown). Fortunately, in the presence of low amounts of the reducing agent TCEP, peptides containing cysteine residues successfully reacted with the fluorescent compound (Figure 2.9A). Fluorescent bands in the location of BSP (1-100) represent successful reaction with cysteine residues. The gel was subsequently stained with Stains-All and silver to determine the presence and relative amount of protein present (Figure 2.9B). All of the BSP (1-100) mutants successfully reacted with the compound, indicating the presence of reactive cysteine residues. Surprisingly, there was a faint fluorescent band corresponding to fluorescein maleimide reacting with wild-type protein containing no cysteine residues. Although the maleimide group is highly specific for cysteine residues between pH 6.5 – 7.5, a low level of reaction with primary amines can possibly occur due to microenvironmental effects surrounding amine groups.

Figure 2.9 Reaction of fluorescein-5-maleimide with each BSP (1-100) peptide.

Each peptide was reduced using TCEP for a period of 30 minutes, followed by a 90-minute incubation with fluorescein-5- maleimide using established protocols [253]. The products were analyzed by SDS-PAGE in a 12% polyacrylamide gel, and imaged on a UV transilluminator (365 nm) **(A).** The gel was subsequently stained with Stains-All and silver to confirm the presence of BSP (1-100) **(B)**.

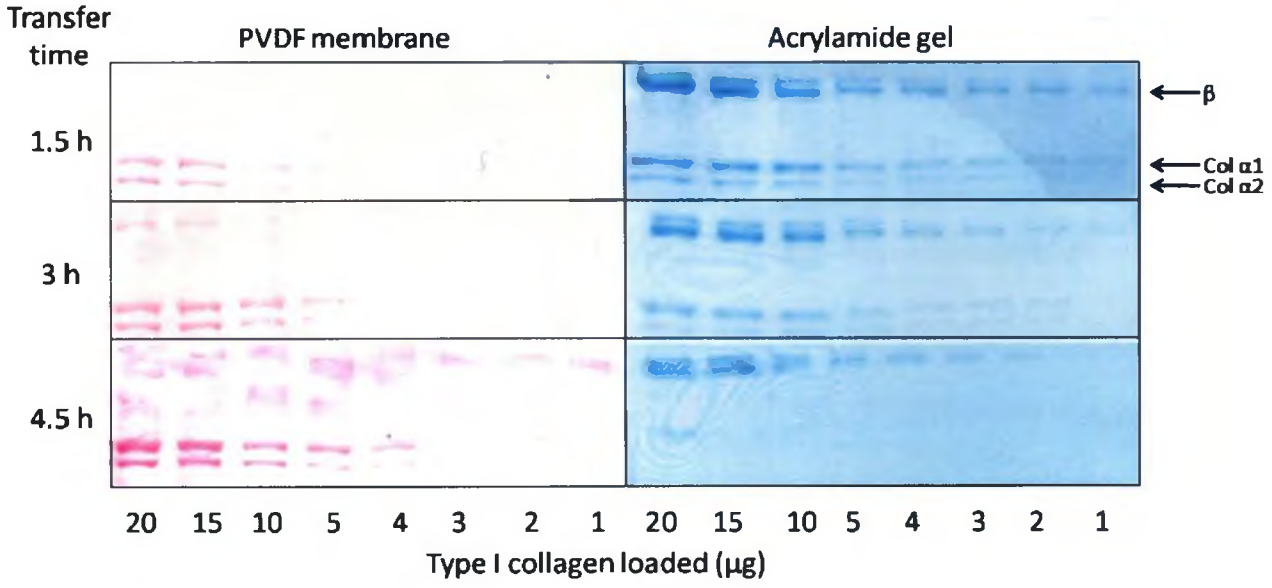

A)
WT
S16C
Y24C
Y31C
F36C
S49C
B)

2.3.4 Optimization of type I collagen transfer to PVDF

Western blot and mass spectrometry are the most common methods used to confirm the presence of a cross-link between two proteins. Because intact collagen is too large for conventional mass spectrometry, Western blot was used in this study. For our purposes, this would require the transfer of collagen and BSP from polyacrylamide gels to PVDF membranes. Because sensitivity is often a significant issue with cross-linking experiments due to low levels of cross-link formation, a high efficiency transfer of protein must occur from the gel to a PVDF membrane. Collagen monomers and aggregates are notoriously difficult to transfer, and therefore experiments were undertaken to determine the conditions that optimize this transfer. Either a carbonate transfer buffer or a Tris-glycine buffer was used; however, there was no appreciable difference in transfer efficiency between the two buffers (data not shown). Various amounts of collagen ranging from 1 – 20 μg were separated by SDS-PAGE, followed by transfers at either 30 V or 100 V for a variety of time periods. Following the transfer, the gel was stained with Coomassie blue, and the PVDF membrane with Ponceau S stain (Figure 2.10). Under all of the conditions tested, a significant amount of collagen remained in the gel. The optimal transfer observed occurred with a transfer at 100 V for 3 h. Although longer transfers yielded marginally better results, the heat generated during transfers of these time periods appeared to alter the binding of collagen to PVDF: The PVDF membranes under these conditions displayed areas where no observable collagen was present, although it was originally present in the gel. As expected, the collagen α-chains had the most efficient transfer, followed by a very low level of β-bands transferring (sets of two α-chains). Under no conditions were any higher-molecular-

Figure 2.10 Optimization of the transfer of type I collagen from polyacrylamide gels to PVDF membranes.

Collagen solutions containing 20 μg, 15 μg, 10 μg, 5 μg, 4 μg, 3 μg, 2 μg, or 1 μg protein were separated by SDS-PAGE in a 6% gel, followed by a 'wet' transfer in Tris-glycine buffer (pH 8.3) for various time periods at 100 V. Following the transfer, PVDF membranes were briefly stained with Ponceau S. The gel was stained with Coomassie brilliant blue. Intensities of the bands were compared to qualitatively assess transfer efficiency under each set of conditions.


Transfer time
PVDF membrane
Acrylamide gel
1.5 h
3 h
4.5 h
β
Col α1
Col α2
20 15 10 5 4 3 2 1 20 15 10 5 4 3 2 1
Type I collagen loaded (μg)

weight aggregates transferred despite their presence in the polyacrylamide gel. Therefore, a three-h transfer in a Tris-glycine buffer at 100 V was used for all transfers involving collagen.

2.3.5 Cross-linking using BPM and APB

Experiments were conducted to cross-link the BSP (1-100) mutants to collagen using the cross-linkers APB and BPM. The reactions were then analyzed by Western blot. Initially, either a polyclonal N-terminal BSP antibody or a commercial anti-His antibody was used for analysis. However, due to the poor specificity and low sensitivity of both of these antibodies, the HisProbe kit from Pierce was used for all Western blot analyses. This is a non-antibody Western blot system that utilizes the nickel affinity of His tags for highly sensitive and specific detection of recombinant His-tagged proteins. Using this system, there was no detectible cross-link formed between any BSP peptide and collagen using either BPM or APB (Figure 2.11). A band corresponding to a successful cross-link of BSP to collagen would most likely be present immediately above the collagen α-chains.

It is interesting to note that each BSP peptides were self-cross-linking, resulting in a ladder-type banding pattern. This was even apparent, though at relatively low levels, with the wild-type protein, presumably due to a low level of cross-linkers reacting non-specifically with amine groups. Also of interest is the fact that collagen appeared to have a blocking effect in the upper mass range, resulting in a negative signal in the Western blot image. Detection of a successful cross-link, if present in very small quantities, may be complicated by this effect. However, the advantage of this blocking effect is the ability to easily visualize the location of the α1 and α2 chains in this analysis.

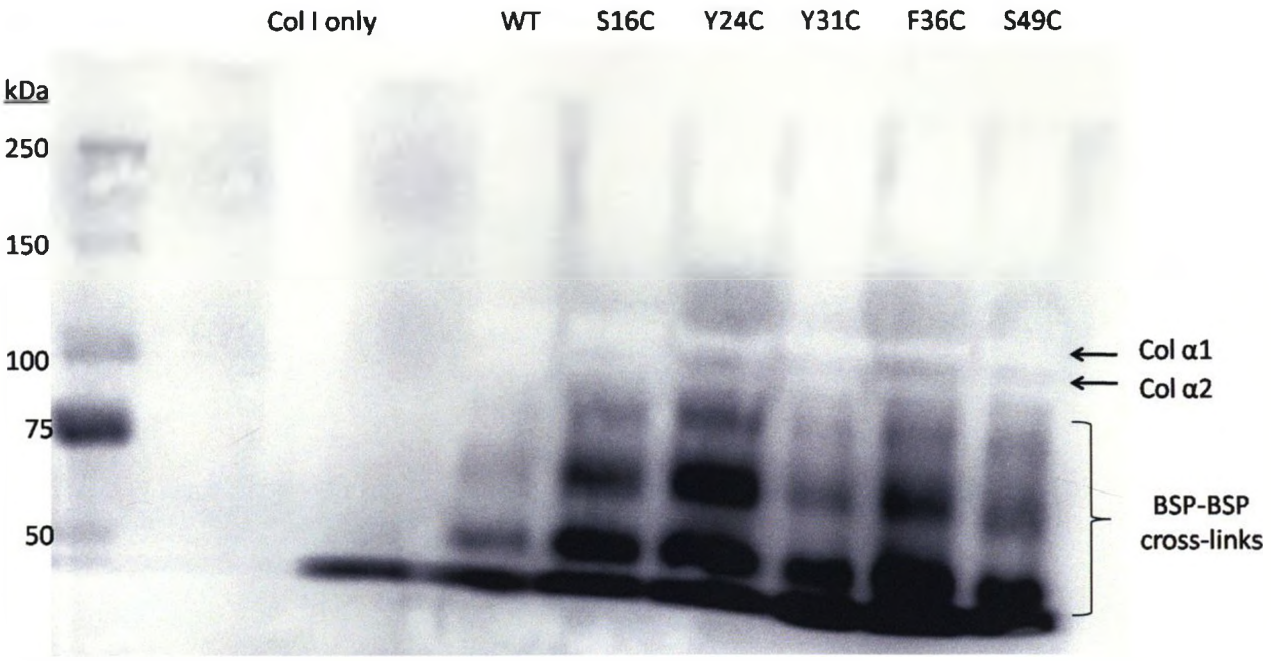
The lack of a cross-link between BSP and collagen could be due to many reasons, such as a spacer-arm that was not sufficiently long enough for a cross-link to occur. Another, but preventable possibility was that the cross-linkers were not initially reacting with the cysteine residues of the BSP (1-100) peptides. An experiment was carried out in which the cross-linkers BPM and APB were reacted with the reduced peptides under conditions specific for those cross-linkers. The efficiency of the reaction between the cross-linkers and the peptides was qualitatively analyzed by the subsequent reaction with fluorescein-5-maleimide and analysis after separation by SDS-PAGE. Fluorescent band intensities were compared with samples in which FM was reacted with peptides that had not been conjugated to cross-linkers. It was confirmed that the cross-linkers were, in fact, reacting with the cysteine residues with high efficiency, indicated by the reduction or complete loss of fluorescence in the presence of cross-linker molecules (Figure 2.12). Therefore some other factor must be preventing the formation of a cross-link, such as the length of the cross-linker or the characteristics of the photoactivatable functional group.

An interesting observation was the fact that each BSP peptide cross-linked to other BSP peptides with different frequencies depending on the cross-linker used. For instance, the mutant Y24C appeared to have the greatest quantity of BSP-BSP cross-links while using BPM, while Y31C had a relatively low level of these cross-links. In contrast, in reactions using APB, Y31C had the greatest amount of BSP-BSP cross-links while levels of Y24C complexes were low in comparison. This may be due to the slightly different lengths of the cross-linkers, or the fact that these cross-linkers possess different photoactivatable groups. These groups may have different propensities to react with the closest amino acids in the BSP-BSP interaction.

Figure 2.11 BSP-collagen cross-linking reactions involving BPM and APB.

Cross-linking reactions were carried out between collagen and each of the BSP (1-100) peptides using the heterobifunctional cross-linking reagents BPM (A) and APB (B). Once complete, reactions were separated in a 6% polyacrylamide gel, and the proteins transferred to PVDF for analysis by Western blot. A band corresponding to a cross-link would be located above the collagen α-bands. The ladder-type banding pattern occurring below the collagen α-chains corresponds to cross-linked aggregates of BSP peptides. No cross-link to collagen was detected with either cross-linker. Collagen has a pronounced blocking effect in the upper mass range.

A) Cross-linking using BPM



B) Cross-linking using APB

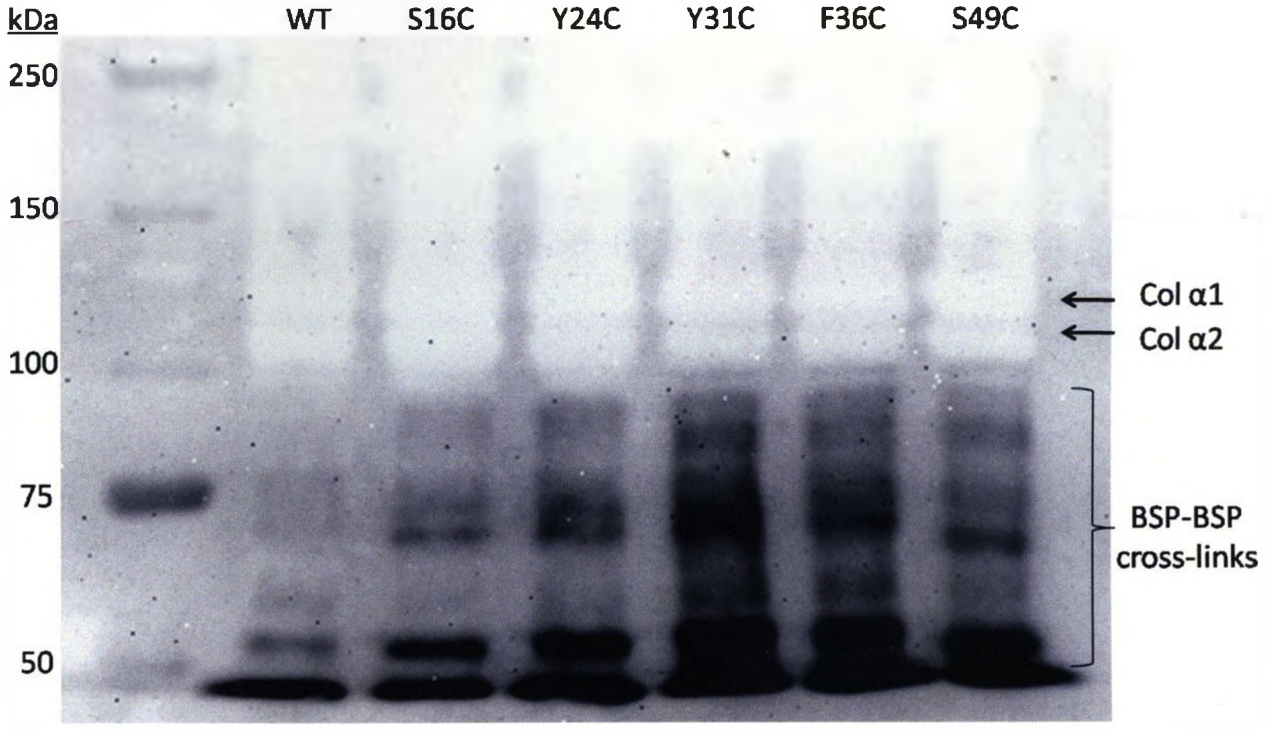
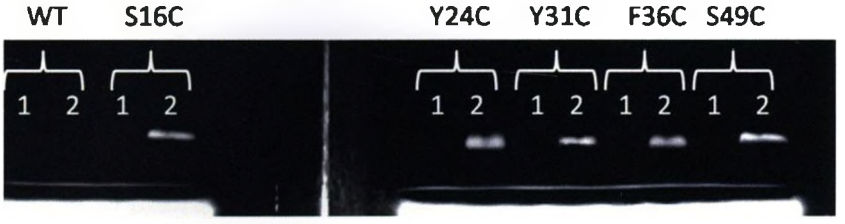


Figure 2.12 Reaction of BSP peptides with the cross-linkers BPM and APB.

Fluorescein maleimide was added to BSP (1-100) peptides before and after reacting with benzophenone-4-maleimide (A) or ρ-azidophenacyl bromide (B). Successful conjugation of cross-linkers to cysteine residues results in the inability of that residue to react with FM, and therefore decreased fluorescence.

A) BPM



B) APB

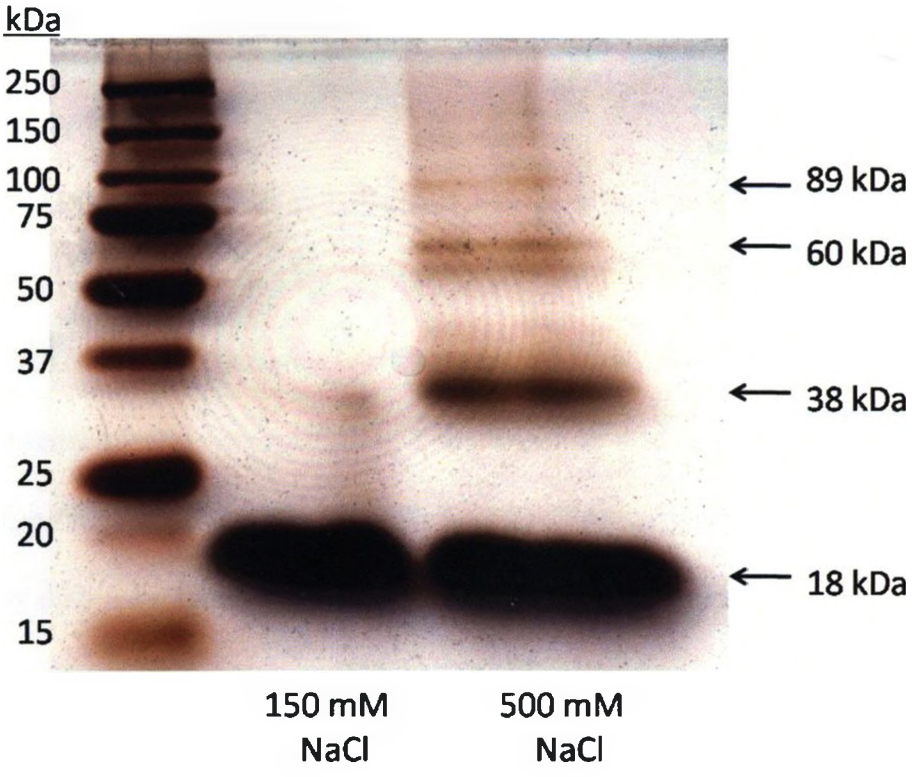
1 - peptide initially reacted with cross-linking reagent, then FM
2 - peptide only reacted with FM

2.3.6 Characterization of the BSP-BSP interaction

It was of interest to characterize the BSP-BSP cross-links observed in the Western blot analysis, as these aggregates result in decreased protein available for cross-linking to collagen. BSP (1-100) is highly negatively charged, although there are three lysine and three arginine residues within the collagen-binding region. Furthermore, the collagen-binding domain of BSP is mainly hydrophobic in nature. Because of these properties, the BSP peptides generated could potentially interact with other BSP peptides via electrostatic interactions and/or hydrophobic interactions. Therefore, it was of interest to characterize the molecular interaction(s) involved to ensure that enough BSP would be in monomeric form to interact with collagen. In order to do this, the most abundant mutant, Y24C, was conjugated to the cross-linker BPM followed by UV irradiation in the presence of physiological (150 mM) or high (500 mM) NaCl concentrations. The reactions were then analyzed by SDS-PAGE and the gel stained with Stains-All and silver. Although a very low amount of BSP-BSP cross-links were present in the lower-salt reaction, the propensity for Y24C to cross-link to other BSP peptides significantly increased under conditions of high salt (Figure 2.13). In high-salt conditions, distinct bands were detected at 18 kDa, 38 kDa, 60 kDa, and 89 kDa. The three highest molecular weight bands presumably represent cross-linked aggregates of 2, 3, and 4 BSP (1-100) Y24C molecules. There is faint evidence of even higher-molecular-weight aggregates in low quantities. In the lower-salt reaction, no aggregates above 38 kDa were observed. The fact that higher electrolyte concentrations did not abolish the BSP-BSP interaction suggests that the interaction is hydrophobic in nature. At 150 mM NaCl, a significant

Figure 2.13 Effect of sodium chloride on the BSP-BSP interaction.

BSP (1-100) Y24C was reduced and conjugated to the cross-linker benzophenone-4-maleimide. Additional buffers containing NaCl were then added, resulting in either 150 mM or 500 mM NaCl. The reactions were irradiated with UV light (365 nm) for a period of 20 minutes and analyzed by SDS-PAGE on a 12% polyacrylamide gel. The gel was stained with Stains-All and silver nitrate prior to imaging.


kDa
250
150
100
75
50
37
25
20
15
89 kDa
60 kDa
38 kDa
18 kDa
150 mM NaCl
500 mM NaCl

amount of BSP (1-100) Y24C is still in its monomeric form, and therefore will be available to cross-link to collagen molecules in close proximity.

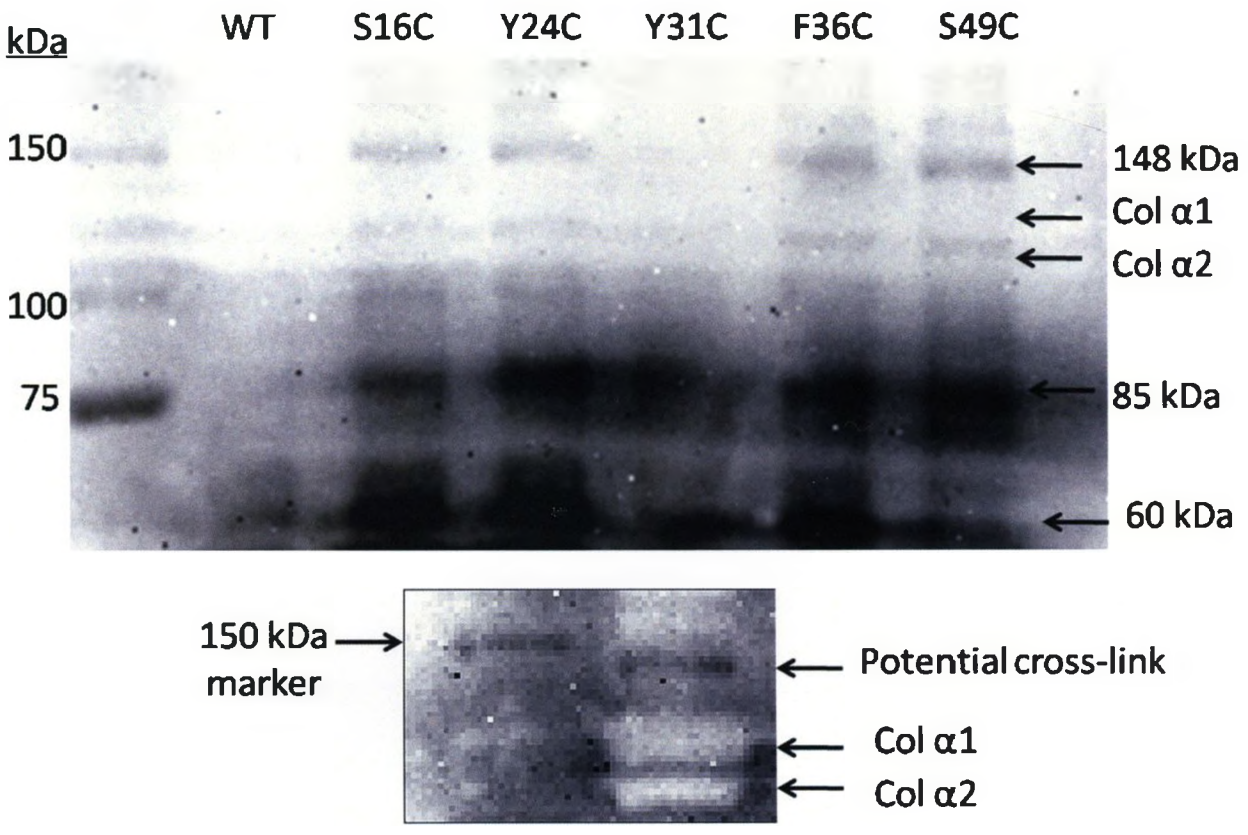
2.3.7 Cross-linking of BSP (1-100) peptides to type I collagen

It was important to use another cross-linker with a longer spacer-arm than both APB and BPM, which are both shorter than 1.2 nm. Therefore reactions were carried out where each BSP peptide was conjugated to the cross-linking molecule APDP (2.1 nm). Dissolved collagen was then added to this reaction. The mixture was irradiated by direct UV light to induce the formation of cross-links, and subsequently analyzed by SDS-PAGE followed by Western blot.

The use of APDP, in contrast to BPM and APB, resulted in BSP-collagen cross-links with the BSP (1-100) mutants S16C, Y24C, F36C, and S49C (Figure 2.14). There was a very low level of cross-linking between mutant Y31C and collagen. Unexpectedly, a relatively low-intensity band corresponding to a cross-link was present in the reaction containing wild-type BSP (1-100) (Figure 2.14, lane 2). This suggests either that APDP is forming dimers by disulfide bonding, resulting in a longer cross-linking agent with a photoactivatable group on either end, or that APDP is somehow non-specifically reacting with the wild-type protein. Subsequent experiments have shown that these cross-links are dependent on the presence of APDP, collagen, and UV irradiation, and that the band does not appear if the same reaction is reduced with β-mercaptoethanol prior to separation by SDS-PAGE (data not shown). Although the cross-link observed appeared very similar in molecular weight to the 150-kDa ladder marker, subsequent cross-linking reactions indicated that the cross-link was slightly lower in apparent weight (Figure 2.14, lower image). As with the previous cross-linkers used, in the presence of type I collagen, BSP (1-100) appears to be cross-linking to itself at a similar level to that seen under conditions

Figure 2.14 BSP-collagen cross-link formation using APDP.

Each BSP (1-100) peptide was conjugated with APDP and incubated with type I collagen prior to UV irradiation. The reactions were then centrifuged and the resulting precipitated-collagen pellet analyzed by SDS-PAGE followed by Western blot using a His-tag probe. Note the cross-linked species running slightly faster than the 150-kDa molecular weight ladder protein. As with previous Western blot analyses, BSP-BSP aggregates were observed and collagen α-chains produced a blocking effect enabling the visualization of their location.


kDa
WT
S16C
Y24C
Y31C
F36C
S49C
150
100
75
148 kDa
Col α1
Col α2
85 kDa
60 kDa
150 kDa
marker
Potential cross-link
Col α1
Col α2

of higher salt. The typical BSP aggregates observed in the previous BSP-BSP interaction study are observed up to the tetramer level in the cross-linking experiments. In contrast to the variable BSP-BSP cross-links observed while using BPM or APB, each single-cysteine mutant conjugated to APDP formed a similar high level of aggregation, presumably due to a less-stringent, longer spacer arm.

2.4 Discussion

The most promising ways to determine where BSP binds to collagen begin with chemical cross-linking. This is due to the ability, through digestions, mass spectrometry and other techniques, to map binding interfaces down to the amino acid level. In these experiments, chemical cross-linking offers the additional advantage that the triple-helical structure of collagen can be maintained throughout the cross-linking process to obtain physiologically relevant results. We have successfully created the reagents and protocols required to obtain a covalent attachment between BSP and collagen. This process required the expression, purification and characterization of single-cysteine peptides of BSP, covalent attachment of these peptides to various cross-linking molecules, UV-induced cross-link to collagen, characterization of BSP-BSP interactions, optimization of the transfer of collagen to PVDF for Western blot analysis, and experimentation with various Western blot detection methods. Through these developments, a cross-link between BSP and collagen has been created with 4 of the 5 single-cysteine BSP-mutants.

2.4.1 BSP (1-100) single-cysteine mutants

Multiple single-cysteine mutants were created in order to maximize the chances of a successful cross-link. The residues mutated to cysteine were chosen based on location with respect to the collagen-binding domain of BSP, as well as the nature of the amino acid side-chain. The three sites chosen within the collagen-binding domain (Y24, Y31, and F36) were bulky, aromatic residues in order to minimize steric hindrance between collagen and the BSP peptides due to the presence of cross-linkers. Interestingly, other researchers established a cross-link between PEDF and collagen after mutating phenylalanine and tyrosine to cysteine [225]. In the case that any mutation within the

collagen-binding domain of BSP interfered with binding, it was important to mutate residues outside this region as well. The two serine-to-cysteine mutations at amino acid positions 16 and 49 are just outside the collagen-binding domain, and therefore steric hindrance was less of a consideration, and maintaining a chemically similar side-chain became more favourable.

Each BSP (1-100) peptide migrated more slowly than predicted by molecular weight in SDS-PAGE. This slow migration is also characteristic of full-length BSP and is poorly understood. One factor that may be responsible for this is the extended, flexible structure of BSP. Although SDS is thought to induce a fairly extended structure in most proteins, it is likely that the degree of denaturation varies and that BSP maintains a more extended structure than most other proteins. The highly acidic nature of BSP may also play a role in polyacrylamide gel migration rates. Acidic residues may repel negatively-charged SDS molecules resulting in a less negatively-charged protein, thus affecting migration rates.

2.4.2 Fluorescein maleimide reactions

Fluorescein-5-maleimide covalently reacts with cysteine residues in much the same way that cross-linking molecules do, yielding a fluorescent product. This reagent was used to ensure that the cysteine residues were available and able to react with cross-linking reagents. It is interesting to note that a low level of fluorescein-5-maleimide appeared to react with wild-type BSP (1-100) peptides despite the fact that they contain no cysteine residues. Theoretically, maleimide reactions are completely specific to thiol groups between pH 6.5 – 7.5. There are two possible contributing factors to this phenomenon. Firstly, as mentioned earlier, the chemical microenvironment surrounding

amino acids can alter pKa values, resulting in protonation/deprotonation at pH values where this would not normally occur, and therefore allowing, for instance, certain amine groups to become more nucleophilic [223]. This would allow the maleimides to react with lysine residues or the N-terminal amino group of the protein. Secondly, through the use of maleimide-based dyes similar to fluorescein-5-maleimide, it has been shown that a large excess of maleimide over thiol groups can result in non-specific reactions with other groups (presumable amines) [253]. In this study, Tyagarajan and coworkers optimized concentrations and molar ratios of maleimide dyes to thiols to the reducing agent TCEP. These concentrations were closely related to those used in this thesis for cross-linking and labeling with maleimide agents. However, FM may have slight differences in chemical reactivity compared with the maleimide-based dyes used in the optimization study. Alternatively, the microenvironment surrounding cysteine residues in the BSP peptides may differ from proteins used in that study, and therefore it is possible that non-specific labeling occurs at a lower maleimide:thiol ratio than that in other systems. Fluorescein maleimide reacted well with the single-cysteine mutants, displaying a large increase in fluorescence over the wild-type peptide. This confirms the presence of reactive cysteine residues that can be used for cross-linking reactions.

2.4.3 Transferring type I collagen to PVDF

In virtually all cross-linking studies, the formation of a cross-link is initially confirmed by mass spectrometry or Western blot prior to further digestion and analysis. Conventional MS techniques do not allow the analysis of large proteins such as type I collagen. Therefore, a Western blot approach was undertaken. Due to its physical and chemical properties, collagen does not transfer well to PVDF or nitrocellulose

membranes. Therefore it was necessary to establish the transfer protocol that would optimize this transfer process. Not surprisingly, the transfer efficiency of collagen remained relatively low under all conditions tested. This is presumably due to the poor mobility of collagen within polyacrylamide gels (type I collagen α-chains run 20 – 35 kDa heavier than they should in SDS-PAGE), the sheer size of type I collagen that inhibits its movement in the gel, as well as the pI of collagen (approximately 8.3). Because the use of a carbonate buffer shown to be effective for basic proteins [255] did not significantly increase transfer efficiency, the former properties of collagen are likely the reason for poor transfer. A lower-percentage gel could potentially be used to increase the level of collagen transferred, but this would result in poorer resolution between collagen α-bands on the one hand and a band corresponding to a cross-link between BSP (1-100) peptides and collagen on the other. Although longer transfer times at 100 V seemed to increase transfer efficiency, the associated increase in temperature caused the loss of collagen from large areas of the membrane. The cooling apparatus used was the Mini-protean 3 system (Bio-rad) that was run in a 4°C environment with an ice pack that was replaced half way through the transfer. Although a more complex apparatus could be tested that would circulate a coolant through the equipment during transfer, the increase in transfer over time between 3 h and 4.5 h did not justify this type of upgrade. It appears at this point as though experiments must be carried out at high enough collagen concentrations that bands will be detectible on PVDF despite poor transfer efficiency.

2.4.4 Reaction of BPM and APB with BSP (1-100) peptides

Because no cross-link was formed using either BPM or APB, it was of interest to determine if these reagents had been successfully conjugated to the BSP (1-100) peptides.

Through the use of fluorescein-5-maleimide, the reactions between cross-linkers APB and BPM and the five BSP (1-100) mutants were confirmed qualitatively. A low basal fluorescence was detectible on both wild-type protein and mutants that had already been conjugated to the cross-linkers, indicating a low level of amine reactivity as observed with the initial fluorescein maleimide reactions. Similar results were obtained when the reactions were carried out in 2x PBS, the solution used for cross-linking experiments. These results indicate that the peptides used in the cross-linking studies were indeed conjugated to the cross-linkers APB and BPM. Therefore another property of these reagents was preventing the formation of a cross-link. Most likely, the spacer-arm length of these agents is not long enough to reach a collagen amino acid from the cysteine residues of our BSP peptides.

2.4.5 BSP-BSP cross-links

In initial cross-linking experiments using each of the three cross-linkers, aggregates of cross-linked BSP peptides were present. It was evident that these were aggregates of BSP due to their blue staining with Stains-All indicative of an acidic protein, as well as their reactivity to the Hisprobe reagent indicating the presence of a His tag. Initially, it was hypothesized that this may be an electrostatic interaction, due to the three lysine and three arginine side-chains within the collagen-binding domain and glutamic acid residues outside of this region on other BSP molecules. It was desirable to maintain the BSP peptides in a monomeric form in order to be available to interact with collagen. If this interaction were electrostatic in nature, then it may have been possible to eliminate aggregation by altering salt concentrations. When the propensity for BSP (1-100) S16C to cross-link to itself was assessed under low (150 mM) and high (500 mM)

NaCl concentrations, it was found that the high-salt condition caused a much higher level of cross-linking than the low-salt condition. As mentioned previously, this result indicates that the interaction is not electrostatic in nature. Since high salt did not decrease or abolish the aggregation of BSP (1-100), it appears that hydrophobic forces are the cause of this interaction. In addition to the basic amino acids, the collagen-binding region of BSP contains a high proportion of hydrophobic residues. There are two theories as to how salt affects protein interaction and aggregation. These are the Hofmeister theory involving chaotropic and kosmotropic ions, and the theory of Debye-Hückel screening.

2.4.5.1 The Hofmeister effect

According to Hofmeister, different ions in solution can affect protein stability differently. Chaotropic ions increase the solubility of non-polar molecules and decrease the forces involved in hydrophobic interactions. Conversely, kosmotropic ions increase solvent surface tension, decrease the solubility of non-polar molecules, and increase hydrophobic forces. According to Hofmeister, anions have a greater effect than cations and are ordered according to Table 2.4 [258]. It is generally understood that ions above chloride in the table are kosmotropic and ions below are chaotropic. However, more recent reports studying the ability of anions in aqueous solutions to disrupt surfactant monolayers suggest that chloride is mildly kosmotropic (chaotropic ions readily penetrate and disrupt such monolayers) [259-261]. Therefore, the excess chloride ions in the high-salt cross-linking experiment could act to increase the hydrophobic attraction between BSP molecules, resulting in increased cross-linking once irradiated by UV light. The mechanisms behind these ion effects are not well understood. Originally, Hofmeister and others proposed that the effect ions have on the solubility and interaction of proteins is

Table 2.4 The Hofmeister ion series.

The ions are listed from the most kosmotropic (promoting hydrophobic interaction) to the most chaotropic (decreasing hydrophobic interaction). In general, anions have much more pronounced effect than cations [258].

Mode of Action	Anion	Cation
Promotes hydrophobic interactions	PO_4^{3-}	NH_4^+
↓	SO_4^{2-}	Rb^+
	CH_3COO^-	K^+
	Cl^-	Na^+
	Br^-	Cs^+
	NO_3^-	Li^+
	ClO_4^-	Mg^{2+}
	I^-	Ca^{2+}
Decreases hydrophobic interactions	SCN^-	Ba^{2+}

due to the effects the ions have on the water molecules in the bulk solution. Specifically, chaotropic ions were thought to decrease hydrophobic forces by decreasing the ordering of water molecules in solution, whereas kosmotropic ions were thought to increase the ordering of water molecules, amplifying hydrophobic forces [262]. However, more recent evidence indicates that the effects of these ions are due to their direct interaction with either the protein itself, or the hydration shell immediately surrounding the macromolecule [263]. Kosmotropic ions were found to polarize water molecules that were hydrogen-bonded to amide groups of a macromolecule, and to raise the surface tension of the macromolecule-solution interface, thereby increasing the cost of hydrophobic hydration [264].

2.4.5.2 Debye-Hückel screening

One other possible explanation for the observed effect of salt on the BSP-BSP interaction is that it is caused by Debye-Hückel screening. This effect occurs when ions in solution aggregate in proximity to oppositely-charged regions of macromolecules (for example, sodium ions aggregating near glutamic acid residues of BSP). This local gathering of ions increases the apparent dielectric constant of water and causes electrostatic forces to decrease much more rapidly with distance than they would otherwise. This would decrease repulsive electrostatic interactions, such as that between two highly negatively charged BSP molecules (or repulsion of positively charged residues within the collagen-binding domain), and increase forces such as hydrophobics that promote intermolecular interactions [265, 266]. Interestingly, recent database analyses indicate that regions of proteins intended for interaction with other proteins are the most prone to non-specific aggregation [267]. It has been theorized that charged

residues within close proximity to aggregation-prone regions of proteins act as "gatekeepers" and prevent aggregation [268-270]. These bioinformatic studies provide strong evidence of this theory, as charged residues are almost always found within close proximity to hydrophobic regions associated with protein interactions [267]. The three conserved lysine residues and three conserved arginine residues within the collagen-binding domain of BSP, as well as glutamic acid residues just outside of this region, could therefore act as 'gatekeepers' that inhibit hydrophobic aggregation with other macromolecules. The increase in NaCl concentrations could cause a Debye-Hückel screening effect, reducing the electrostatic repulsion and increasing the aggregating potential of BSP.

The presence of ions including sodium and chloride have been implicated, *in vitro*, in the aggregation of amyloid fibrils implicated in Alzheimer's disease [271]. Similar amyloid aggregates are also involved in pathologies such as atherosclerosis and type II diabetes [272]. Thus the mechanisms by which ions in solution affect biological molecules is an important area of study that is still in its early stages. However, for the purposes of this thesis, it is safe to assume that both theories discussed indicate a hydrophobic interaction between BSP molecules in these cross-linking studies. Salt concentrations above physiological levels should therefore be avoided in studies using BSP in order to prevent non-physiological aggregation.

2.4.5.3 Collagen-induced BSP-BSP cross-links

With regards to experiments involving the chemical cross-linking of BSP to type I collagen, a high level of BSP self-cross-linking was observed despite the use of a buffer containing only 150 mM NaCl. The only possible source of this effect is the 1 μg/μL

collagen in 5 mM acetic acid added to the BSP (1-100) peptide solutions prior to cross-linking. Although the acetate ion is a slightly stronger kosmotrope than chloride, it is not present in high enough final concentrations to elicit an effect (2.5 mM). A significant drop in pH could also alter protein interactions; however, the pH of the phosphate-buffered mixture is at neutrality after addition of the collagen solution. There may be excess salt present in spite of the exhaustive dialysis used in the preparation of type I collagen prior to lyophilization, and therefore actual ionic concentrations may be inflated over what is dissolved in the buffers alone. Otherwise, collagen itself must be somehow affecting BSP-BSP interactions. Although unlikely, multiple BSP molecules could bind to a similar region on each collagen triple-helix, and cross-links could then form between BSP peptides. In this case, however, it would be expected that cross-links would also occur between BSP peptides and collagen. Many cross-linking reactions were performed in the presence of collagen where BSP cross-linked aggregates were present, without detection of a cross-link to collagen. Also, past work from our laboratory involving binding curves of BSP to collagen fit well to a one-site binding model, which makes it unlikely that many BSP molecules are binding to collagen at the same location.

It appears as though there is not any realistic way to prevent BSP-BSP aggregates from forming in these cross-linking reactions with the available reagents. Fortunately, a majority of BSP ultimately remains in monomeric form and is therefore available for interaction with collagen.

2.4.6 BSP-collagen cross-link

2.4.6.1 Establishing and analyzing a BSP-collagen cross-link

It is puzzling that no cross-link was detected with collagen when using the cross-linkers benzophenone-4-maleimide and ρ-azidophenacyl bromide, as many successful cross-linking experiments involving other protein complexes have been characterized using these cross-linkers (reviewed in section 2.1). Since it was shown that these cross-linkers successfully react with cysteine residues on BSP (1-100) peptides, the photoactivatable functional groups must not be reacting with collagen. In the past, benzophenone groups have been shown to have a preference for methionine residues in what has been termed a 'magnet effect' of methionines towards the benzophenone [273]. Methionines are fairly spread out within the type I collagen molecule, and therefore if no methionine is present near the cross-linker, than it would be less likely to form a cross-link upon UV irradiation. However if the benzophenone group were the sole reason for a lack of cross-linked species, then one would expect APB, with an only slightly shorter spacer arm (0.9 nm vs. 1.14 nm) and a different photoactive functional group, to be able to form a cross-link. This is further supported by the fact that a successful cross-link was established using APDP, which also contains a reactive nitrene similar to APB. It seems as though the spacer arm of both of these molecules is not sufficiently long enough to reach collagen from any of the five locations tested. Alternatively, it is possible that cross-links are being formed at such low efficiencies that the current detection limits are not low enough to detect them.

The goal of this project was to establish the reagents and protocols needed to cross-link BSP to collagen. This appears to have been successfully accomplished with the

use of APDP as a cross-linking reagent. A successful cross-link would be on the order of 110 kDa, however due to the poor migration of both collagen and BSP, the observed cross-link migrated as though it were about 148 kDa when analyzed by Western blot. The molecular weight of each collagen α-chain is approximately 97 kDa, with very little difference between α1 and α2 monomers. However, when separated by SDS-PAGE, they have markedly different migration rates (migrating some 15 kDa apart), and both migrate much more slowly than their molecular weight would suggest (running at 135 and 120 kDa, respectively). As previously discussed, the BSP (1-100) peptides also migrated more slowly than expected (18 kDa versus a MW of 12.2 kDa). While the migration of a successful cross-link is difficult to predict, an approximate range in which the cross-link might appear would be anywhere from 132 kDa to 155 kDa. The observed cross-link at 148 kDa appears within this range. Because this band corresponds to the presence of a His-tag, and is not observed in either the presence of a reducing agent or in the absence of collagen or APDP, it is highly likely that it corresponds to a cross-link between BSP and collagen. Interestingly, another group of researchers successfully established a cross-link between type I collagen and another protein, PEDF, using APDP [225].

2.4.6.2 Further characterization

It is possible that the four BSP peptides successfully cross-linked to collagen have been linked to various sites on both α1 and α2 chains: BSP at any location on collagen is in close proximity to three α-chains. Emsley and coworkers have solved the crystal structure of a complex between a 21-residue synthetic type I collagen triple-helical peptide bound to an integrin ligand, and determined that amino acids from all three subunits of the triple-helix participate in the interaction [215]. Since the presence of all

three chains in a triple helix is also important for binding BSP [45], it is probable that at least 2 chains have some type of interaction with BSP. Although various sites of cross-links on collagen would make data analysis more difficult, it could provide us with a more comprehensive understanding of the way BSP binds to collagen. Comparing the cross-links formed from locations within the collagen-binding region of BSP to those outside of it, while taking into account the microfibrillar structure of collagen solved by Orgel and colleagues [49], would provide information about conformation of BSP upon binding. Similar comparisons can be made between the S16C mutant and the S49C mutant, for example, to visualize the orientation of BSP on collagen. As BSP has higher HA nucleation potency upon binding type I collagen [45], this type of conformational study would provide valuable information about protein secondary structures that promote mineralization. This information would be of high importance in the development of bone-repair therapeutics.

The visualized region of the triple-helix to which BSP binds could then be mapped onto the theoretical collagen model developed by San Antonio and colleagues [248] and compared with other collagen-binding ligands. It is thought that BSP will be found to bind to collagen within the collagen matrix-interacting domain proposed by San Antonio and, more specifically, bind to collagen adjacent to the hole zones where initial mineralization occurs *in vivo*.

Further experiments involving the digestion of the cross-linked complex and characterization by a method such as mass spectrometry should be carried out to characterize these findings. However, it should be taken into consideration that characterizing CNBr digestion fragments of collagen by mass spectrometry is difficult

due to heterogeneous protein modifications [257]. However difficult it may be, this digestion process is essential to determining the region of collagen that interacts with BSP. The protocols involved in this process are further described in the following section.

2.5 Future directions

The goal of this research is to ultimately elucidate the region of type I collagen to which BSP binds. Up to this point, a covalent cross-link has been obtained between a BSP peptide and collagen through the use of the cross-linking agent APDP. In order to locate the specific region of collagen to which BSP has been cross-linked, additional research needs to be conducted. There are a variety of ways that this could be accomplished and some preliminary studies have already been undertaken. Firstly, it would be of interest to determine the α-chain(s) of collagen involved in the cross-linked complex.

Yasui and Koide have cross-linked PEDF to collagen using APDP, and went on to determine that the α2 chain was the sole collagen component of the cross-link. This was accomplished by taking advantage of the fact that APDP acts through disulfide exchange to react with cysteine residues. A 2-dimensional electrophoresis experiment was undertaken in which the intact cross-link was separated in one dimension and the gel subsequently incubated in reducing conditions. This breaks the disulfide bond between their single-cysteine mutant and APDP, leaving collagen with a thiol group at the free end of APDP. As collagen contains no endogenous cysteine residues after procollagen fragments have been cleaved, a sole thiol is now present on the collagen molecule where the cross-link was initially formed. The reduced sample was separated in the second dimension and the gel incubated in a solution containing fluorescein-5-maleimide. When the gel was analyzed under UV light, the electrophoretic mobility of the now fluorescent collagen molecule was compared against the mobility of each α-chain (once stained with Coomassie blue) to determine that the α2 chain was the one involved in the cross-link. In

additional experiments, the cross-linked collagen was cleaved with MMP-1, resulting in two fragments of ¾ and ¼ of the full-length collagen molecule. It was determined that PEDF bound to the C-terminal quarter of collagen.

This method looks promising; however, this particular protein-collagen cross-link appeared to occur with much higher efficiency than those obtained by cross-linking BSP to collagen. As this type of fluorescence requires a significant amount of protein present to be detected, sensitivity becomes a limiting factor.

A common issue with cross-linking experiments is the sensitivity of detection, as cross-linking efficiencies are usually very low. In order to circumvent this limitation, some researchers have used trifunctional cross-linkers containing a biotin group, and subsequently enriched the cross-linked products through affinity chromatography [218, 274]. Although effective, this approach requires the use of a large amount of protein to obtain the desired cross-linked species in high quantities if the cross-linking efficiency is low. One advantage of the cross-linker APDP that has not yet been exploited in these studies is its ability to be radioiodinated. A radioisotope would permit greater detection limits, for instance, in SDS-PAGE. In a similar manner to how Yasui and colleagues labeled collagen with the thiol group of APDP at the PEDF-binding region [225], it would be possible to perform a label transfer reaction where the radio-labeled APDP is cleaved from BSP with a reducing agent, labeling the interacting region of collagen with APDP containing ^{125}I. This type of study would provide some valuable preliminary information as to the general region of collagen to which BSP binds with significantly increased sensitivity.

To answer questions as to exactly how and where BSP binds to collagen, the structure it adopts after doing so, and how this results in increased mineralization, a highly sensitive and high resolution detection system is needed. A starting point for this type of study would be to cleave collagen into many fragments. Because the triple-helix of collagen is highly resistant to enzymatic cleavage, it has traditionally been digested using cyanogen bromide, cleaving the peptide bond at the carboxyl end of methionine residues. The CNBr fragments of collagen are well characterized in the literature and the positions of methionine residues are highly conserved between species. A schematic of the CNBr cleavage sites on each α-chain is depicted in Figure 2.15 [275].

Rauterberg and Kuhn initially developed a protocol to cleave very large quantities (on the gram scale) of lyophilized collagen with CNBr in order to better characterize the resulting fragments [256]. With the more recent advances in technology such as mass spectrometry, characterization of peptides can be carried out on much smaller quantities. More recently, Rauterberg's protocols have been adapted to much smaller quantities of lyophilized collagen (0.5 – 1 mg) [257]. However, obtaining lyophilized cross-linked product may not be feasible as the process of dialysis prior to lyophilization would inevitably result in sample loss. With an already low cross-linking efficiency, sample loss could reduce our desired product below detection limits. Therefore it was necessary to further adapt the collagen-CNBr digestion protocol to allow cleavage in the cross-linking solutions directly. Because the reaction must take place in 70% formic acid, this involved significantly increasing the volume of the solution in which the cleavage took place. This resulted in a successful CNBr digest of collagen comparable to results obtained in other studies [257, 275-278] (Figure 2.16).

Figure 2.15 Schematic diagram of the CNBr peptides of type I collagen.

Horizontal lines indicate the location of methionine residues that react with CNBr. Image adapted from [275].

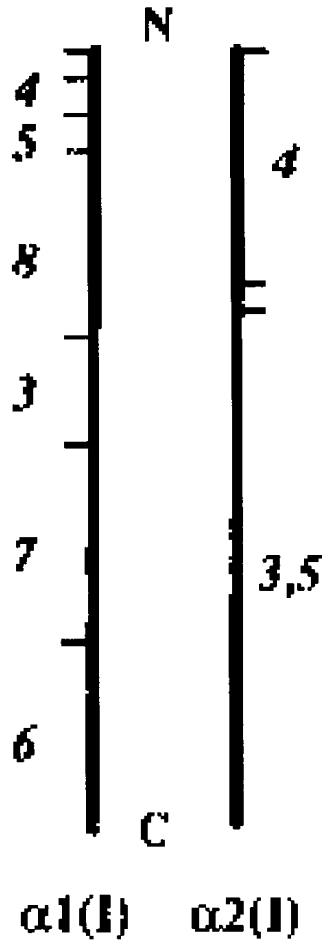
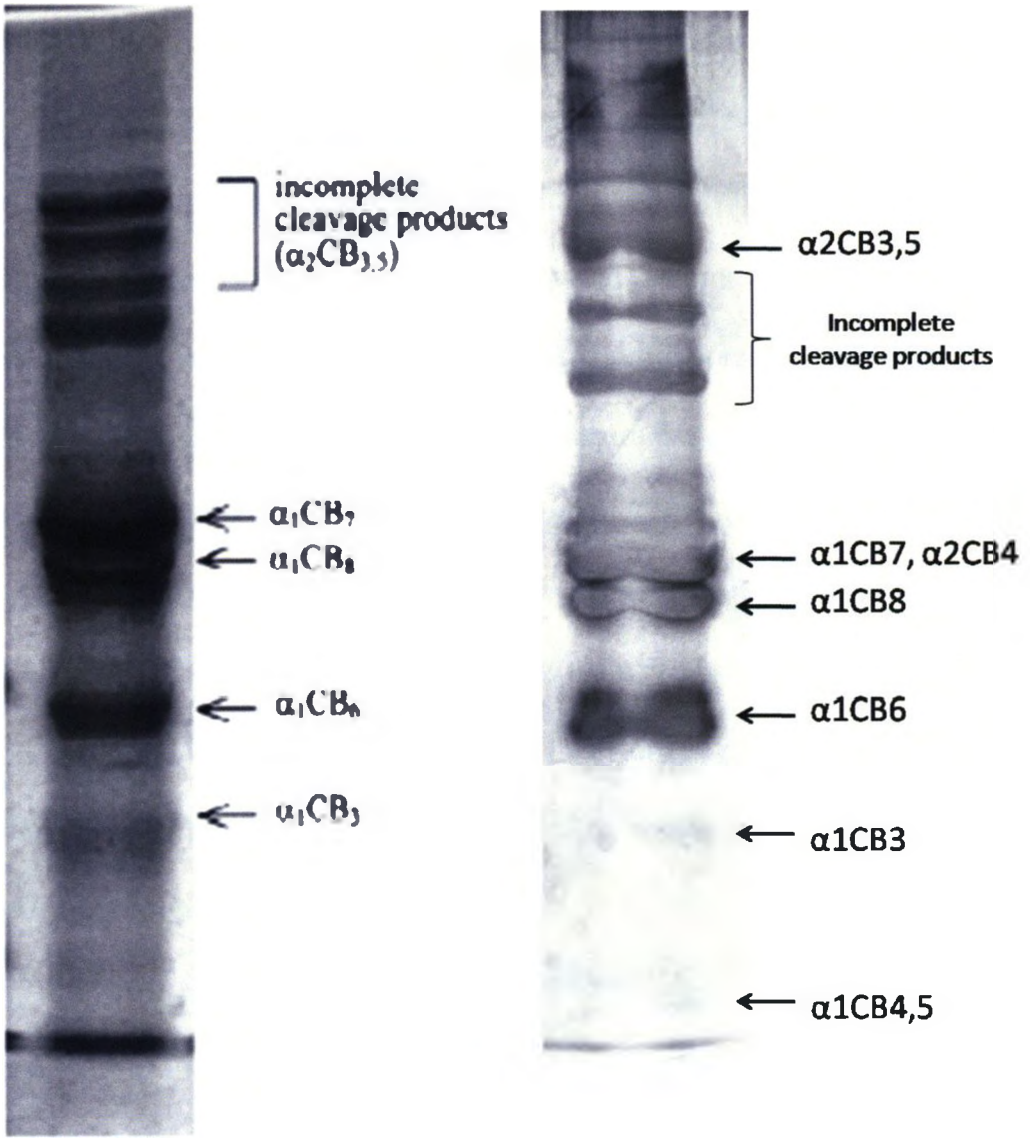

N
4
5
8
3
7
6
C
4
3,5
α1(I)
α2(I)

Figure 2.16 CNBr digestion profile of type I collagen.

The cyanogen bromide reaction was carried out using a protocol adapted from [257] that could be applied to BSP-collagen cross-linking reactions. Peptides were separated using 12% polyacrylamide gels. Left: A typical CNBr cleavage pattern upon survey of the literature [277]. Right: Our CNBr cleavage of collagen from cross-linking type solutions. Note: The image on the left was adapted from [277], where inaccuracies in labeling of the collagen CNBr fragments were corrected for based on more extensive studies involving rat tail collagen [278].


incomplete
cleavage products
($\alpha_2CB_{3,5}$)
α_1CB_7
α_1CB_8
α_1CB_6
α_1CB_3
α2CB3,5
Incomplete
cleavage products
α1CB7, α2CB4
α1CB8
α1CB6
α1CB3
α1CB4,5

Following CNBr cleavage, there are a number of ways to determine the collagen fragment involved in the cross-link. One of these methods is through mass spectrometry. The CNBr fragments of type I collagen have been well characterized using mass spectrometry, including the possible modifications imparted during the CNBr digestion process [257]. A cross-link could be identified by mass spectrometry through the appearance of a new MS peak among the collagen CNBr products. Subsequently, this cross-linked species could be analyzed through MS/MS to obtain the amino acid sequence of the collagen fragment involved in the cross-link. Pearson and colleagues performed a similar experiment with cross-linking reactions involving cytochrome c and ribonuclease A, where LC/ESI-MS/MS was used to select the cross-linked species of interest, impart collision-induced dissociation (CID), and subsequently identify the b and y ions from the fragmentation to determine the exact amino acids involved in the cross-link [279]. Similarly, Sinz and colleagues used this approach to identify the amino acids involved in the cross-link of calmodulin to a peptide of skeletal muscle myosin light chain kinase [218]. This type of study would provide very detailed information about the cross-link(s) between BSP and collagen at the amino acid level.

Because of the extensive modifications that can occur to collagen CNBr peptides as described in [257], it is possible that a cross-linked species may be difficult to detect by MS. In this case, the CNBr peptides could be separated by SDS-PAGE. Any additional band(s) observed after cross-linking could be excised from the gel for characterization through additional digestion and MS, or directly through MS/MS to sequence the peptide(s) involved.

Once it is determined exactly where BSP binds to collagen, structural studies could be carried out to elucidate the secondary structure formed, if any, upon binding. For instance, a triple-helical peptide of collagen mimicking the BSP-binding region could be synthesized, as Emsley and coworkers have for an integrin-binding region of collagen [215]. Subsequent X-ray crystallography, NMR, or circular dichroism experiments could be conducted to detect the structure of the complex.

Alternatively, information obtained about the cross-links could be used as constraints in molecular-modeling techniques to obtain detailed structural information. This type of study was recently performed by a group of researchers who obtained three distinct cross-links between the proteins latexin and carboxypeptidase A4 [247]. In this proof-of-principle study, mass spectrometry was used to locate the cross-links at the amino acid level, and this information was used as constraints in molecular-docking experiments that generate a predicted structure. This structure was then compared with the already-characterized crystal structure of the complex, and it was found to be in very close agreement [247]. We have generated 4 different cross-links between BSP and collagen from distinct locations on the BSP peptides. Constraints generated from the analysis of these cross-links could be used to accurately predict the structure of BSP upon binding collagen. This type of structure could then be mimicked in therapeutics such as bone implants to enhance calcification during the healing process.

Other potential studies that could be conducted involve the use of cross-linking reagents for purposes other than the location of binding regions. Some of the steady-state experiments conducted in our lab involve free BSP within reconstituted fibrillar collagen gels [45]. As protein binding events are in an equilibrium state, an increased amount of

BSP could be bound to collagen through chemical cross-linking to potentially increase HA-nucleation rates. Additionally, past experiments involving the implantation of collagenous plugs containing BSP into animal models of wound healing [29] could be repeated with the covalent attachment of BSP to collagen so as to maintain a locally high concentration of the HA-nucleating protein. Similar experiments have been conducted to increase angiogenesis within collagen matrices with the presence of VEGF. Covalently cross-linking VEGF to the collagen matrix significantly increased angiogenesis over samples where no cross-link was formed [280].

The development of the reagents and protocols that have led to a successful cross-link between BSP and collagen is an exciting and important step in the elucidation of the BSP-binding region of type I collagen. In the future, it is hoped that the cross-links formed can be extensively analyzed to glean valuable information about this important interaction, with the ultimate goal of improving bone healing therapeutics through enhanced biomineralization.

References

1. Lia Addadi, S.W. *Control and Design Principles in Biological Mineralization.* Angewandte Chemie International Edition in English, 1992. **31**(2): p. 153-169.
2. Lowenstam, A.H., Weiner, S. *On Biomineralization.* 1989, New York: Oxford University Press.
3. Lepot, K., Benzerara, K., Brown, G.E., and Philippot, P. *Microbially influenced formation of 2,724-million-year-old stromatolites.* Nat Geosci, 2008. **1**(2): p. 118-21.
4. Bengtson, S. *Palaeontology: a ghost with a bite.* Nature, 2006. **442**(7099): p. 146-7.
5. Gsponer, J., Futschik, M.E., Teichmann, S.A., and Babu, M.M. *Tight regulation of unstructured proteins: from transcript synthesis to protein degradation.* Science, 2008. **322**(5906): p. 1365-8.
6. Omelon, S.J. and Grynpas, M.D. *Relationships between polyphosphate chemistry, biochemistry and apatite biomineralization.* Chem Rev, 2008. **108**(11): p. 4694-715.
7. Fratzl, P., Fratzl-Zelman, N., Klaushofer, K., Vogl, G., and Koller, K. *Nucleation and growth of mineral crystals in bone studied by small-angle X-ray scattering.* Calcif Tissue Int, 1991. **48**(6): p. 407-13.
8. Weiner, S. and Price, P.A. *Disaggregation of bone into crystals.* Calcif Tissue Int, 1986. **39**(6): p. 365-75.
9. Robinson, R.A. *An electron-microscopic study of the crystalline inorganic component of bone and its relationship to the organic matrix.* J Bone Joint Surg Am, 1952. **34-A**(2): p. 389-435; passim.
10. Moradian-Oldak, J., Weiner, S., Addadi, L., Landis, W.J., and Traub, W. *Electron imaging and diffraction study of individual crystals of bone, mineralized tendon and synthetic carbonate apatite.* Connect Tissue Res, 1991. **25**(3-4): p. 219-28.
11. Burger, C., Zhou, H.W., Wang, H., Sics, I., Hsiao, B.S., Chu, B., Graham, L., and Glimcher, M.J. *Lateral packing of mineral crystals in bone collagen fibrils.* Biophys J, 2008. **95**(4): p. 1985-92.
12. Glimcher, M.J., Hodge, A.J., and Schmitt, F.O. *Macromolecular Aggregation States in Relation to Mineralization: the Collagen-Hydroxyapatite System as Studied in Vitro.* Proc Natl Acad Sci U S A, 1957. **43**(10): p. 860-7.
13. De Jong, A.S., Hak, T.J., and Van Duijn, P. *The dynamics of calcium phosphate precipitation studied with a new polyacrylamide steady state matrix-model: influence of pyrophosphate collagen and chondroitin sulfate.* Connect Tissue Res, 1980. **7**(2): p. 73-9.
14. Hunter, G.K., Nyburg, S.C., and Pritzker, K.P. *Hydroxyapatite formation in collagen, gelatin, and agarose gels.* Coll Relat Res, 1986. **6**(3): p. 229-38.
15. Glimcher, M.J. *Mechanism of calcification: role of collagen fibrils and collagen-phosphoprotein complexes in vitro and in vivo.* Anat Rec, 1989. **224**(2): p. 139-53.
16. Veis, A. and Perry, A. *The phosphoprotein of the dentin matrix.* Biochemistry, 1967. **6**(8): p. 2409-16.

17. Veis, A. *Phosphoproteins of dentin and bone. Do they have a role in matrix mineralization?, in The Chemistry and Biology of Mineralized Tissue*, ed. W.T. Butler. 1985, Birmingham, AL.: Ebsco Press. p. 170-7.
18. Hunter, G.K., Hauschka, P.V., Poole, A.R., Rosenberg, L.C., and Goldberg, H.A. *Nucleation and inhibition of hydroxyapatite formation by mineralized tissue proteins.* Biochem J, 1996. **317 (Pt 1)**: p. 59-64.
19. Bianco, P., Riminucci, M., Silvestrini, G., Bonucci, E., Termine, J.D., Fisher, L.W., and Robey, P.G. *Localization of bone sialoprotein (BSP) to Golgi and post-Golgi secretory structures in osteoblasts and to discrete sites in early bone matrix.* J Histochem Cytochem, 1993. **41**(2): p. 193-203.
20. Olsen, B.R., Reginato, A.M., and Wang, W. *Bone development.* Annu Rev Cell Dev Biol, 2000. **16**: p. 191-220.
21. Wagner, E.F. and Karsenty, G. *Genetic control of skeletal development.* Curr Opin Genet Dev, 2001. **11**(5): p. 527-32.
22. Provot, S. and Schipani, E. *Molecular mechanisms of endochondral bone development.* Biochem Biophys Res Commun, 2005. **328**(3): p. 658-65.
23. Brighton, C.T. and Hunt, R.M. *Early histologic and ultrastructural changes in microvessels of periosteal callus.* J Orthop Trauma, 1997. **11**(4): p. 244-53.
24. Brighton, C.T. and Hunt, R.M. *Early histological and ultrastructural changes in medullary fracture callus.* J Bone Joint Surg Am, 1991. **73**(6): p. 832-47.
25. Ham, A.W., and Harris, W.R. *Repair and Transplantation of Bone.* The biochemistry and physiology of bone. 1971, New York: Academic Press.
26. Brighton, C.T. and Hunt, R.M. *Histochemical localization of calcium in the fracture callus with potassium pyroantimonate. Possible role of chondrocyte mitochondrial calcium in callus calcification.* J Bone Joint Surg Am, 1986. **68**(5): p. 703-15.
27. Samartzis, D., Khanna, N., Shen, F.H., and An, H.S. *Update on bone morphogenetic proteins and their application in spine surgery.* J Am Coll Surg, 2005. **200**(2): p. 236-48.
28. Graf, H.L., Stoeva, S., Armbruster, F.P., Neuhaus, J., and Hilbig, H. *Effect of bone sialoprotein and collagen coating on cell attachment to TICER and pure titanium implant surfaces.* Int J Oral Maxillofac Surg, 2008. **37**(7): p. 634-40.
29. Wang, J., Zhou, H.Y., Salih, E., Xu, L., Wunderlich, L., Gu, X., Hofstaetter, J.G., Torres, M., and Glimcher, M.J. *Site-specific in vivo calcification and osteogenesis stimulated by bone sialoprotein.* Calcif Tissue Int, 2006. **79**(3): p. 179-89.
30. George, A. and Veis, A. *Phosphorylated proteins and control over apatite nucleation, crystal growth, and inhibition.* Chem Rev, 2008. **108**(11): p. 4670-93.
31. Anderson, H.C. *Matrix vesicles and calcification.* Curr Rheumatol Rep, 2003. **5**(3): p. 222-6.
32. Anderson, H.C. *Vesicles associated with calcification in the matrix of epiphyseal cartilage.* J Cell Biol, 1969. **41**(1): p. 59-72.
33. Ali, S.Y., Sajdera, S.W., and Anderson, H.C. *Isolation and characterization of calcifying matrix vesicles from epiphyseal cartilage.* Proc Natl Acad Sci U S A, 1970. **67**(3): p. 1513-20.
34. Kanabe, S., Hsu, H.H., Cecil, R.N., and Anderson, H.C. *Electron microscopic localization of adenosine triphosphate (ATP)-hydrolyzing activity in isolated*

matrix vesicles and reconstituted vesicles from calf cartilage. J Histochem Cytochem, 1983. **31**(4): p. 462-70.

35. Caswell, A.M., Ali, S.Y., and Russell, R.G. *Nucleoside triphosphate pyrophosphatase of rabbit matrix vesicles, a mechanism for the generation of inorganic pyrophosphate in epiphyseal cartilage.* Biochim Biophys Acta, 1987. **924**(2): p. 276-83.
36. Montessuit, C., Caverzasio, J., and Bonjour, J.P. *Characterization of a Pi transport system in cartilage matrix vesicles. Potential role in the calcification process.* J Biol Chem, 1991. **266**(27): p. 17791-7.
37. Boskey, A.L., Spevak, L., Paschalis, E., Doty, S.B., and Mckee, M.D. *Osteopontin deficiency increases mineral content and mineral crystallinity in mouse bone.* Calcif Tissue Int, 2002. **71**(2): p. 145-54.
38. Price, P.A., Toroian, D., and Lim, J.E. *Mineralization by Inhibitor Exclusion: THE Calcification of collagen with fetuin.* J Biol Chem, 2009. **284**(25): p. 17092-101.
39. Terkeltaub, R.A. *Inorganic pyrophosphate generation and disposition in pathophysiology.* Am J Physiol Cell Physiol, 2001. **281**(1): p. C1-C11.
40. Whyte, M.P. *Hypophosphatasia and the role of alkaline phosphatase in skeletal mineralization.* Endocr Rev, 1994. **15**(4): p. 439-61.
41. Murshed, M., Harmey, D., Millan, J.L., Mckee, M.D., and Karsenty, G. *Unique coexpression in osteoblasts of broadly expressed genes accounts for the spatial restriction of ECM mineralization to bone.* Genes Dev, 2005. **19**(9): p. 1093-104.
42. Posner, A.S., Betts, F. *Molecular Control of Tissue Mineralization in The Chemistry and Biology of Mineralized Connective Tissues.* 1981, Amsterdam: Elsevier.
43. Cuisinier, F.J.G., Steuer, P., Brisson, A., and Voegel, J.C. *High-Resolution Electron-Microscopy Study of Crystal-Growth Mechanisms in Chicken Bone Composites.* J Cryst Growth, 1995. **156**(4): p. 443-53.
44. Hunter, G.K. and Goldberg, H.A. *Nucleation of hydroxyapatite by bone sialoprotein.* Proc Natl Acad Sci U S A, 1993. **90**(18): p. 8562-5.
45. Baht, G.S., Hunter, G.K., and Goldberg, H.A. *Bone sialoprotein-collagen interaction promotes hydroxyapatite nucleation.* Matrix Biol, 2008. **27**(7):600-8
46. Gerstenfeld, L.C., Chipman, S.D., Kelly, C.M., Hodgens, K.J., Lee, D.D., and Landis, W.J. *Collagen expression, ultrastructural assembly, and mineralization in cultures of chicken embryo osteoblasts.* J Cell Biol, 1988. **106**(3): p. 979-89.
47. Kratochwil, K., Von Der Mark, K., Kollar, E.J., Jaenisch, R., Mooslehner, K., Schwarz, M., Haase, K., Gmachl, I., and Harbers, K. *Retrovirus-induced insertional mutation in Mov13 mice affects collagen I expression in a tissue-specific manner.* Cell, 1989. **57**(5): p. 807-16.
48. Prockop, D.J. and Kivirikko, K.I. *Collagens: molecular biology, diseases, and potentials for therapy.* Annu Rev Biochem, 1995. **64**: p. 403-34.
49. Orgel, J.P., Irving, T.C., Miller, A., and Wess, T.J. *Microfibrillar structure of type I collagen in situ.* Proc Natl Acad Sci U S A, 2006. **103**(24): p. 9001-5.
50. Traub, W. and Piez, K.A. *The chemistry and structure of collagen.* Adv Protein Chem, 1971. **25**: p. 243-352.

51. Hulmes, D.J. and Miller, A. *Molecular packing in collagen.* Nature, 1981. **293**(5829): p. 239-4.
52. Hulmes, D.J., Miller, A., Parry, D.A., Piez, K.A., and Woodhead-Galloway, J. *Analysis of the primary structure of collagen for the origins of molecular packing.* J Mol Biol, 1973. **79**(1): p. 137-48.
53. Berg, R.A. and Prockop, D.J. *The thermal transition of a non-hydroxylated form of collagen. Evidence for a role for hydroxyproline in stabilizing the triple-helix of collagen.* Biochem Biophys Res Commun, 1973. **52**(1): p. 115-20.
54. Berg, R.A. and Prockop, D.J. *Purification of (14C) protocollagen and its hydroxylation by prolyl-hydroxylase.* Biochemistry, 1973. **12**(18): p. 3395-401.
55. Bruckner, P. and Prockop, D.J. *Proteolytic enzymes as probes for the triple-helical conformation of procollagen.* Anal Biochem, 1981. **110**(2): p. 360-8.
56. Ramshaw, J.A., Shah, N.K., and Brodsky, B. *Gly-X-Y tripeptide frequencies in collagen: a context for host-guest triple-helical peptides.* J Struct Biol, 1998. **122**(1-2): p. 86-91.
57. Wess, T.J., Hammersley, A.P., Wess, L., and Miller, A. *A consensus model for molecular packing of type I collagen.* J Struct Biol, 1998. **122**(1-2): p. 92-100.
58. Hodge, A.J., Petruska, J.A. *Recent Studies with the Electron Microscope on Ordered Aggregates of the Tropocollagen Macromolecule. in Aspects of Protein Structure*, ed. G.N. Ramachandran. 1963, New York: Academic Press. p. 289.
59. Berthet-Colominas, C., Miller, A., and White, S.W. *Structural study of the calcifying collagen in turkey leg tendons.* J Mol Biol, 1979. **134**(3): p. 431-45.
60. Arsenault, A.L. *Crystal-collagen relationships in calcified turkey leg tendons visualized by selected-area dark field electron microscopy.* Calcif Tissue Int, 1988. **43**(4): p. 202-12.
61. Landis, W.J., Song, M.J., Leith, A., Mcewen, L., and Mcewen, B.F. *Mineral and organic matrix interaction in normally calcifying tendon visualized in three dimensions by high-voltage electron microscopic tomography and graphic image reconstruction.* J Struct Biol, 1993. **110**(1): p. 39-54.
62. Vuorio, E. and De Crombrugghe, B. *The family of collagen genes.* Annu Rev Biochem, 1990. **59**: p. 837-72.
63. Ramirez, F. and Di Liberto, M. *Complex and diversified regulatory programs control the expression of vertebrate collagen genes.* Faseb J, 1990. **4**(6): p. 1616-23.
64. De Wet, W., Bernard, M., Benson-Chanda, V., Chu, M.L., Dickson, L., Weil, D., and Ramirez, F. *Organization of the human pro-alpha 2(I) collagen gene.* J Biol Chem, 1987. **262**(33): p. 16032-6.
65. Bornstein, P. and Sage, H. *Regulation of collagen gene expression.* Prog Nucleic Acid Res Mol Biol, 1989. **37**: p. 67-106.
66. Slack, J.L., Liska, D.J., and Bornstein, P. *Regulation of expression of the type I collagen genes.* Am J Med Genet, 1993. **45**(2): p. 140-51.
67. Fessler, J.H., Doege, K.J., Duncan, K.G., and Fessler, L.I. *Biosynthesis of collagen.* J Cell Biochem, 1985. **28**(1): p. 31-7.
68. Von Der Mark, K. *Structure, Biosynthesis, and Gene Regulation of Collagens in Cartilage and Bone. in Dynamics of Bone and Cartilage Metabolism.* 1999, New York: Academic Press. p. 3.

69. Nagata, K. *Hsp47: a collagen-specific molecular chaperone.* Trends Biochem Sci, 1996. **21**(1): p. 22-6.
70. Pihlajaniemi, T., Myllyla, R., and Kivirikko, K.I. *Prolyl 4-hydroxylase and its role in collagen synthesis.* J Hepatol, 1991. **13 Suppl 3**: p. S2-7.
71. Kivirikko, K.I. and Pihlajaniemi, T. *Collagen hydroxylases and the protein disulfide isomerase subunit of prolyl 4-hydroxylases.* Adv Enzymol Relat Areas Mol Biol, 1998. **72**: p. 325-98.
72. Kivirikko, K.I. and Myllyharju, J. *Prolyl 4-hydroxylases and their protein disulfide isomerase subunit.* Matrix Biol, 1998. **16**(7): p. 357-68.
73. Myllyla, R., Pihlajaniemi, T., Pajunen, L., Turpeenniemi-Hujanen, T., and Kivirikko, K.I. *Molecular cloning of chick lysyl hydroxylase. Little homology in primary structure to the two types of subunit of prolyl 4-hydroxylase.* J Biol Chem, 1991. **266**(5): p. 2805-10.
74. Beck, K., Boswell, B.A., Ridgway, C.C., and Bachinger, H.P. *Triple helix formation of procollagen type I can occur at the rough endoplasmic reticulum membrane.* J Biol Chem, 1996. **271**(35): p. 21566-73.
75. Engel, J. and Prockop, D.J. *The zipper-like folding of collagen triple helices and the effects of mutations that disrupt the zipper.* Annu Rev Biophys Biophys Chem, 1991. **20**: p. 137-52.
76. Bulleid, N.J. *Novel approach to study the initial events in the folding and assembly of procollagen.* Semin Cell Dev Biol, 1996. **7**(5): p. 667-72.
77. Bulleid, N.J., Dalley, J.A., and Lees, J.F. *The C-propeptide domain of procollagen can be replaced with a transmembrane domain without affecting trimer formation or collagen triple helix folding during biosynthesis.* Embo J, 1997. **16**(22): p. 6694-701.
78. Doege, K.J. and Fessler, J.H. *Folding of carboxyl domain and assembly of procollagen I.* J Biol Chem, 1986. **261**(19): p. 8924-35.
79. Wilson, R., Lees, J.F., and Bulleid, N.J. *Protein disulfide isomerase acts as a molecular chaperone during the assembly of procollagen.* J Biol Chem, 1998. **273**(16): p. 9637-43.
80. Mclaughlin, S.H. and Bulleid, N.J. *Thiol-independent interaction of protein disulphide isomerase with type X collagen during intra-cellular folding and assembly.* Biochem J, 1998. **331 (Pt 3)**: p. 793-800.
81. Prockop, D.J., Sieron, A.L., and Li, S.W. *Procollagen N-proteinase and procollagen C-proteinase. Two unusual metalloproteinases that are essential for procollagen processing probably have important roles in development and cell signaling.* Matrix Biol, 1998. **16**(7): p. 399-408.
82. Rawlings, N.D. and Barrett, A.J. *Evolutionary families of metallopeptidases.* Methods Enzymol, 1995. **248**: p. 183-228.
83. Paglia, L.M., Wiestner, M., Duchene, M., Ouellette, L.A., Horlein, D., Martin, G.R., and Muller, P.K. *Effects of procollagen peptides on the translation of type II collagen messenger ribonucleic acid and on collagen biosynthesis in chondrocytes.* Biochemistry, 1981. **20**(12): p. 3523-7.
84. Paglia, L., Wilczek, J., De Leon, L.D., Martin, G.R., Horlein, D., and Muller, P. *Inhibition of procollagen cell-free synthesis by amino-terminal extension peptides.* Biochemistry, 1979. **18**(22): p. 5030-4.

85. Wu, C.H., Donovan, C.B., and Wu, G.Y. *Evidence for pretranslational regulation of collagen synthesis by procollagen propeptides.* J Biol Chem, 1986. **261**(23): p. 10482-4.
86. Fisher, L.W., Torchia, D.A., Fohr, B., Young, M.F., and Fedarko, N.S. *Flexible structures of SIBLING proteins, bone sialoprotein, and osteopontin.* Biochem Biophys Res Commun, 2001. **280**(2): p. 460-5.
87. Rowe, P.S., De Zoysa, P.A., Dong, R., Wang, H.R., White, K.E., Econs, M.J., and Oudet, C.L. *MEPE, a new gene expressed in bone marrow and tumors causing osteomalacia.* Genomics, 2000. **67**(1): p. 54-68.
88. Ganss, B., Kim, R.H., and Sodek, J. *Bone sialoprotein.* Crit Rev Oral Biol Med, 1999. **10**(1): p. 79-98.
89. Macdougall, M. *Refined mapping of the human dentin sialophosphoprotein (DSPP) gene within the critical dentinogenesis imperfecta type II and dentin dysplasia type II loci.* Eur J Oral Sci, 1998. **106 Suppl 1**: p. 227-33.
90. Macdougall, M., Gu, T.T., and Simmons, D. *Dentin matrix protein-1, a candidate gene for dentinogenesis imperfecta.* Connect Tissue Res, 1996. **35**(1-4): p. 267-72.
91. Malaval, L., Wade-Gueye, N.M., Boudiffa, M., Fei, J., Zirngibl, R., Chen, F., Laroche, N., Roux, J.P., Burt-Pichat, B., Duboeuf, F., Boivin, G., Jurdic, P., Lafage-Proust, M.H., Amedee, J., Vico, L., Rossant, J., and Aubin, J.E. *Bone sialoprotein plays a functional role in bone formation and osteoclastogenesis.* J Exp Med, 2008. **205**(5): p. 1145-53.
92. Gowen, L.C., Petersen, D.N., Mansolf, A.L., Qi, H., Stock, J.L., Tkalcevic, G.T., Simmons, H.A., Crawford, D.T., Chidsey-Frink, K.L., Ke, H.Z., Mcneish, J.D., and Brown, T.A. *Targeted disruption of the osteoblast/osteocyte factor 45 gene (OF45) results in increased bone formation and bone mass.* J Biol Chem, 2003. **278**(3): p. 1998-2007.
93. Feng, J.Q., Huang, H., Lu, Y., Ye, L., Xie, Y., Tsutsui, T.W., Kunieda, T., Castranio, T., Scott, G., Bonewald, L.B., and Mishina, Y. *The Dentin matrix protein 1 (Dmp1) is specifically expressed in mineralized, but not soft, tissues during development.* J Dent Res, 2003. **82**(10): p. 776-80.
94. Ling, Y., Rios, H.F., Myers, E.R., Lu, Y., Feng, J.Q., and Boskey, A.L. *DMP1 depletion decreases bone mineralization in vivo: an FTIR imaging analysis.* J Bone Miner Res, 2005. **20**(12): p. 2169-77.
95. Ye, L., Mishina, Y., Chen, D., Huang, H., Dallas, S.L., Dallas, M.R., Sivakumar, P., Kunieda, T., Tsutsui, T.W., Boskey, A., Bonewald, L.F., and Feng, J.Q. *Dmp1-deficient mice display severe defects in cartilage formation responsible for a chondrodysplasia-like phenotype.* J Biol Chem, 2005. **280**(7): p. 6197-203.
96. Sreenath, T., Thyagarajan, T., Hall, B., Longenecker, G., D'souza, R., Hong, S., Wright, J.T., Macdougall, M., Sauk, J., and Kulkarni, A.B. *Dentin sialophosphoprotein knockout mouse teeth display widened predentin zone and develop defective dentin mineralization similar to human dentinogenesis imperfecta type III.* J Biol Chem, 2003. **278**(27): p. 24874-80.
97. George, A., Sabsay, B., Simonian, P.A., and Veis, A. *Characterization of a novel dentin matrix acidic phosphoprotein. Implications for induction of biomineralization.* J Biol Chem, 1993. **268**(17): p. 12624-30.

98. He, G., Dahl, T., Veis, A., and George, A. *Dentin matrix protein 1 initiates hydroxyapatite formation in vitro.* Connect Tissue Res, 2003. **44 Suppl 1**: p. 240-5.
99. He, G., Dahl, T., Veis, A., and George, A. *Nucleation of apatite crystals in vitro by self-assembled dentin matrix protein 1.* Nat Mater, 2003. **2**(8): p. 552-8.
100. He, G., Gajjeraman, S., Schultz, D., Cookson, D., Qin, C., Butler, W.T., Hao, J., and George, A. *Spatially and temporally controlled biomineralization is facilitated by interaction between self-assembled dentin matrix protein 1 and calcium phosphate nuclei in solution.* Biochemistry, 2005. **44**(49): p. 16140-8.
101. Cross, K.J., Huq, N.L., and Reynolds, E.C. *Protein dynamics of bovine dentin phosphophoryn.* J Pept Res, 2005. **66**(2): p. 59-67.
102. Long, J.R., Shaw, W.J., Stayton, P.S., and Drobny, G.P. *Structure and dynamics of hydrated statherin on hydroxyapatite as determined by solid-state NMR.* Biochemistry, 2001. **40**(51): p. 15451-5.
103. Goobes, G., Goobes, R., Schueler-Furman, O., Baker, D., Stayton, P.S., and Drobny, G.P. *Folding of the C-terminal bacterial binding domain in statherin upon adsorption onto hydroxyapatite crystals.* Proc Natl Acad Sci U S A, 2006. **103**(44): p. 16083-8.
104. Fujisawa, R. and Kuboki, Y. *Affinity of bone sialoprotein and several other bone and dentin acidic proteins to collagen fibrils.* Calcif Tissue Int, 1992. **51**(6): p. 438-42.
105. Fujisawa, R., Nodasaka, Y., and Kuboki, Y. *Further characterization of interaction between bone sialoprotein (BSP) and collagen.* Calcif Tissue Int, 1995. **56**(2): p. 140-4.
106. Lee, J.Y., Choo, J.E., Choi, Y.S., Park, J.B., Min, D.S., Lee, S.J., Rhyu, H.K., Jo, I.H., Chung, C.P., and Park, Y.J. *Assembly of collagen-binding peptide with collagen as a bioactive scaffold for osteogenesis in vitro and in vivo.* Biomaterials, 2007. **28**(29): p. 4257-67.
107. He, G. and George, A. *Dentin matrix protein 1 immobilized on type I collagen fibrils facilitates apatite deposition in vitro.* J Biol Chem, 2004. **279**(12): p. 11649-56.
108. Traub, W., Arad, T., and Weiner, S. *Origin of mineral crystal growth in collagen fibrils.* Matrix, 1992. **12**(4): p. 251-5.
109. George, A. and Hao, J. *Role of phosphophoryn in dentin mineralization.* Cells Tissues Organs, 2005. **181**(3-4): p. 232-40.
110. Kaartinen, M.T., Pirhonen, A., Linnala-Kankkunen, A., and Maenpaa, P.H. *Cross-linking of osteopontin by tissue transglutaminase increases its collagen binding properties.* J Biol Chem, 1999. **274**(3): p. 1729-35.
111. Kaartinen, M.T., Sun, W., Kaipatur, N., and Mckee, M.D. *Transglutaminase crosslinking of SIBLING proteins in teeth.* J Dent Res, 2005. **84**(7): p. 607-12.
112. Keykhosravani, M., Doherty-Kirby, A., Zhang, C., Brewer, D., Goldberg, H.A., Hunter, G.K., and Lajoie, G. *Comprehensive identification of post-translational modifications of rat bone osteopontin by mass spectrometry.* Biochemistry, 2005. **44**(18): p. 6990-7003.
113. Tartaix, P.H., Doulaverakis, M., George, A., Fisher, L.W., Butler, W.T., Qin, C., Salih, E., Tan, M., Fujimoto, Y., Spevak, L., and Boskey, A.L. *In vitro effects of*

dentin matrix protein-1 on hydroxyapatite formation provide insights into in vivo functions. J Biol Chem, 2004. **279**(18): p. 18115-20.

114. Stetler-Stevenson, W.G. and Veis, A. *Bovine dentin phosphophoryn: composition and molecular weight.* Biochemistry, 1983. **22**(18): p. 4326-35.
115. Huq, N.L., Cross, K.J., Talbo, G.H., Riley, P.F., Loganathan, A., Crossley, M.A., Perich, J.W., and Reynolds, E.C. *N-terminal sequence analysis of bovine dentin phosphophoryn after conversion of phosphoseryl to S-propylcysteinyl residues.* J Dent Res, 2000. **79**(11): p. 1914-9.
116. Saito, T., Arsenault, A.L., Yamauchi, M., Kuboki, Y., and Crenshaw, M.A. *Mineral induction by immobilized phosphoproteins.* Bone, 1997. **21**(4): p. 305-11.
117. Torres-Quintana, M.A., Lecolle, S., and Goldberg, M. *Effects of inositol hexasulphate, a casein kinase inhibitor, on dentine phosphorylated proteins in organ culture of mouse tooth germs.* Arch Oral Biol, 1998. **43**(8): p. 597-610.
118. Langdon, A., Wignall, G.R., Rogers, K., Sorensen, E.S., Denstedt, J., Grohe, B., Goldberg, H.A., and Hunter, G.K. *Kinetics of calcium oxalate crystal growth in the presence of osteopontin isoforms: an analysis by scanning confocal interference microcopy.* Calcif Tissue Int, 2009. **84**(3): p. 240-8.
119. Andrews, A.T., Herring, G.M., and Kent, P.W. *Some studies on the composition of bovine cortical-bone sialoprotein.* Biochem J, 1967. **104**(3): p. 705-15.
120. Fisher, L.W., Whitson, S.W., Avioli, L.V., and Termine, J.D. *Matrix sialoprotein of developing bone.* J Biol Chem, 1983. **258**(20): p. 12723-7.
121. Fisher, L.W., Hawkins, G.R., Tuross, N., and Termine, J.D. *Purification and partial characterization of small proteoglycans I and II, bone sialoproteins I and II, and osteonectin from the mineral compartment of developing human bone.* J Biol Chem, 1987. **262**(20): p. 9702-8.
122. Zhang, Q., Domenicucci, C., Goldberg, H.A., Wrana, J.L., and Sodek, J. *Characterization of fetal porcine bone sialoproteins, secreted phosphoprotein I (SPPI, osteopontin), bone sialoprotein, and a 23-kDa glycoprotein. Demonstration that the 23-kDa glycoprotein is derived from the carboxyl terminus of SPPI.* J Biol Chem, 1990. **265**(13): p. 7583-9.
123. Gorski, J.P. and Shimizu, K. *Isolation of new phosphorylated glycoprotein from mineralized phase of bone that exhibits limited homology to adhesive protein osteopontin.* J Biol Chem, 1988. **263**(31): p. 15938-45.
124. Kinne, R.W. and Fisher, L.W. *Keratan sulfate proteoglycan in rabbit compact bone is bone sialoprotein II.* J Biol Chem, 1987. **262**(21): p. 10206-11.
125. Shintani, S., Kamakura, N., Kobata, M., Toyosawa, S., Onishi, T., Sato, A., Kawasaki, K., Weiss, K.M., and Ooshima, T. *Identification and characterization of integrin-binding sialoprotein (IBSP) genes in reptile and amphibian.* Gene, 2008. **424**(1-2): p. 11-7.
126. Chen, J.K., Shapiro, H.S., Wrana, J.L., Reimers, S., Heersche, J.N., and Sodek, J. *Localization of bone sialoprotein (BSP) expression to sites of mineralized tissue formation in fetal rat tissues by in situ hybridization.* Matrix, 1991. **11**(2): p. 133-43.
127. Chen, J., Shapiro, H.S., and Sodek, J. *Development expression of bone sialoprotein mRNA in rat mineralized connective tissues.* J Bone Miner Res, 1992. **7**(8): p. 987-97.

128. Oldberg, A., Franzen, A., and Heinegard, D. *The primary structure of a cell-binding bone sialoprotein.* J Biol Chem, 1988. **263**(36): p. 19430-2.
129. Tye, C.E., Hunter, G.K., and Goldberg, H.A. *Identification of the type I collagen-binding domain of bone sialoprotein and characterization of the mechanism of interaction.* J Biol Chem, 2005. **280**(14): p. 13487-92.
130. Fisher, L.W., Mcbride, O.W., Termine, J.D., and Young, M.F. *Human bone sialoprotein. Deduced protein sequence and chromosomal localization.* J Biol Chem, 1990. **265**(4): p. 2347-51.
131. Kerr, J.M., Fisher, L.W., Termine, J.D., Wang, M.G., Mcbride, O.W., and Young, M.F. *The human bone sialoprotein gene (IBSP): genomic localization and characterization.* Genomics, 1993. **17**(2): p. 408-15.
132. Macdougall, M., Dupont, B.R., Simmons, D., and Leach, R.J. *Assignment of DMP1 to human chromosome 4 band q21 by in situ hybridization.* Cytogenet Cell Genet, 1996. **74**(3): p. 189.
133. Kim, R.H., Shapiro, H.S., Li, J.J., Wrana, J.L., and Sodek, J. *Characterization of the human bone sialoprotein (BSP) gene and its promoter sequence.* Matrix Biol, 1994. **14**(1): p. 31-40.
134. Yang, R. and Gerstenfeld, L.C. *Signal transduction pathways mediating parathyroid hormone stimulation of bone sialoprotein gene expression in osteoblasts.* J Biol Chem, 1996. **271**(47): p. 29839-46.
135. Li, J.J. and Sodek, J. *Cloning and characterization of the rat bone sialoprotein gene promoter.* Biochem J, 1993. **289 (Pt 3)**: p. 625-9.
136. Kim, R.H., Li, J.J., Ogata, Y., Yamauchi, M., Freedman, L.P., and Sodek, J. *Identification of a vitamin D3-response element that overlaps a unique inverted TATA box in the rat bone sialoprotein gene.* Biochem J, 1996. **318 (Pt 1)**: p. 219-26.
137. Oldberg, A., Jirskog-Hed, B., Axelsson, S., and Heinegard, D. *Regulation of bone sialoprotein mRNA by steroid hormones.* J Cell Biol, 1989. **109**(6 Pt 1): p. 3183-6.
138. Kasugai, S., Todescan, R., Jr., Nagata, T., Yao, K.L., Butler, W.T., and Sodek, J. *Expression of bone matrix proteins associated with mineralized tissue formation by adult rat bone marrow cells in vitro: inductive effects of dexamethasone on the osteoblastic phenotype.* J Cell Physiol, 1991. **147**(1): p. 111-20.
139. Yao, K.L., Todescan, R., Jr., and Sodek, J. *Temporal changes in matrix protein synthesis and mRNA expression during mineralized tissue formation by adult rat bone marrow cells in culture.* J Bone Miner Res, 1994. **9**(2): p. 231-40.
140. Rickard, D.J., Sullivan, T.A., Shenker, B.J., Leboy, P.S., and Kazhdan, I. *Induction of rapid osteoblast differentiation in rat bone marrow stromal cell cultures by dexamethasone and BMP-2.* Dev Biol, 1994. **161**(1): p. 218-28.
141. Sodek, J., Li, I.W.S., Li, H., Bellows, C.G., Mcculloch, C.A.G., Tenenbaum, H.C., and Ellen, R.P. *The Role of Tgf-Beta and Bmp-7 in Regenerating Bone and Soft-Tissues.* Mat Sci Eng C-Bio S, 1994. **2**(1-2): p. 19-26.
142. Mcculloch, C.A. and Tenenbaum, H.C. *Dexamethasone induces proliferation and terminal differentiation of osteogenic cells in tissue culture.* Anat Rec, 1986. **215**(4): p. 397-402.

143. Maniatopoulos, C., Sodek, J., and Melcher, A.H. *Bone formation in vitro by stromal cells obtained from bone marrow of young adult rats.* Cell Tissue Res, 1988. **254**(2): p. 317-30.
144. Bellows, C.G., Heersche, J.N., and Aubin, J.E. *Determination of the capacity for proliferation and differentiation of osteoprogenitor cells in the presence and absence of dexamethasone.* Dev Biol, 1990. **140**(1): p. 132-8.
145. Reddi, A.H. *Regulation of cartilage and bone differentiation by bone morphogenetic proteins.* Curr Opin Cell Biol, 1992. **4**(5): p. 850-5.
146. Kato, N., Nakayama, Y., Nakajima, Y., Samoto, H., Saito, R., Yamanouchi, F., Masunaga, H., Shimizu, E., and Ogata, Y. *Regulation of bone sialoprotein (BSP) gene transcription by lipopolysaccharide.* J Cell Biochem, 2006. **97**(2): p. 368-79.
147. Klein-Nulend, J., Roelofsen, J., Semeins, C.M., Bronckers, A.L., and Burger, E.H. *Mechanical stimulation of osteopontin mRNA expression and synthesis in bone cell cultures.* J Cell Physiol, 1997. **170**(2): p. 174-81.
148. Bikle, D.D., Halloran, B.P., Cone, C.M., Globus, R.K., and Morey-Holton, E. *The effects of simulated weightlessness on bone maturation.* Endocrinology, 1987. **120**(2): p. 678-84.
149. Vico, L., Chappard, D., Alexandre, C., Palle, S., Minaire, P., Riffat, G., Morukov, B., and Rakhmanov, S. *Effects of a 120 day period of bed-rest on bone mass and bone cell activities in man: attempts at countermeasure.* Bone Miner, 1987. **2**(5): p. 383-94.
150. Shimizu, E., Matsuda-Honjyo, Y., Samoto, H., Saito, R., Nakajima, Y., Nakayama, Y., Kato, N., Yamazaki, M., and Ogata, Y. *Static magnetic fields-induced bone sialoprotein (BSP) expression is mediated through FGF2 response element and pituitary-specific transcription factor-1 motif.* J Cell Biochem, 2004. **91**(6): p. 1183-96.
151. Nakajima, Y., Kato, N., Nakayama, Y., Kim, D.S., Takai, H., Arai, M., Saito, R., Samoto, H., Shimizu, E., and Ogata, Y. *Effect of chlorpromazine on bone sialoprotein (BSP) gene transcription.* J Cell Biochem, 2006. **97**(6): p. 1198-206.
152. Bianco, P., Fisher, L.W., Young, M.F., Termine, J.D., and Robey, P.G. *Expression of bone sialoprotein (BSP) in developing human tissues.* Calcif Tissue Int, 1991. **49**(6): p. 421-6.
153. Shapiro, H.S., Chen, J., Wrana, J.L., Zhang, Q., Blum, M., and Sodek, J. *Characterization of porcine bone sialoprotein: primary structure and cellular expression.* Matrix, 1993. **13**(6): p. 431-40.
154. Arai, N., Ohya, K., Kasugai, S., Shimokawa, H., Ohida, S., Ogura, H., and Amagasa, T. *Expression of bone sialoprotein mRNA during bone formation and resorption induced by colchicine in rat tibial bone marrow cavity.* J Bone Miner Res, 1995. **10**(8): p. 1209-17.
155. Macneil, R.L., Sheng, N., Strayhorn, C., Fisher, L.W., and Somerman, M.J. *Bone sialoprotein is localized to the root surface during cementogenesis.* J Bone Miner Res, 1994. **9**(10): p. 1597-606.
156. Chenu, C. and Delmas, P.D. *Platelets contribute to circulating levels of bone sialoprotein in human.* J Bone Miner Res, 1992. **7**(1): p. 47-54.
157. Chen, J., Zhang, Q., Mcculloch, C.A., and Sodek, J. *Immunohistochemical localization of bone sialoprotein in foetal porcine bone tissues: comparisons with*

secreted phosphoprotein 1 (SPP-1, osteopontin) and SPARC (osteonectin). Histochem J, 1991. **23**(6): p. 281-9.

158. Horton, M.A., Nesbit, M.A., and Helfrich, M.H. *Interaction of osteopontin with osteoclast integrins.* Ann N Y Acad Sci, 1995. **760**: p. 190-200.
159. Seibel, M.J., Woitge, H.W., Pecherstorfer, M., Karmatschek, M., Horn, E., Ludwig, H., Armbruster, F.P., and Ziegler, R. *Serum immunoreactive bone sialoprotein as a new marker of bone turnover in metabolic and malignant bone disease.* J Clin Endocrinol Metab, 1996. **81**(9): p. 3289-94.
160. Riminucci, M., Silvestrini, G., Bonucci, E., Fisher, L.W., Gehron Robey, P., and Bianco, P. *The anatomy of bone sialoprotein immunoreactive sites in bone as revealed by combined ultrastructural histochemistry and immunohistochemistry.* Calcif Tissue Int, 1995. **57**(4): p. 277-84.
161. Chen, J., Mckee, M.D., Nanci, A., and Sodek, J. *Bone sialoprotein mRNA expression and ultrastructural localization in fetal porcine calvarial bone: comparisons with osteopontin.* Histochem J, 1994. **26**(1): p. 67-78.
162. Ingram, R.T., Clarke, B.L., Fisher, L.W., and Fitzpatrick, L.A. *Distribution of noncollagenous proteins in the matrix of adult human bone: evidence of anatomic and functional heterogeneity.* J Bone Miner Res, 1993. **8**(9): p. 1019-29.
163. Kasugai, S., Nagata, T., and Sodek, J. *Temporal studies on the tissue compartmentalization of bone sialoprotein (BSP), osteopontin (OPN), and SPARC protein during bone formation in vitro.* J Cell Physiol, 1992. **152**(3): p. 467-77.
164. Bellahcene, A., Maloujahmoum, N., Fisher, L.W., Pastorino, H., Tagliabue, E., Menard, S., and Castronovo, V. *Expression of bone sialoprotein in human lung cancer.* Calcif Tissue Int, 1997. **61**(3): p. 183-8.
165. Bellahcene, A., Merville, M.P., and Castronovo, V. *Expression of bone sialoprotein, a bone matrix protein, in human breast cancer.* Cancer Res, 1994. **54**(11): p. 2823-6.
166. Bellahcene, A., Albert, V., Pollina, L., Basolo, F., Fisher, L.W., and Castronovo, V. *Ectopic expression of bone sialoprotein in human thyroid cancer.* Thyroid, 1998. **8**(8): p. 637-41.
167. Waltregny, D., Bellahcene, A., Van Riet, I., Fisher, L.W., Young, M., Fernandez, P., Dewe, W., De Leval, J., and Castronovo, V. *Prognostic value of bone sialoprotein expression in clinically localized human prostate cancer.* J Natl Cancer Inst, 1998. **90**(13): p. 1000-8.
168. Fedarko, N.S., Fohr, B., Robey, P.G., Young, M.F., and Fisher, L.W. *Factor H binding to bone sialoprotein and osteopontin enables tumor cell evasion of complement-mediated attack.* J Biol Chem, 2000. **275**(22): p. 16666-72.
169. Karadag, A., Ogbureke, K.U., Fedarko, N.S., and Fisher, L.W. *Bone sialoprotein, matrix metalloproteinase 2, and alpha(v)beta3 integrin in osteotropic cancer cell invasion.* J Natl Cancer Inst, 2004. **96**(12): p. 956-65.
170. Bellahcene, A., Bonjean, K., Fohr, B., Fedarko, N.S., Robey, F.A., Young, M.F., Fisher, L.W., and Castronovo, V. *Bone sialoprotein mediates human endothelial cell attachment and migration and promotes angiogenesis.* Circ Res, 2000. **86**(8): p. 885-91.

171. Bellahcene, A., Menard, S., Bufalino, R., Moreau, L., and Castronovo, V. *Expression of bone sialoprotein in primary human breast cancer is associated with poor survival.* Int J Cancer, 1996. **69**(4): p. 350-3.
172. Bosse, A., Wuisman, P., Jones, D.B., and Schwarz, K. *Noncollagenous proteins in heterotopic ossification. Immunohistochemical analysis in 15 paraplegies.* Acta Orthop Scand, 1993. **64**(6): p. 634-8.
173. Mckee, M.D., Giachelli, C.M., and Nanci, A. *Matrix-mineral relationships in calcifying human atherosclerotic plaque: ultrastructural immunodetection of osteopontin and bone sialoprotein at calcification sites (abstract).* J Bone Miner Res, 1996. **11**: p. S330.
174. Wuttke, M., Muller, S., Nitsche, D.P., Paulsson, M., Hanisch, F.G., and Maurer, P. *Structural characterization of human recombinant and bone-derived bone sialoprotein. Functional implications for cell attachment and hydroxyapatite binding.* J Biol Chem, 2001. **276**(39): p. 36839-48.
175. Tye, C.E., Rattray, K.R., Warner, K.J., Gordon, J.A., Sodek, J., Hunter, G.K., and Goldberg, H.A. *Delineation of the hydroxyapatite-nucleating domains of bone sialoprotein.* J Biol Chem, 2003. **278**(10): p. 7949-55.
176. Zaia, J., Boynton, R., Heinegard, D., and Barry, F. *Posttranslational modifications to human bone sialoprotein determined by mass spectrometry.* Biochemistry, 2001. **40**(43): p. 12983-91.
177. Stubbs, J.T., 3rd, Mintz, K.P., Eanes, E.D., Torchia, D.A., and Fisher, L.W. *Characterization of native and recombinant bone sialoprotein: delineation of the mineral-binding and cell adhesion domains and structural analysis of the RGD domain.* J Bone Miner Res, 1997. **12**(8): p. 1210-22.
178. Midura, R.J., Mcquillan, D.J., Benham, K.J., Fisher, L.W., and Hascall, V.C. *A rat osteogenic cell line (UMR 106-01) synthesizes a highly sulfated form of bone sialoprotein.* J Biol Chem, 1990. **265**(9): p. 5285-91.
179. Hunter, G.K. and Goldberg, H.A. *Modulation of crystal formation by bone phosphoproteins: role of glutamic acid-rich sequences in the nucleation of hydroxyapatite by bone sialoprotein.* Biochem J, 1994. **302 (Pt 1)**: p. 175-9.
180. Goldberg, H.A., Warner, K.J., Stillman, M.J., and Hunter, G.K. *Determination of the hydroxyapatite-nucleating region of bone sialoprotein.* Connect Tissue Res, 1996. **35**(1-4): p. 385-92.
181. Boskey, A.L., Doty, S.B., Kudryashov, V., Mayer-Kuckuk, P., Roy, R., and Binderman, I. *Modulation of extracellular matrix protein phosphorylation alters mineralization in differentiating chick limb-bud mesenchymal cell micromass cultures.* Bone, 2008. **42**(6): p. 1061-71.
182. Heinegard, D., Hultenby, K., Oldberg, A., Reinholt, F., and Wendel, M. *Macromolecules in bone matrix.* Connect Tissue Res, 1989. **21**(1-4): p. 3-11; discussion 12-4.
183. Salih, E., Zhou, H.Y., and Glimcher, M.J. *Phosphorylation of purified bovine bone sialoprotein and osteopontin by protein kinases.* J Biol Chem, 1996. **271**(28): p. 16897-905.
184. Salih, E. and Fluckiger, R. *Complete topographical distribution of both the in vivo and in vitro phosphorylation sites of bone sialoprotein and their biological implications.* J Biol Chem, 2004. **279**(19): p. 19808-15.

185. Saad, F.A., Salih, E., Wunderlich, L., Fluckiger, R., and Glimcher, M.J. *Prokaryotic expression of bone sialoprotein and identification of casein kinase II phosphorylation sites.* Biochem Biophys Res Commun, 2005. **333**(2): p. 443-7.
186. Oldberg, A., Franzen, A., Heinegard, D., Pierschbacher, M., and Ruoslahti, E. *Identification of a bone sialoprotein receptor in osteosarcoma cells.* J Biol Chem, 1988. **263**(36): p. 19433-6.
187. Prince, C.W., Dickie, D., and Krumdieck, C.L. *Osteopontin, a substrate for transglutaminase and factor XIII activity.* Biochem Biophys Res Commun, 1991. **177**(3): p. 1205-10.
188. Miyauchi, A., Alvarez, J., Greenfield, E.M., Teti, A., Grano, M., Colucci, S., Zambonin-Zallone, A., Ross, F.P., Teitelbaum, S.L., Cheresh, D., and Et Al. *Recognition of osteopontin and related peptides by an alpha v beta 3 integrin stimulates immediate cell signals in osteoclasts.* J Biol Chem, 1991. **266**(30): p. 20369-74.
189. Ross, F.P., Chappel, J., Alvarez, J.I., Sander, D., Butler, W.T., Farach-Carson, M.C., Mintz, K.A., Robey, P.G., Teitelbaum, S.L., and Cheresh, D.A. *Interactions between the bone matrix proteins osteopontin and bone sialoprotein and the osteoclast integrin alpha v beta 3 potentiate bone resorption.* J Biol Chem, 1993. **268**(13): p. 9901-7.
190. Flores, M.E., Norgard, M., Heinegard, D., Reinholt, F.P., and Andersson, G. *RGD-directed attachment of isolated rat osteoclasts to osteopontin, bone sialoprotein, and fibronectin.* Exp Cell Res, 1992. **201**(2): p. 526-30.
191. Grano, M., Zigrino, P., Colucci, S., Zambonin, G., Trusolino, L., Serra, M., Baldini, N., Teti, A., Marchisio, P.C., and Zallone, A.Z. *Adhesion properties and integrin expression of cultured human osteoclast-like cells.* Exp Cell Res, 1994. **212**(2): p. 209-18.
192. Gordon, J.A., Sodek, J., Hunter, G.K., and Goldberg, H.A. *Bone sialoprotein stimulates focal adhesion-related signaling pathways: Role in migration and survival of breast and prostate cancer cells.* J Cell Biochem, 2009.
193. Chenu, C., Colucci, S., Grano, M., Zigrino, P., Barattolo, R., Zambonin, G., Baldini, N., Vergnaud, P., Delmas, P.D., and Zallone, A.Z. *Osteocalcin induces chemotaxis, secretion of matrix proteins, and calcium-mediated intracellular signaling in human osteoclast-like cells.* J Cell Biol, 1994. **127**(4): p. 1149-58.
194. Raynal, C., Delmas, P.D., and Chenu, C. *Bone sialoprotein stimulates in vitro bone resorption.* Endocrinology, 1996. **137**(6): p. 2347-54.
195. Zhou, H.Y., Takita, H., Fujisawa, R., Mizuno, M., and Kuboki, Y. *Stimulation by bone sialoprotein of calcification in osteoblast-like MC3T3-E1 cells.* Calcif Tissue Int, 1995. **56**(5): p. 403-7.
196. Gordon, J.A., Tye, C.E., Sampaio, A.V., Underhill, T.M., Hunter, G.K., and Goldberg, H.A. *Bone sialoprotein expression enhances osteoblast differentiation and matrix mineralization in vitro.* Bone, 2007. **41**(3): p. 462-73.
197. Goldberg, H.A., Warner, K.J., Li, M.C., and Hunter, G.K. *Binding of bone sialoprotein, osteopontin and synthetic polypeptides to hydroxyapatite.* Connect Tissue Res, 2001. **42**(1): p. 25-37.
198. Harris, N.L., Rattray, K.R., Tye, C.E., Underhill, T.M., Somerman, M.J., D'errico, J.A., Chambers, A.F., Hunter, G.K., and Goldberg, H.A. *Functional analysis of*

bone sialoprotein: identification of the hydroxyapatite-nucleating and cell-binding domains by recombinant peptide expression and site-directed mutagenesis. Bone, 2000. **27**(6): p. 795-802.

199. Silverman, L. and Boskey, A.L. *Diffusion systems for evaluation of biomineralization.* Calcif Tissue Int, 2004. **75**(6): p. 494-501.
200. Chen, Y., Bal, B.S., and Gorski, J.P. *Calcium and collagen binding properties of osteopontin, bone sialoprotein, and bone acidic glycoprotein-75 from bone.* J Biol Chem, 1992. **267**(34): p. 24871-8.
201. San Antonio, J.D., Lander, A.D., Karnovsky, M.J., and Slayter, H.S. *Mapping the heparin-binding sites on type I collagen monomers and fibrils.* J Cell Biol, 1994. **125**(5): p. 1179-88.
202. Keene, D.R., San Antonio, J.D., Mayne, R., Mcquillan, D.J., Sarris, G., Santoro, S.A., and Iozzo, R.V. *Decorin binds near the C terminus of type I collagen.* J Biol Chem, 2000. **275**(29): p. 21801-4.
203. Xu, Y., Gurusiddappa, S., Rich, R.L., Owens, R.T., Keene, D.R., Mayne, R., Hook, A., and Hook, M. *Multiple binding sites in collagen type I for the integrins alpha1beta1 and alpha2beta1.* J Biol Chem, 2000. **275**(50): p. 38981-9.
204. Rosenberg, K., Olsson, H., Morgelin, M., and Heinegard, D. *Cartilage oligomeric matrix protein shows high affinity zinc-dependent interaction with triple helical collagen.* J Biol Chem, 1998. **273**(32): p. 20397-403.
205. Traub, W., Jodaikin, A., Arad, T., Veis, A., and Sabsay, B. *Dentin phosphophoryn binding to collagen fibrils.* Matrix, 1992. **12**(3): p. 197-201.
206. Dahl, T., Sabsay, B., and Veis, A. *Type I collagen-phosphophoryn interactions: specificity of the monomer-monomer binding.* J Struct Biol, 1998. **123**(2): p. 162-8.
207. Scott, J.E. *Proteoglycan-fibrillar collagen interactions.* Biochem J, 1988. **252**(2): p. 313-23.
208. Staatz, W.D., Walsh, J.J., Pexton, T., and Santoro, S.A. *The alpha 2 beta 1 integrin cell surface collagen receptor binds to the alpha 1 (I)-CB3 peptide of collagen.* J Biol Chem, 1990. **265**(9): p. 4778-81.
209. Beher, D., Hesse, L., Masters, C.L., and Multhaup, G. *Regulation of amyloid protein precursor (APP) binding to collagen and mapping of the binding sites on APP and collagen type I.* J Biol Chem, 1996. **271**(3): p. 1613-20.
210. Kleinman, H.K., Mcgoodwin, E.B., Martin, G.R., Klebe, R.J., Fietzek, P.P., and Woolley, D.E. *Localization of the binding site for cell attachment in the alpha1(I) chain of collagen.* J Biol Chem, 1978. **253**(16): p. 5642-6.
211. Gullberg, D., Gehlsen, K.R., Turner, D.C., Ahlen, K., Zijenah, L.S., Barnes, M.J., and Rubin, K. *Analysis of alpha 1 beta 1, alpha 2 beta 1 and alpha 3 beta 1 integrins in cell--collagen interactions: identification of conformation dependent alpha 1 beta 1 binding sites in collagen type I.* Embo J, 1992. **11**(11): p. 3865-73.
212. Somasundaram, R., Ruehl, M., Tiling, N., Ackermann, R., Schmid, M., Riecken, E.O., and Schuppan, D. *Collagens serve as an extracellular store of bioactive interleukin 2.* J Biol Chem, 2000. **275**(49): p. 38170-5.
213. Fujisawa, R., Zhou, H., and Kuboki, Y. *In vitro and in vivo association of dentin phosphophoryn with alpha1CB6 peptide of type I collagen.* Connect Tissue Res, 1994. **31**(1): p. 1-10.

214. Farndale, R.W., Lisman, T., Bihan, D., Hamaia, S., Smerling, C.S., Pugh, N., Konitsiotis, A., Leitinger, B., De Groot, P.G., Jarvis, G.E., and Raynal, N. *Cell-collagen interactions: the use of peptide toolkits to investigate collagen-receptor interactions.* Biochem Soc Trans, 2008. **36**(Pt 2): p. 241-50.
215. Emsley, J., Knight, C.G., Farndale, R.W., and Barnes, M.J. *Structure of the integrin alpha2beta1-binding collagen peptide.* J Mol Biol, 2004. **335**(4): p. 1019-28.
216. Dzamba, B.J., Wu, H., Jaenisch, R., and Peters, D.M. *Fibronectin binding site in type I collagen regulates fibronectin fibril formation.* J Cell Biol, 1993. **121**(5): p. 1165-72.
217. Knight, C.G., Morton, L.F., Peachey, A.R., Tuckwell, D.S., Farndale, R.W., and Barnes, M.J. *The collagen-binding A-domains of integrins alpha(1)beta(1) and alpha(2)beta(1) recognize the same specific amino acid sequence, GFOGER, in native (triple-helical) collagens.* J Biol Chem, 2000. **275**(1): p. 35-40.
218. Sinz, A., Kalkhof, S., and Ihling, C. *Mapping protein interfaces by a trifunctional cross-linker combined with MALDI-TOF and ESI-FTICR mass spectrometry.* J Am Soc Mass Spectrom, 2005. **16**(12): p. 1921-31.
219. Rappsilber, J., Siniossoglou, S., Hurt, E.C., and Mann, M. *A generic strategy to analyze the spatial organization of multi-protein complexes by cross-linking and mass spectrometry.* Anal Chem, 2000. **72**(2): p. 267-75.
220. Bunnett, J.F. *Nucleophilic Reactivity.* Annu Rev Phys Chem, 1963. **14**(1): p. 271-290.
221. Edwards, J.O. and Pearson, R.G. *The Factors Determining Nucleophilic Reactivities.* J Am Chem Soc, 1962. **84**(1): p. 16-24.
222. Pearson, R.G., Sobel, H.R., and Songstad, J. *Nucleophilic reactivity constants toward methyl iodide and trans-dichlorodi(pyridine)platinum(II).* . J Am Chem Soc, 1968. **90**(2): p. 319-326.
223. Hermanson, G.T. *Bioconjugate Techniques, Second Edition.* 2 ed. 2008, London, UK: Academic Press, Elsevier.
224. Sinz, A. *Chemical cross-linking and mass spectrometry for mapping three-dimensional structures of proteins and protein complexes.* J Mass Spectrom, 2003. **38**(12): p. 1225-37.
225. Yasui, N. and Koide, T. *Collagen-Protein Interactions Mapped by Phototriggered Thiol Introduction.* . J Am Chem Soc, 2003. **125**(51): p. 15728-15729.
226. Tao, T., Lamkin, M., and Scheiner, C.J. *The conformation of the C-terminal region of actin: a site-specific photocrosslinking study using benzophenone-4-maleimide.* Arch Biochem Biophys, 1985. **240**(2): p. 627-34.
227. Guo, L.W., Muradov, H., Hajipour, A.R., Sievert, M.K., Artemyev, N.O., and Ruoho, A.E. *The inhibitory gamma subunit of the rod cGMP phosphodiesterase binds the catalytic subunits in an extended linear structure.* J Biol Chem, 2006. **281**(22): p. 15412-22.
228. Giron-Monzon, L., Manelyte, L., Ahrends, R., Kirsch, D., Spengler, B., and Friedhoff, P. *Mapping protein-protein interactions between MutL and MutH by cross-linking.* J Biol Chem, 2004. **279**(47): p. 49338-45.

229. Wood, K.S. and Dunn, S.D. *Role of the asymmetry of the homodimeric b2 stator stalk in the interaction with the F1 sector of Escherichia coli ATP synthase.* J Biol Chem, 2007. **282**(44): p. 31920-7.
230. Luo, Y., Wu, J.L., Li, B., Langsetmo, K., Gergely, J., and Tao, T. *Photocrosslinking of benzophenone-labeled single cysteine troponin I mutants to other thin filament proteins.* J Mol Biol, 2000. **296**(3): p. 899-910.
231. Cole, R.D., Stein, W.H., and Moore, S. *On the Cysteine Content of Human Hemoglobin.* J. Biol. Chem., 1958. **233**(6): p. 1359-63.
232. Gilchrist, T.L., and Rees, C.W. *Carbenes, Nitrenes, and Arynes* Studies in Modern Chemistry. 1969, London: Nelson Publishers.
233. Soubannier, V., Rusconi, F., Vaillier, J., Arselin, G., Chaignepain, S., Graves, P.-V., Schmitter, J.-M., Zhang, J.L., Mueller, D., and Velours, J. *The Second Stalk of the Yeast ATP Synthase Complex: Identification of Subunits Showing Cross-Links with Known Positions of Subunit 4 (Subunit b).* Biochemistry, 1999. **38**(45): p. 15017-24.
234. Hixson, S.H. and Hixson, S.S. *P-Azidophenacyl bromide, a versatile photolabile bifunctional reagent. Reaction with glyceraldehyde-3-phosphate dehydrogenase.* Biochemistry, 1975. **14**(19): p. 4251-4.
235. Gu, S.Q., Jockel, J., Beinker, P., Warnecke, J., Semenkov, Y.P., Rodnina, M.V., and Wintermeyer, W. *Conformation of 4.5S RNA in the signal recognition particle and on the 30S ribosomal subunit.* RNA, 2005. **11**(9): p. 1374-84.
236. Kassabov, S.R. and Bartholomew, B. *Site-directed histone-DNA contact mapping for analysis of nucleosome dynamics.* Methods Enzymol, 2004. **375**: p. 193-210.
237. Rausch, J.W., Sathyanarayana, B.K., Bona, M.K., and Le Grice, S.F. *Probing contacts between the ribonuclease H domain of HIV-1 reverse transcriptase and nucleic acid by site-specific photocross-linking.* J Biol Chem, 2000. **275**(21): p. 16015-22.
238. Konarska, M.M. *Site-specific derivatization of RNA with photocrosslinkable groups.* Methods, 1999. **18**(1): p. 22-8.
239. Mclachlin, D.T., Coveny, A.M., Clark, S.M., and Dunn, S.D. *Site-directed cross-linking of b to the alpha, beta, and a subunits of the Escherichia coli ATP synthase.* J Biol Chem, 2000. **275**(23): p. 17571-7.
240. Wang, Y. and Margoliash, E. *Enzymic activities of covalent 1:1 complexes of cytochrome c and cytochrome c peroxidase.* Biochemistry, 1995. **34**(6): p. 1948-58.
241. Zecherle, G.N., Oleinikov, A., and Traut, R.R. *The C-terminal domain of Escherichia coli ribosomal protein L7/L12 can occupy a location near the factor-binding domain of the 50S subunit as shown by cross-linking with N-[4-(p-azidosalicylamido)butyl]-3-(2'-pyridyldithio)propionamide.* Biochemistry, 1992. **31**(40): p. 9526-32.
242. Van Voorst, F., Vereyken, I.J., and De Kruijff, B. *The high affinity ATP binding site modulates the SecA-precursor interaction.* FEBS Lett, 2000. **486**(1): p. 57-62.
243. Kinns, H. and Howorka, S. *The surface location of individual residues in a bacterial S-layer protein.* J Mol Biol, 2008. **377**(2): p. 589-604.
244. Traut, R.R., Dey, D., Bochkariov, D.E., Oleinikov, A.V., Jokhadze, G.G., Hamman, B., and Jameson, D. *Location and domain structure of Escherichia coli*

ribosomal protein L7/L12: site specific cysteine crosslinking and attachment of fluorescent probes. Biochem Cell Biol, 1995. **73**(11-12): p. 949-58.

245. Seebacher, J., Mallick, P., Zhang, N., Eddes, J.S., Aebersold, R., and Gelb, M.H. *Protein cross-linking analysis using mass spectrometry, isotope-coded cross-linkers, and integrated computational data processing.* J Proteome Res, 2006. **5**(9): p. 2270-82.
246. Gao, Q., Xue, S., Doneanu, C.E., Shaffer, S.A., Goodlett, D.R., and Nelson, S.D. *Pro-CrossLink. Software tool for protein cross-linking and mass spectrometry.* Anal Chem, 2006. **78**(7): p. 2145-9.
247. Mouradov, D., King, G., Ross, I.L., Forwood, J.K., Hume, D.A., Sinz, A., Martin, J.L., Kobe, B., and Huber, T. *Protein structure determination using a combination of cross-linking, mass spectrometry, and molecular modeling.* Methods Mol Biol, 2008. **426**: p. 459-74.
248. Di Lullo, G.A., Sweeney, S.M., Korkko, J., Ala-Kokko, L., and San Antonio, J.D. *Mapping the ligand-binding sites and disease-associated mutations on the most abundant protein in the human, type I collagen.* J Biol Chem, 2002. **277**(6): p. 4223-31.
249. Sweeney, S.M., Orgel, J.P., Fertala, A., Mcauliffe, J.D., Turner, K.R., Di Lullo, G.A., Chen, S., Antipova, O., Perumal, S., Ala-Kokko, L., Forlino, A., Cabral, W.A., Barnes, A.M., Marini, J.C., and San Antonio, J.D. *Candidate cell and matrix interaction domains on the collagen fibril, the predominant protein of vertebrates.* J Biol Chem, 2008. **283**(30): p. 21187-97.
250. Williams, B.R., Gelman, R.A., Poppke, D.C., and Piez, K.A. *Collagen fibril formation. Optimal in vitro conditions and preliminary kinetic results.* J Biol Chem, 1978. **253**(18): p. 6578-85.
251. Chandrakasan, G., Torchia, D.A., and Piez, K.A. *Preparation of intact monomeric collagen from rat tail tendon and skin and the structure of the nonhelical ends in solution.* J Biol Chem, 1976. **251**(19): p. 6062-7.
252. Goldberg, H.A. and Warner, K.J. *The staining of acidic proteins on polyacrylamide gels: enhanced sensitivity and stability of "Stains-all" staining in combination with silver nitrate.* Anal Biochem, 1997. **251**(2): p. 227-33.
253. Tyagarajan, K., Pretzer, E., and Wiktorowicz, J.E. *Thiol-reactive dyes for fluorescence labeling of proteomic samples.* Electrophoresis, 2003. **24**(14): p. 2348-58.
254. Laemmli, U.K. *Cleavage of structural proteins during the assembly of the head of bacteriophage T4.* Nature, 1970. **227**(5259): p. 680-5.
255. Dunn, S.D. *Effects of the modification of transfer buffer composition and the renaturation of proteins in gels on the recognition of proteins on Western blots by monoclonal antibodies.* Anal Biochem, 1986. **157**(1): p. 144-53.
256. Rauterberg, J. and Kuhn, K. *Acid soluble calf skin collagen. Characterization of the peptides obtained by cyanogen bromide cleavage of its alpha-1-chain.* Eur J Biochem, 1971. **19**(3): p. 398-407.
257. Henkel, W. and Dreisewerd, K. *Cyanogen bromide peptides of the fibrillar collagens I, III, and V and their mass spectrometric characterization: detection of linear peptides, peptide glycosylation, and cross-linking peptides involved in*

formation of homo- and heterotypic fibrils. J Proteome Res, 2007. **6**(11): p. 4269-89.

258. Hofmeister, F. *[About the science of the effect of salts].* Arch Exp Pathol Pharmakol, 1888. **24**: p. 247 - 60.
259. Gurau, M.C., Lim, S.M., Castellana, E.T., Albertorio, F., Kataoka, S., and Cremer, P.S. *On the mechanism of the hofmeister effect.* J Am Chem Soc, 2004. **126**(34): p. 10522-3.
260. Sachs, J.N. and Woolf, T.B. *Understanding the Hofmeister effect in interactions between chaotropic anions and lipid bilayers: molecular dynamics simulations.* J Am Chem Soc, 2003. **125**(29): p. 8742-3.
261. Schnell, B., Schurhammer, R., and Wipff, G. *Distribution of Hydrophobic Ions and Their Counterions at an Aqueous Liquid;Liquid Interface: A Molecular Dynamics Investigation.* J Phys Chem B, 2004. **108**(7): p. 2285-94.
262. Collins, K.D. and Washabaugh, M.W. *The Hofmeister effect and the behaviour of water at interfaces.* Q Rev Biophys, 1985. **18**(4): p. 323-422.
263. Zhang, Y. and Cremer, P.S. *Interactions between macromolecules and ions: The Hofmeister series.* Curr Opin Chem Biol, 2006. **10**(6): p. 658-63.
264. Zhang, Y., Furyk, S., Bergbreiter, D.E., and Cremer, P.S. *Specific ion effects on the water solubility of macromolecules: PNIPAM and the Hofmeister series.* J Am Chem Soc, 2005. **127**(41): p. 14505-10.
265. Rieskautt, M. and Ducruix, A. *Inferences drawn from physicochemical studies of crystallogenesis and precrystalline state*, in *Macromolecular Crystallography, Pt A*. 1997, Academic Press Inc: San Diego. p. 23-59.
266. Creighton, T.E. *Proteins: Structures and Molecular Properties*. 2 ed. 1993, New York: W.H. Freeman.
267. Pechmann, S., Levy, E.D., Tartaglia, G.G., and Vendruscolo, M. *Physicochemical principles that regulate the competition between functional and dysfunctional association of proteins.* Proc Natl Acad Sci U S A, 2009. **106**(25):10159-64
268. Tartaglia, G.G., Pawar, A.P., Campioni, S., Dobson, C.M., Chiti, F., and Vendruscolo, M. *Prediction of aggregation-prone regions in structured proteins.* J Mol Biol, 2008. **380**(2): p. 425-36.
269. Rousseau, F., Serrano, L., and Schymkowitz, J.W. *How evolutionary pressure against protein aggregation shaped chaperone specificity.* J Mol Biol, 2006. **355**(5): p. 1037-47.
270. Vendruscolo, M. and Dobson, C.M. *Chemical biology: More charges against aggregation.* Nature, 2007. **449**(7162): p. 555.
271. Klement, K., Wieligmann, K., Meinhardt, J., Hortschansky, P., Richter, W., and Fandrich, M. *Effect of different salt ions on the propensity of aggregation and on the structure of Alzheimer's abeta(1-40) amyloid fibrils.* J Mol Biol, 2007. **373**(5): p. 1321-33.
272. Westermark, P., Benson, M.D., Buxbaum, J.N., Cohen, A.S., Frangione, B., Ikeda, S., Masters, C.L., Merlini, G., Saraiva, M.J., and Sipe, J.D. *Amyloid: toward terminology clarification. Report from the Nomenclature Committee of the International Society of Amyloidosis.* Amyloid, 2005. **12**(1): p. 1-4.

273. Wittelsberger, A., Thomas, B.E., Mierke, D.F., and Rosenblatt, M. *Methionine acts as a "magnet" in photoaffinity crosslinking experiments.* FEBS Lett, 2006. **580**(7): p. 1872-6.
274. Rao, G., Santhoshkumar, P., and Sharma, K.K. *Anti-chaperone betaA3/A1(102-117) peptide interacting sites in human alphaB-crystallin.* Mol Vis, 2008. **14**: p. 666-74.
275. Rossi, A., Zuccarello, L.V., Zanaboni, G., Monzani, E., Dyne, K.M., Cetta, G., and Tenni, R. *Type I collagen CNBr peptides: species and behavior in solution.* Biochemistry, 1996. **35**(19): p. 6048-57.
276. Dalgleish, R. *The human type I collagen mutation database.* Nucleic Acids Res, 1997. **25**(1): p. 181-7.
277. Deyl, Z. and Miksik, I. *Advanced separation methods for collagen parent alpha-chains, their polymers and fragments.* J Chromatogr B Biomed Sci Appl, 2000. **739**(1): p. 3-31.
278. Brennan, M. *Changes in the cross-linking of collagen from rat tail tendons due to diabetes.* J Biol Chem, 1989. **264**(35): p. 20953-60.
279. Pearson, K.M., Pannell, L.K., and Fales, H.M. *Intramolecular cross-linking experiments on cytochrome c and ribonuclease A using an isotope multiplet method.* Rapid Commun Mass Spectrom, 2002. **16**(3): p. 149-59.
280. Koch, S., Yao, C., Grieb, G., Prevel, P., Noah, E.M., and Steffens, G.C. *Enhancing angiogenesis in collagen matrices by covalent incorporation of VEGF.* J Mater Sci Mater Med, 2006. **17**(8): p. 735-41.