This is to certify that the
dissertation entitled

The X-ray Crystallographic Structures of the Angiogenesis Inhibitor Angiostatin Bound to a Peptide from the Group A Streptococcal Surface Protein PAM and the Metal-Bound Conantokins Con-G and Con-T[K7gamma]

presented by

Sara Elizabeth Cnudde

has been accepted towards fulfillment
of the requirements for the

Ph.D. degree in Biochemistry

Major Professor's Signature

11/9/07

Date

THE X-RAY CRYSTALLOGRAPHIC STRUCTURES OF THE ANGIOGENESIS INHIBITOR ANGIOSTATIN BOUND TO A PEPTIDE FROM THE GROUP A STREPTOCOCCAL SURFACE PROTEIN PAM AND THE METAL-BOUND CONANTOKINS CON-G AND CON-T[K7GAMMA]

By

Sara Elizabeth Cnudde

A DISSERTATION

Submitted to
Michigan State University
in partial fulfillment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Department of Biochemistry

2007

ABSTRACT

THE X-RAY CRYSTALLOGRAPHIC STRUCTURES OF THE ANGIOGENESIS INHIBITOR ANGIOSTATIN BOUND TO A PEPTIDE FROM THE GROUP A STREPTOCOCCAL SURFACE PROTEIN PAM AND THE METAL-BOUND CONANTOKINS CON-G AND CON-T[K7GAMMA]

By

Sara Elizabeth Cnudde

Angiostatin is a fragment of plasminogen encompassing the first three kringle domains. Plasminogen is the zymogen of plasmin that is known to bind C-terminal lysine residues in fibrin through the kringle domain lysine binding site. The first structure of a multi-kringle containing compound bound to a ligand was not done until the structure of the angiostatin/VEK-30 complex was determined and refined to 2.3 Å resolution that of which is described herein. It provides a model of the interaction between plasminogen and streptococcal-derived pathogenic proteins during infection. VEK-30 contains a through-space isostere for C-terminal lysine, wherein Arg and Glu side chains, separated by one helical turn, bind within the bipolar angiostatin kringle 2 (K2)-domain lysine-binding site. VEK-30 also makes several contacts with K2 residues that exist outside of the canonical LBS and are not conserved among the other plasminogen kringles, thus providing a molecular basis for the selectivity of VEK-30 for K2. The structure also shows that plasminogen kringle domains undergo significant structural rearrangement relative to one another, and reveals dimerization between two molecules of angiostatin/VEK-30 related by crystallographic symmetry. This dimerization, which only exists in the crystal

structure, is consistent with the parallel coiled-coil full-length PAM dimer expected from sequence similarities and homology modeling.

The challenge of generating a small, unstructured peptide capable of metal ion-triggered helix formation and self-association has been satisfied in nature with the peptide conantokin-G (con-G). Con-G antagonizes the N-methyl-D-aspartate (NMDA) receptors. Con-G is 17 residues and contains five γ-carboxyglumates (Gla). A variety of metals can promote a conformational change from random coil to a helix, but only Ca^{2+} allows for the formation of a dimeric con-G complex. From these data, we proposed a model for the complex in which antiparallel con-G strands are stabilized solely through Ca^{2+}-bridging of Gla headgroups within the helix-helix interface. This model represents a heretofore unknown motif which we define as the "metallo-zipper." A second member from the conantokin family, conantokin-T (con-T), shares some sequence identity to con-G. However, several primary and secondary structural differences exist between con-G and con-T, specificially at position 7, which is occupied by Lys in con-T and Gla in con-G. While con-T does not undergo Ca^{2+}-induced self-assembly, replacing the Lys with a Gla (con-T[K7γ]) allows it to form an antiparallel helix dimer in the presence of Ca^{2+}. However the dimer interface is substantially different from con-G.

We were able to further understand metal-dependant dimerization and helix stabilization by determining the structure of the peptide bound to different metals, such as Cd^{2+} and Mg^{2+}. X-ray structures have been determined of Ca^{2+}-complexed con-G and con-T[K7γ] at high resolution (1.2 Å for con-G and 1.6 Å for con-T[K7γ]) as well as Mg^{2+}/Cd^{2+}- complexed con-T[K7γ] at 1.2 Å.

To my husband, Jeremy

ACKNOWLEDGEMENTS

I would like to thank my research advisor Dr. James Geiger for all of his guidance, encouragement and support throughout the course of my graduate study. I sincerely thank him for all of the invaluable experiences that I have had in his lab. I would also like to thank our collaborator Dr. Frank Castellino and Dr. Mary Prorok for providing the proteins and peptides that were used in both the angiostatin and conantokin studies. I would also like to thank them for helping to write and review the published papers. I would also like to thank my biochemistry advisor Dr. Bill Henry for all of his help and support during my oral and defense. I am in gratitude to my committee members as well Dr. Michael Garavito, Dr. John Wang, and Dr. Hongao Yan.

I would like to acknowledge and thank all former and current members of the Geiger lab, Dr. Stacy Hovde, Dr. Marta Abad, Dr. Xiangshu Jin, Dr. Soheila Vaezeslami, Dr. Andrej Hanzlowsky, Paul Booth, Mike Wolf, Erika Mathes, Aimee Brooks, Lei Feng, Xaofei, Susie, Tyra, Blanka, and Dorothy Tappenden. It has been a pleasure to know you and work with you. A time in my life I will never forget. No lab will ever be like the Geiger lab. I would also like to give a special thanks to Dr. Stacy Hovde. Your help and friendship has meant a great deal to me. Stacy thank you for teaching me to knit. I would also like to thank Dr. Jorge Rios and Dr. Raghuvir Arni for their help learning crystallography and refinement techniques. Thank you everyone and I wish you good luck in your scientific careers.

I would also like to thank my parents who both pushed me to get an education and provided financial support as well as unconditional love.

Last but certainly not least I would like to thank my husband. If not for your unconditional love and continuing support I might not have finished. I am forever grateful. I love you.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER I: INTRODUCTION

CHAPTER II: X-RAY STRUCTURE DETERMINATION

CHAPTER III: THE THREE DIMENSIONAL STRUCTURE OF THE ANGIOSTATIN/VEK-30 COMPLEX

LIST OF FIGURES

Images in this dissertation are presented in color.

CHAPTER I: INTRODUCTION

CHAPTER II: X-RAY STRUCTURE DETERMINATION

CHAPTER III: THE THREE DIMENSIONAL STRUCTURE OF THE ANGIOSTATIN/VEK-30 COMPLEX

LIST OF ABBREVIATIONS

A – alanine
Ala - alanine
apoA – apolipoprotein A
Arg - arginine
Asn – asparagine
Asp – aspartic acid

bicine – N,N Bis(2-hydroxyethyl) glycine

C – cysteine
Cys - cysteine
CC – correlation coefficient
C terminal – carboxy terminus
CHO cells – Chinese hamster ovary cells
con-G - conantokin G
con-T – conantokin T

D – aspartic acid

E – glutamic acid
EACA – ε-aminocaproic acid

F – phenylalanine
Fcalc – calculated structure factors
F_H – structure factor contribution of the derivative
F_{hkl} – structure factor for a reflection labeled *hkl*
Fobs – observed structure factors
F_{PH} – structure factor for protein plus derivative
FAK – focal adhesion kinase

G –glycine
Gla - γ-carboxyglutamate
Gln - glutamine
Glu – glutamic acid
glu-Pg – full length plasminogen
Gly - glycine

H – histidine
hepes – N-[2-hydroxyethyl] piperazine-N'-[ethane sulfonic acid]
His - histidine

I – isoleucine

I – intensity
Iso - isoleucine

K – lysine
Kg – kringle

L – leucine
LBS – lysine binding site
Leu - leucine
LLG – log likelihood gain
Lys - lysine

M – methionine
Met - methionine
MPD – 2-Methyl-2,4-pentanediol
mL – milliliter
min - minute

N – asparagines
NOE – nuclear overhauser effect
N terminal – amino terminal
NTD – N terminal domain
NMDA receptor - N-methyl-D-aspartate receptor

P – proline
PAM – plasminogen binding group A *Streptococcal* M-like surface protein
PEG – polyethylene glycol
Phe - phenylalanine
PP – phasing power
Pm – plasmin
Pg – plasminogen
PGK – phosphoglycerate kinase
Pro - proline
PSA – prostate specific antigen

Q – glutamine
R – arginine
rmsd – root mean square deviation

S – serine
SAD – single wavelength anomalous dispersion
SBC – structural biology center
SDS-PAGE – sodium dodecyl sulfate polyacrylamide gel electrophoresis
Ser - serine
SeMet – selenium methionine
SIR – single isomorphous replacement

SK - streptokinase

T – threonine
Thr - threonine
TM - transmembrane
tPA – tissue type plasminogen activator
Trp - tryptophan
Tyr - tyrosine

uPA – urinary type plasminogen activator

V – valine
Val - valine

W – tryptophan
WT – wild type

Y - tyrosine

Images in this thesis/dissertation are presented in color.

Chapter I

Introduction

1.1 Introduction to the Angiostatin/VEK-30 Complex Structure

1.1.1 Plasminogen

A critical reaction in the generation of the fibrinolytic response is the production of the serine protease plasmin (Pm) from the activation of the zymogen, plasminogen (Pg). Pm catalyzes the proteolysis of the fibrin network, resulting in the dissolution of blood clots. Conversion of Pg to Pm results from cleavage of the Arg561-Val562 peptide bond by tissue-type plasminogen activator (tPA) or urinary-type plasminogen activator (uPA). Cleavage at this site in Pg results in the formation of two-chain Pm, which is composed of a heavy chain and light chain linked by two disulfide bonds. The heavy chain consists of the N-terminal domain (NTD) followed by five consecutive homologous triple disulfide-bonded kringle (Kg or K) domains. The light chain possesses the C-terminal serine protease catalytic unit (Figure 1-1)[7].

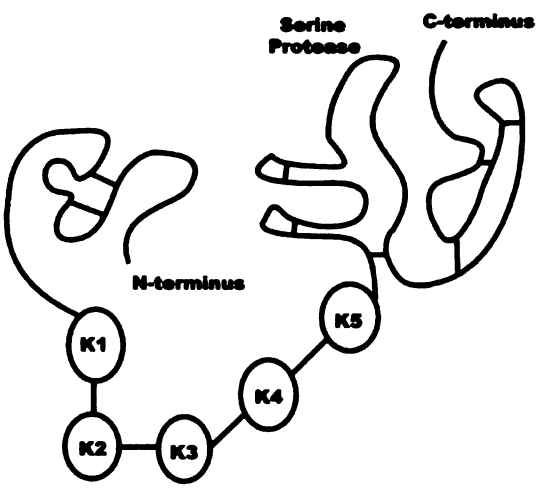

Figure 1-1 Schematic representation of the structure of human plasminogen.

Kringle domains are found in a variety of proteins including Pg[8], prothrombin[9], the unactivated form of thrombin, tPA[10] and uPA[11] and apolipoprotein A (apoA)[12]. Kringle modules have been shown to be protein recognition modules in virtually all cases where a function has been identified. Kringle domains are compact structures that are held together by three critical disulfide bonds. An example of the overall kringle domain structure is shown in Figure 1-2. The five kringle domains of plasminogen share significant sequence and structural homology. A superposition of the Pg kringle domains yields an rmsd value of about 0.40 Å.

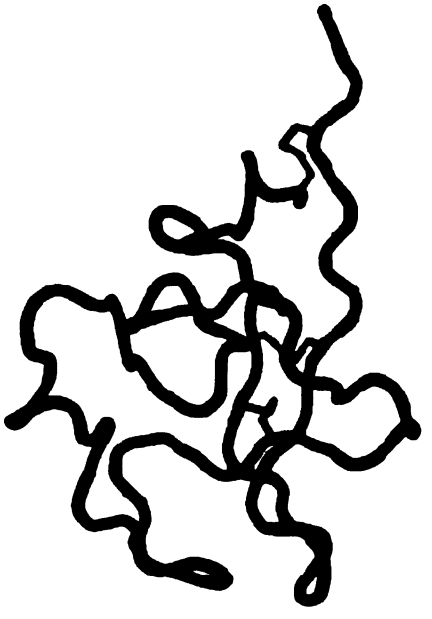

Figure 1-2 The structure of plasminogen K5 (PDB ID 5HPG)[3]. The disulfide bonds are colored red.

Many of these kringles display an affinity for C-terminal lysine residues on proteins, and for small molecules such as epsilon-amino caproic acid (EACA) that mimic C-terminal lysines[13-15]. Their binding modes for lysine-like ligands have been extensively studied both structurally and by site-directed mutagenesis. The specific region responsible for ligand-binding is the lysine-binding site (LBS). Structures of several kringle domains bound to EACA are known and define the LBS as bipolar with a cationic and anionic center that stabilizes the carboxyl and amino group of a C-terminal lysine residue[2,3,16,17]. The LBS is defined by residues R115, R153, D137 and D139 in K1; R234, D219 and E221 in K2; R290, R324, D309 and H317 in K3; D411, D413 and R426 in K4; D516, D518 and R530 in K5[2]. Between the two charged regions of the binding site is a hydrophobic region consisting of two aromatic residues that act as forceps for the intervening hydrophobic methylene chain of the C-terminal lysine. Figure 1-3 shows K1 bound to EACA. Comparison of all kringle structures known to date shows that all the structures are quite similar with root mean square deviations of no more than 1-2 Å. Pg-K1, -K4, and -K5 all show reasonably high affinity for EACA while Pg-K2 has significantly lower affinity for EACA. Pg-K3 displays no affinity for any of the C-terminal lysine mimics[18]. However, specific and high affinity interactions between Pg kringle domains and proteins lacking C-terminal lysines have also been identified. These include tetranectin,[19-21] which binds to Pg-K4, the streptococcal surface protein PAM (Pg-binding group A streptococcal M-like protein),[6,22-24] which binds to Pg-K2, DANCE (developmental arteries and neural crest epidermal growth factor), which binds to Kringle IV type 2 domains of apo(a)[25], factor Va, which binds to K1 and K2 of

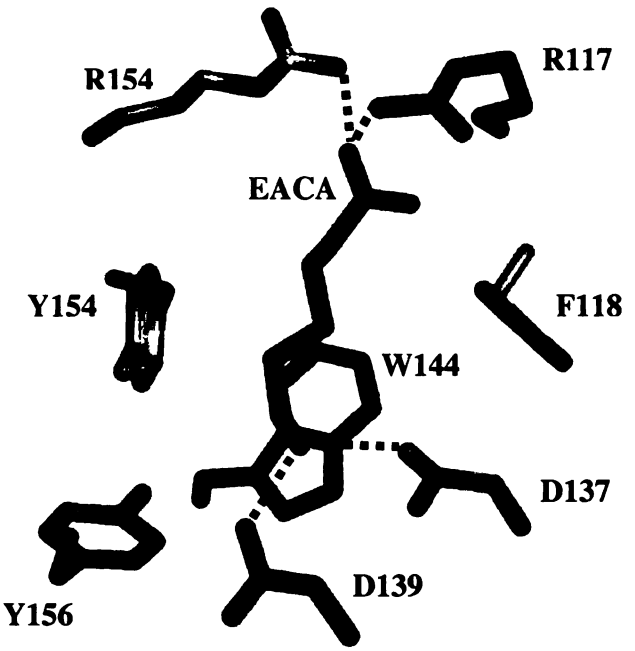

Figure 1-3 LBS of K1-EACA (PDB ID 1CEA)2. K1 is green and the residues are colored by atom (nitrogen, blue; oxygen, red; carbon, green) and EACA is colored magenta. Side chains are labeled using plasminogen numbering.

prothrombin[26,27], and the NG2 proteoglycan, which binds to angiostatin[28,29]. A ligand for Pg-K3 has yet to be identified. The five kringle domains that make up most of the N-terminus of Pg are the most studied examples of kringle domains.

The interaction between the C-terminal lysine of fibrin and the lysine binding kringles of Pm or Pg serve to localize Pg to fibrin, thereby promoting proteolytic dissolution of the fibrin knot by Pm[30]. The five Pg kringle domains have an additional regulatory function involving a dramatic conformational change[31,32]. The full-length Pg (glu-Pg) exists as a tightly compact structure in the presence of Cl^- and is relatively inactive toward activation by Pg activators[6,31,33-35]. Upon EACA binding this structure

becomes significantly less compact and far more susceptible to activation[36]. This compact conformation depends, at least in-part, on interactions between Pg-K5 and the N-terminal 77-residue domain (NTD), since recombinant Pg bearing LBS mutations in K5 or N-terminal lysine substitutions exist only in the extended conformation and show no activation in response to EACA[37]. The LBS's of K1 and K4 also play an important role in the maintenance of the closed conformation, though the target for these binding sites is less clear. Therefore the five kringle domains of Pg must be capable of significant structural rearrangement relative to one another.

1.1.2 PAM

The group A streptococcal surface protein PAM, a 43 kDa member of the M-protein family, binds Pg-K2 with high affinity[38]. M and M-like proteins account for several interactions between group A streptococci and plasma proteins and are known to act as virulence factors by inhibiting phagocytosis[38,39]. In fact, PAM is required for infection by several streptococcal strains[40,41]. It acts by localizing Pm to the bacterial surface, inhibiting fibrin encapsulation during infection. Pm then catalyzes both extra-cellular matrix and fibrin degeneration, thwarting bacterial encapsulation during infection. M and M-like proteins are highly related structurally, consisting of a continuous α–helix encompassing most of the structure, with a membrane-binding domain on the N-terminus. The continuous α–helices then dimerize, forming extended parallel coiled-coil structures that extend tens of angstroms from the cell surface. These M-protein protrusions are easily visualized in EM projections[42]. A region of PAM, spanning amino acids 91 - 116, contains two direct repeat sequences and is responsible for Pg binding by PAM. VEK-30, an α-helical peptide derived from residues 85-113 of

PAM and containing the first and most of the second direct repeat, possesses a high affinity binding site for Pg-K2 (K_d = 460 nM),[6,23,24,38] even though it does not contain a C-terminal lysine residue. Figure 1-4 shows the primary sequence of VEK-30.

VEK-30:

85 95 105
VEKLTADAELQRLKNERHEEAELERLKSEY

Figure 1-4 Primary sequence of VEK-30. The underlined residues are the direct repeat.

VEK-30 specifically binds Pg-K2, having no measurable affinity for any of the other isolated Pg kringle domains and displays almost identical affinity for K2 compared to the full length protein, indicating that it contains most of the PAM protein's interface with K2[23,24]. The crystal structure of Pg-mK2 (Figure 1-5), mutated to contain an upregulated (an order of magnitude higher) lysine binding site (C169G/E221D/L237Y), bound to VEK-30 demonstrates a novel kringle LBS interaction, where the C-terminal lysine is mimicked by an argininyl and glutamyl side-chain residues displaced by almost one helical turn[6]. This arrangement of residues is named a "pseudo-lysine" moiety. This structure shows that kringle domains can make interactions with a variety of bipolar protein ligands and enlarges significantly the possible targets for kringle interaction.

1.1.3 Angiostatin

Angiostatin is an internal fragment of Pg containing the first three or four kringle modules[43,44]. It was one of the first angiogenesis inhibitors to be identified. Angiogenesis is the sprouting of blood vessels from pre-existing capillary beds and is essential for development and wound healing. In 1994, Folkman and coworkers first

identified the plasminogen fragment angiostatin containing K1-4 through its anti-tumor effects in mice[43]. Angiogenesis is critical for the growth of most solid tumors, as a blood supply must be recruited to stimulate significant growth[45,46]. For this reason, intense interest has been focused on angiogenesis inhibitors for use as potential anti-cancer agents.

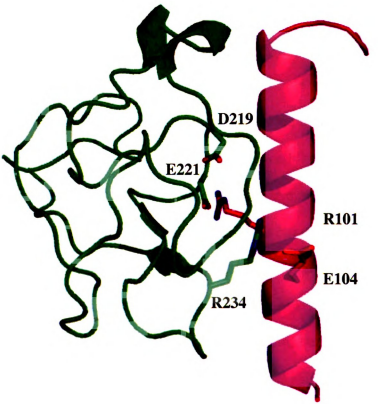

Figure 1-5 X-ray crystallographic structure of mK2/VEK-30 (PDB ID 1I5K)[6]. mK2 is colored green and the VEK-30 peptide is colored magenta. mK2 side chains are labeled using plasminogen numbering and VEK-30 pseudo-lysine labeled using PAM numbering.

Although the function of angiostatin in angiogenesis inhibition is uncertain, agents containing Pg-K1-3, -K1-4, and -K1-5 show potent anti-angiogenic and/or anti-tumor growth activity in animal models. These fragments, as well as individual kringle modules, are also inhibitory toward endothelial cell migration and/or proliferation *in vitro*. Later studies showed, however, that angiostatin corresponding to K1-3 engenders all the determinants responsible for maximal inhibition of cell proliferation and motility[47].

The Production of Angiostatin

Until recently, the extracellular component involved in disulfide reduction of Pm in Chinese hamster ovary cells (CHO cells) or HT1080 cells remained elusive. Hogg and coworkers discovered that the Pm disulfide bonds were reduced by a reductase secreted by CHO or HT1080 cells and that this disulfide bond reduction triggered the proteolysis of Pm generating angiostatin[48]. They believed that Pm reduction is the first step in the formation of angiostatin fragments consisting of kringle domains 1-3, 1-4, and 1-4$^1/_2$[48-52]. The proteolysis of Pm is thought to occur in three stages and a model was proposed by Hogg and coworkers[48-52]. The model proposes that the first step involves the reduction of disulfide bonds in K5 by a Pm reductase. The Pm reductase requires reduced glutathione or cysteine as cofactors. After disulfide reduction, proteolysis occurs at peptide bonds in K5 by a serine proteinase producing K1-4$^1/_2$. Finally, the fragment kringle domain consisting of K1-4$^1/_2$ is cleaved by various matrix metalloproteinases to produce either kringle fragments 1-3 or 1-4[53-58]. The Pm reductase

involved in Pm processing has been shown to be phosphoglycerate kinase (PGK), a glycolytic enzyme[52]. High levels of PGK have been observed in plasma of mice bearing fibrosarcoma tumors and treatment with recombinant PGK inhibits tumor growth between 50-70%, depending on the tumor type. The reduction of Pm by PGK is thought to occur independently of thiols[50].

Other molecules exist as well that mediate the cleavage of plasminogen to angiostatin. The proteolytic activity of elastase catalyzes the production of angiostatin in Lewis lung carcinoma[43,59]. Tumor cells are believed to up regulate elastase production. In human prostate carcinoma, Pm generates angiostatin in the presence of free sulfhydral donors[60]. Prostate specific antigen (PSA) has also been shown to cleave plasminogen to form angiostatin in vitro[61]. Most recently PSA has been shown to be an inhibitor of angiogensis[62]. Thioredoxin and protein disulfide isomerase are also thought to be involved in the generation of angiostatin[48].

The Structure of Angiostatin

The X-ray crystallographic structure of angiostatin was previously determined to a resolution of 1.75 Å[63]. As shown in Figure 1-6, the three kringle domains of angiostatin come together to form a triangular bowl-like structure. An inter-kringle disulfide bond between K2 residue C169 and K3 residue C297 contributes significantly to the relative orientation of K2 and K3, while K1's position is somewhat constrained by the short three-residue inter-kringle peptide linking K1 with K2. However, fixing the orientation of K2 and K3 by disulphide linkage is not required for angiostatin's activity as disruption of the linkage has little effect on its antiproliferative activity[64]. Numerous interactions between the inter-kringle peptides and the kringles appear to stabilize the

arrangement seen in the structure. However, there are no direct interactions between any of the three kringle domains, which raises the possibility that significant motion may occur between the three domains. This motion is limited somewhat by the inter-kringle disulfide bond between K2 and K3.

The crystal structure of angiostatin provides important insights regarding the varying binding site preferences of the three kringle LBS's because each of the three LBS's is bound to a molecule of bicine (N,N Bis (2-hydroxyethyl) glycine) in the structure[63] (Figure 1-6 and 1-7). Interestingly, the LBS of K1, which does have high affinity for C-terminal lysine residues, is located on the opposite face of the molecule relative to K2

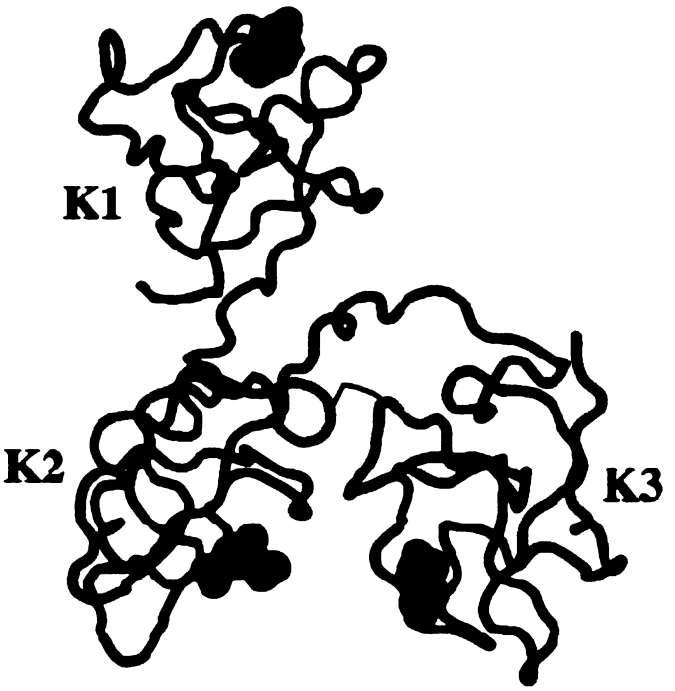

Figure 1-6 The overall structure of angiostatin (PDB ID 1KIO). The interkringle disulfide connecting K2 and K3 (C169-C297) is shown in red. The green spheres are bicine molecules.

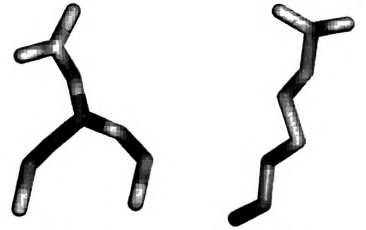

Figure 1-7 The left molecule is bicine and the right molecule is EACA. Both are colored by atom.

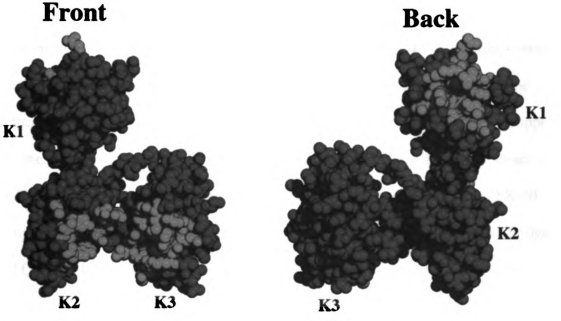

Figure 1-8 Space-filling view of angiostatin. The LBS in each of the three kringles is colored red. All other atoms are blue.

and K3 (Figure 1-8). Since the structures of K1 bound to EACA[2] and K2 bound to VEK-30[6] are known, detailed comparisons can be made between the structures. The K1 LBS (Figure 1-9) binds bicine in an orientation reminiscent of its interaction with EACA, the carboxylate of the bicine interacting with the positive end of the LBS defined by two positively charged residues and the hydrophobic chains of bicine tracking through the V-shaped cleft formed by two aromatic side chains (W144 and Y154). One of the bicine hydroxyl groups then makes a hydrogen bond with one of the aspartates (D137) that defines the negative end of the bipolar binding site. The situation is similar in the K2 LBS (Figure 1-10), with a bicine similarly oriented in the LBS. However, D219, one of the two acidic residues that define the negative end of the K2 LBS, is flipped out of the binding site of K2 and instead makes a tight salt bridge with R220, a residue that is not conserved in other kringle domains. The altered conformation of D219 may explain the lower affinity for EACA that K2 displays when compared to other lysine-binding kringle domains. In contrast, D219 is found to be flipped into the LBS in the structure of K2 bound to the VEK-30 helix. D219 is flipped into the active site in this structure, recapitulating the LBS while R220 does interact with VEK-30. In fact, if the K2 structure of angiostatin is overlayed onto the K2/VEK-30 complex structure (Figure 1-11), a collision occurs between the D219/R220 salt bridge and Q95 and L94 of VEK-30. We surmise that binding of angiostatin to VEK-30 induces this conformational "switch" recapitulating the LBS of K2 by sterically interfering with this salt bridge and interacting with D219 and R220.

Although the K3 LBS (Figure 1-12) is also bound by bicine, the binding is completely different. This is because the K3 LBS is structurally distinct from all other

kringle domain LBS's known. One of the two acidic residues that define the negative end of the LBS is mutated to lysine in K3. This lysine (K311) not only neutralizes the negative end of the LBS, it also tracks across the hydrophobic cleft through the middle of the LBS occluding half of the LBS. Overlaying the K3 LBS over the K1/EACA structure shows that a significant collision would occur between K311 and EACA in the K3 LBS (Figure 1-13), explaining its complete lack of affinity for EACA or C-terminal lysines. However, bicine is still able to bind in the LBS, but it does so by rotating almost 90 degrees relative to its orientation in K1 and K2, inserting its carboxylate group into the remaining binding pocket, which is now essentially positively charged. This interaction may mimic the interaction of K2 with carboxylate side chains from other proteins. Taken together, this indicates that the cleft between K2 and K3 may act by binding to opposite sides of an angiostatin ligand, with the K2 side interacting with some bipolar surface possibly reminiscent of VEK-30's arrangement of glutamate and arginine residues one helical turn away from one another, while the K3 side of the cleft would interact with a protruding carboxylate sidechain. Together, the two kringles K2 and K3 may act like "molecular forceps" specifically interacting with a ligand on two sides. The role of K1 in this binding model is not clear as its LBS is on the opposite face of angiostatin relative to the K2 and K3 LBS's (Figure 1-8).

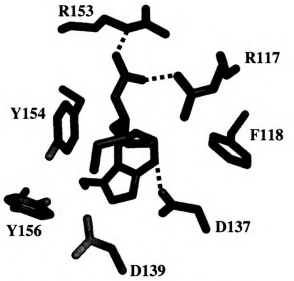

Figure 1-9 Interaction of the angiostatin K1 LBS with bicine. K1 residues are colored green and the bicine molecule is colored magenta. Residues are labeled with plasminogen numbering and atom coloring.

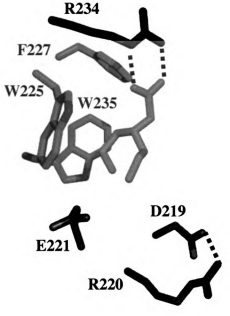

Figure 1-10 Interaction of the angiostatin K2 with bicine. K2 is colored green and bicine is colored magenta. Residues are labeled with plasminogen numbering and atom coloring.

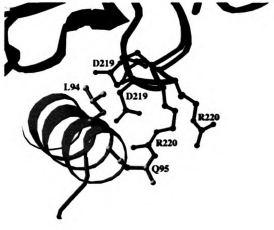

Figure 1-11 Overlay of the angiostatin K2 onto the mK2/VEK-30 structure. Residues are labeled with plasminogen numbering and atom coloring.

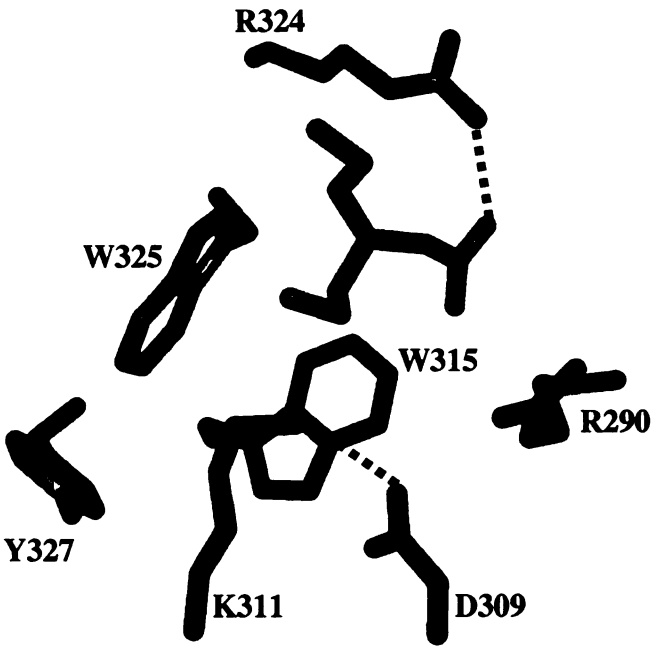

Figure 1-12 Interaction between angiostatin K3 and bicine. K3 is colored green and bicine is colored magenta. Residues are labeled with plasminogen numbering.

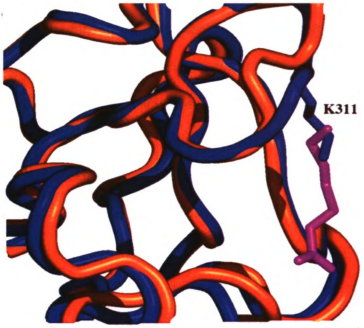

Figure 1-13 Overlay of K1/EACA onto angiostatin K3. K3 is colored blue and K1 is colored orange. EACA is colored lavender. Residues are labeled with plasminogen numbering.

How Does Angiostatin Inhibit Angiogenesis

Blood vessels can be formed by two different processes, vasculogenesis and angiogenesis. Vasculogenesis is the formation of new blood vessels when there are no pre-existing ones. Specifically, vasculogenesis involves the formation of new blood vessels through the differentiation of precursor cells into endothelial cells, forming a vascular network[65]. Vasculogenesis occurs during embryonic development. However, angiogenesis is the process by which blood vessels form from pre-existing capillary beds[46]. Angiogenesis is triggered by a stimulus that forces endothelial dormant cells into the cell cycle. This process starts with basement membrane degradation in endothelial cells and the formation of the lumen. Simultaneously, the endothelial cells

change morphology, proliferate, migrate, form microtubes, and sprout new capillaries. This process is highly regulated involving muliple controls that can be either turned on or off. Both activators and inhibitors of angiogenesis are often found together in tissues that are both active in angiogenesis and those that are quiescent leading to the notion of an "angiogenic switch," where the switch is "off" when angiogenic inhibitors predominate and "on" when the balance is tilted toward activators[66].

Angiogenesis is essential for development and wound healing. While it is intimately involved in the development of the vascular system, angiogenesis is for the most part not active in adults. There are, however, several pathologies that appear to require aberrant angiogenesis. These include rheumatoid arthritis, diabetic retinopathy, cardiovascular diseases, and tumorigenesis[45]. In rheumatoid arthritis, capillaries invade the joints and destroy cartilage. In diabetic retinopathy, new blood vessels invade the retina and cause blindness. Ocular vascularization is the most common cause of blindness. Tumor growth and metastasis depend on angiogenesis as well. Avascular tumors can not grow beyond 2-3 mm^3 without recruiting its own blood supply. But once the tumor is vascularized, the tumor grows rapidly[45,46]. Not only do the blood vessels within the tumor promote tumor growth, it also serves as a portal for tumor cells to enter the blood stream and metastasize.

Once it was determined that tumor growth and metastasis were angiogenesis dependent processes, it was proposed that blocking angiogenesis was a strategy for cancer treatment. Antiangiogenic therapy is a form of cancer treatment that uses angiogenesis inhibitors that specifically stop new blood vessel growth, which in turn starves the tumor. Unlike conventional chemotherapy, antiangiogenic therapy attacks

only new blood vessel cells selectively without affecting normal vasculature and does not kill cancer cells directly. Chemotherapy kills all cells which leads to severe side effects. In order to keep cancers from regrowing, patients may need to take angiogenesis inhibitors for the rest of their lives. Currently, more than 300 angiogenesis inhibitors, including angiostatin, have been discovered so far. Angiostatin has been shown to inhibit tumor growth and metastasis dissemination in animal models[67]. Angiostatin causes no side effects, toxicity, or weight loss and is currently in phase II clinical trials at the Indiana University School of Medicine.

In spite of the intense interest in angiogenesis inhibitors, little is known about the precise mechanism of action of angiostatin or indeed of most of the other angiogenesis inhibitors, but several potential targets of action have been identified. Pizzo and coworkers have discovered that angiostatin binds directly to the mitochondrial F(1)-F(0) ATP synthase and further that the ATPase can be found on the surface of endothelial cells[68]. They have also shown that angiostatin is a potent inhibitor of both the cell-surface and purified mitochondrial forms of the enzyme[69]. The authors speculate that endothelial cell-surface F(1)-F(0) ATP synthase plays an important role in the maintenance of intracellular pH in the acidic tumor cell environment[68,69]. Angiostatin's role is to inhibit the proton pump, allowing the intracellular pH to drop, triggering apoptotic events in endothelial cells[69,70].

Alternatively Takada and coworkers have found that angiostatin specifically binds to the angiogenesis activator $\alpha_v\beta_3$ integrin on the surface of CHO cells and bovine arterial endothelial cells (BAE cells) in an EACA dependent fashion[71]. Further, they have shown that Pm, in contrast to Pg, is also a ligand for $\alpha_v\beta_3$ integrin and that it

induced endothelial cell migration[72]. Angiostatin, an RGD-containing peptide (a known ligand and inhibitor of $\alpha_v\beta_3$) and a serine protease inhibitor all effectively blocked Pm-induced cell migration in these assays[72]. The mechanistic conclusion to be drawn from these studies is that localization of Pm's protease domain to the integrin on the cellular surface is necessary for its activity. Angiostatin, and potentially many of the other Pg kringle domains, would then block this association. These results are quite intriguing as $\alpha_v\beta_3$ integrin is a known angiogenesis activator, though it is not absolutely required for angiogenesis because $\alpha_v\beta_3$-knockout cells still show all the hallmarks of angiogenesis[73].

A third potential target for angiostatin has also been identified. Angiomotin was originally identified in a yeast two hybrid screen using a placental cDNA library[74]. Paradoxically, angiomotin does not appear to be a cell surface receptor nor does it contain a signal sequence for secretion, but nevertheless appears to bind angiostatin on the surface of endothelial cells. Angiomotin is thought to consist of conserved coiled-coil and PDZ binding domains[75]. Cells that contained angiomotin were able to both bind and internalize angiostatin leading to an induction of focal adhesion kinase (FAK) activity[74]. FAK regulates cell motility and adhesion-dependent cell survival and FAK plaques are involved in proton transport[76,77]. Independently, angiostatin was previously shown to activate FAK activity possibly resulting in the inhibition of migration and apoptosis[78]. However, little else is known regarding the function of angiomotin.

Interestingly, an exposed helix with a positive and negative residue separated by approximately one helical turn is found on the β_3 subunit of $\alpha_v\beta_3$ integrin, where lysine and glutamate are found on the exposed side of the helix shown in Figure 1-14[79]. Simply docking angiostatin onto this helix, assuming a similar conformation and

interaction to that seen in the K2/PAM-30 structure indicates that angiostatin could bind in this position and would significantly occlude the integrin's binding site (Figure 1-15). However, without further data this postulate remains highly speculative. It is also of interest that $\alpha_v\beta_3$ integrin and angiostatin complexation is potentially inhibited by concentrations of EACA sufficient to fully saturate both K1 and K2 LBS's[71]. A similar K2 binding site can also be found on the exposed C-terminal helices of both α and β subunits of F1-ATPase encompassing residues K496 and E499 or K472 and E475 in the α and β subunits respectively (Figure 1-16)[80]. Unfortunately, the importance of the LBS's for binding of angiostatin to either angiomotin or F(1)-F(0) ATP synthase has yet to be determined by EACA competition experiments.

1.1.4 Physiological Importance of the Structure

To further understand kringle domain function and kringle domain specificity, angiostatin bound to the VEK-30 peptide was crystallized and its three-dimensional structure determined to a resolution of 2.3 Å. The angiostatin/VEK-30 structure is the first example of an interaction between a multiple kringle-containing protein and a biologically relevant ligand to be visualized at atomic resolution. Furthermore, as the only existing structure of an angiostatin/ligand complex, it provides a model at atomic resolution to test the genesis of anti-angiogenic activity following angiostatin/ligand binding. Structures of other multi-kringle domain complexes will be required to understand how these domains cooperatively interact and undergo conformational rearrangement in order to provide a better understanding of plasminogen structure and function. Plasminogen is an important serine protease that binds to the surface of bacteria. An interesting example resides with flesh-eating bacteria. In order for the

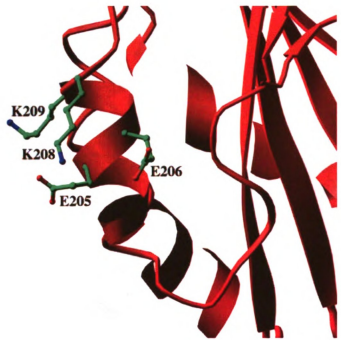

Figure 1-14 Close view of the section of $\alpha_v\beta_3$ integrin that harbors the residues involved in possible angiostatin binding. The β_3 subunit is shown in red and the residues are shown in green with atom coloring.

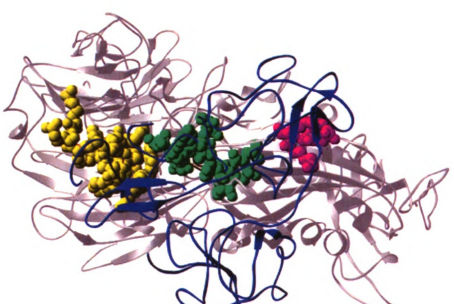

Figure 1-15 Top view of $\alpha_v\beta_3$ integrin with angiostatin docked onto a helix in the β_3 subunit assuming the conformation seen in the mK2/VEK-30 structure. $\alpha_v\beta_3$ integrin is silver, angostatin blue, residues in angiostatin binding are colored magenta, residues in the β_3 subunit responsible for binding are green while the residues in the α_v subunit are shown in yellow.

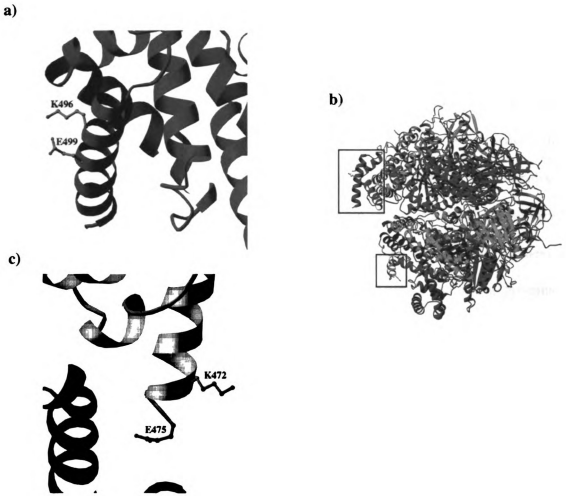

Figure 1-16 (a) Close view of the section of the alpha subunit of F1-ATPase that harbors residues for possible angiostatin binding. (b) Structure of bovine mitochondrial F1-ATPase (131). (c) Enlarged view of the beta subunit F1-ATPase that harbors residues for possible angiostatin binding.

bacterial cell to invade the extracellular matrix of human skin, it must bind an abundant serine protease, plasminogen. The structure of the angiostatin/VEK-30 complex will provide further insight into plasminogen function during bacterial infection. A mechanism for plasminogen function during bacterial infection is proposed herein.

1.2 Introduction to the Metal Bound Structures of con-G and con-T[K7γ]

1.2.1 Protein Design

Knowledge of the interactions that stabilize local protein structure is critical for a global understanding of protein folding. To this end, much effort has been invested in resolving the physical and chemical principles that direct the formation of the α-helix, the most widespread secondary structural motif, and its supersecondary and tertiary structural derivates, which include coiled-coils and helix bundles. Short (≈ 20 residues), linear peptides corresponding to structured segments of naturally occurring proteins would clearly be instructive in this regard, owing to the ease with which their primary sequences can be manipulated and their structures interpreted. However, excepting a few notable examples[81,82], when examined in isolation in aqueous milieu, such abbreviated protein segments rarely adopt the conformations observed within the context of the full-length protein. The *de novo* peptide/protein design field has largely emerged to overcome this problem. Within this discipline, considerable success has been achieved in the design of peptides that are sufficiently short to be synthetically practical, yet capable of adopting stable helices that can further organize into higher order complexes. Among the various strategies that have been employed for directing the formation of monomeric and/or interacting helices, the most tractable and effective approach involves the use of metal ions. For instance, helix nucleation and the attendant folding of unstructured or partially structured peptides can be readily facilitated through metal ion bridging of appropriately spaced natural and unnatural metal-binding amino acids[83-85]. In these particular cases, introduction of the metal induces 80-90% helicity in a 17-mer[83,84] and over 80% helicity in an 11-residue peptide[85], depending on the metal

ion. Similarly, bridging interactions have been exploited to stabilize contacts between two or more helical strands. Many of these studies have focused on the metal-assisted assembly of prefolded *de novo* structures *via* pendant metal ligands, namely porphyrin and bipyridyl derivatives[86,87].

A more daunting challenge in *de novo* construction of higher order structures involves the metal-assisted organization of random coils into helix-helix assemblies. This has been accomplished by induction of a triple-stranded coiled-coil following transition metal ion binding to His or Cys residues contained within a four heptad repeat peptide of a random structure[88,89]. Further, design of Zn^{2+}-chelating sites into a model four-helix bundle resulted in a highly stabilized, ordered structure compared to the apo-variant[90]. A particularly dramatic example of metal ion-induced folding was the design of two disulfide-linked 35-mers that underwent a random coil to coiled-coil transition in the presence of various metal ions[91]. The molecular interactions pivoting this self-organization is interstrand metal ion bridging between γ-carboxyglutamate (Gla) residues occupying positions *e* and *g* in the α–helical heptad repeat. In all of these examples, the metal ion is crucial to initiating strand association, but the predominant forces driving and stabilizing complex formation rely on interstrand hydrophobic contacts. In some of these studies, covalent bonds between the helix-forming strands are necessary for complex stabilization[90,91]. Hence, the metal-ion triggered association of peptide chains in these instances cannot be regarded as unqualifiedly autonomous. Additionally, structural characterization of such complexes has proven elusive, raising questions as to the precise nature of stabilizing interhelix interactions.

A substantial body of evidence has recently been analyzed to suggest that the challenge of generating a small, unstructured peptide capable of metal ion-triggered helix formation and self-association has been satisfied by nature[5]. This peptide is known as conantokin-G (con-G) and is described further in this thesis.

1.2.2 Overview of Cone Snails and Conotoxins

The cone snails (genus *Conus*) are venomous marine mollusks that use small structured peptide toxins (conotoxins) for prey capture, defense, and competitor deterrence[92]. Each of the approximately 500 different *Conus* species can express about 100 different conotoxins. In order to efficiently impose the desired physiological effect on the prey, multiple conopeptides act together in a synergistic fashion. Cone snails can be deadly to humans, specifically, the species *Conus geographus*. The recognition that this snail could kill people led to the initial scientific investigation of cone snail venoms.

There are two broad divisions of cone snail venom components: disulfide-rich conopeptides and non-disulfide rich conopeptides. Conopeptides are usually small containing 12-30 amino acids unlike other polypeptidic toxins from other venoms 40-80 amino acids in length. Conopeptides have a high frequency of posttranslationally modified residues including hydroxyproline, D-amino acids[93], sulfated-tyrosine[94], 6-Br-tryptophan[95], O-glycosylated serine or threonine[96], and γ-carboxy-glutamic acid[97].

The molecular targets of the individual components of *Conus* venom are diverse, including G protein-coupled receptors and neurotransmitter transporters. However, the majority of the *Conus* venom components that have been presently characterized target either ligand-gated ion channels or voltage-gated ion channels. These peptides are able

to discriminate between molecular isoforms within a certain ion channel family. This unprecedented selectivity is an important tool for defining ion channel function.

The voltage-gated ion channel family consists of membrane-bound proteins activated by changes in the transmembrane voltage. Their most important physiological role involves the generation, shaping, and transduction of the cells' electrical signals. They are generally divided into Na^+, K^+, and Ca^{++} channels. Upon activation to the open state, voltage-gated ion channels undergo a conformational change which allows cations through the pore of the channel. An additional conformational change to a closed state causes deactivation and thereby a nonconducing state. Three different conopeptide families are known to target voltage-gated sodium channels: μ-conotoxins, δ-conotoxins, and μO-conotoxins. κ-conotoxins target K channels and ω-conotoxins target Ca channels.

The ligand-gated ion channels are membrane-bound proteins that mediate fast synaptic transmission. One major group of ligand-gated ion channels are those activated by acetylcholine, serotonin, GABA, or glycine. These all belong to the same gene superfamily. α-conotoxins and ψ-conotoxins are known to target the nicotinic receptors. The other gene superfamily of ligand-gated ion channels is the glutamate receptors, including the N-methyl-D-aspartate (NMDA) and non-NMDA (kainite/AMPA) receptors. The conantokins are an unusual class of *Conus* peptides containing a high content of γ-carboxyglutamate (Gla) residues and no Cys residues and are antagonists of the NMDA receptors. A third family of ligand-gated ion channels is the ATP receptors.

1.2.3 con-G

The peptide, conantokin-G (con-G), is a component of the venom of the predatory marine snail, *Conus geographus*[98]. It was originally purified using a behavioral assay. The peptide put young mice to sleep elucidating the nickname "sleeper peptide" but made older mice hyperactive[99,100]. In mammals, con-G antagonizes the N-methyl-D-aspartate (NMDA) subtype of ionotropic glutamate receptors, the dysfunction of which is linked to numerous chronic and acute neuropathologies such as acute neuronal cell death after ischemic stroke and Ca^{2+}-related neurodegeneration in Parkinson's and Alzheimer's diseases[101]. Con-G is 17 amino acids in length, contains no disulfides, and contains five Gla (γ) residues at positions 3, 4, 7, 10 and 14. Figure 1-16 shows the primary sequence of con-G. Of these, only Gla4 is an absolute requirement for NMDA receptor activity[102]. Gla is synthesized from glutamate by a vitamin K-dependent γ-glutamyl carboxylase. The discovery of Gla in con-G established for the first time the presence of Gla in invertebrates[103]. Because of the absence of disulfides in this peptide, it is the Gla residues that provide the structural framework for forming a helical conformation through an interaction with a variety of metal cations. Table 1.1 displays the metal ion effects on the secondary structure and M_{app} of con-G and con-G analogs[1,5]. While a variety of di- and trivalent metals such as Ca^{2+}, Mg^{2+}, and Zn^{2+} can promote a conformational change from random coil to a helix content of 50% and higher, only Ca^{2+} allows for the formation of a dimeric con-G complex with a stoichiometry of 2-3 molecules per mole of peptide[5]. The Ca^{2+}-induced dimerization of con-G was also

shown to be dependant on pH and calcium concentration[5]. The measured binding constant, Kd, of con-G and calcium is 2.8 mM[104]. In the absence of metal cations, con-G has virtually no secondary structure. The calculated apo monomeric molecular weight of con-G is 2265 Da. Interestingly, metal cations such as Mg^{2+} bind tighter to con-G but do not induce metal-mediated dimerization. Thiol-disulfide rearrangement experiments with Cys-containing con-G variants are congruous with an antiparallel alignment of peptide chains in the Ca^{2+}-containing dimer, and Ala-replacement studies have confirmed the participation of Gla residues 3, 7, 10 and 14 in the dimerization event. From these data, a model for the complex in which antiparallel con-G strands are stabilized was proposed solely through Ca^{2+}-bridging of Gla headgroups within the helix-helix interface (Figure 1-18). This model represents a heretofore unknown motif among naturally-occurring proteins or peptides.

1.2.4 con-T

After the discovery of con-G, another member of the conantokin family conantokin-T (con-T), from the marine snail *Conus tulipa* was isolated[105]. Con-T shares sequence identity with con-G at eight sequence positions, including four positions occupied by Gla residues. For con-T, Gla residues exist at positions 3, 4, 10, and 14. However, unlike con-G, con-T at position seven contains a Lys residue instead of a Gla. Also, con-T is a longer peptide containing 21 residues whereas con-G only has 17 residues. Figure 1-16 shows the primary sequence of con-T. The two peptides also have pronounced secondary structural characteristics. As stated above, con-G exhibits minimal structure in its metal-free form, yet adopts an end-to-end helix dimer structure in the presence of divalent metal cations[5]. Table 1.2 shows the metal ion effects on the

Table 1.1 Metal ion effects secondary structure and M_{app} of con-G and analogs[5]. The percentage of helicity was derived from CD measurements, the M_{app} was derived from ultracentrifugation experiments, and Kd was derived from ITC experiments.

Peptide	Metal Ion	% Helix	M_{app} (Da)	Kd (μM)	Number of Molecules
Con-G	None	<2	2550		
	Ca^{2+}	50	3950	2800	2-3
	Mg^{2+}	72	2700	46	2
	Zn^{2+}	68	3110	0.2	3
	Mn^{2+}	67	2875	3.9	2
	Ba^{2+}	7	4200		
	Sr^{2+}	11	3480		
Con-G[γ3A]	None	5	2330		
	Ca^{2+}	31	2530		
Con-G[γ4A]	None	<2	2480		
	Ca^{2+}	36	3120		
Con-G[γ7A]	None	5	2480		
	Ca^{2+}	24	2650		
Con-G[γ7K]	None	10	2470		
	Ca^{2+}	32	2450		
Con-G[γ10A]	None	7	2410		
	Ca^{2+}	14	2310		
Con-G[γ14A]	None	10	2110		
	Ca^{2+}	27	2160		

secondary structure of con-T and con-T analogs. Similarly, con-T also adopts a full α-helical conformation in the presence of metal cations[104,106,107]. However, this change is less dramatic since con-T already has high helical content in the absence of metal cations. Another large difference exists between con-G and con-T. In the presence of Ca^{2+}, con-G undergoes calcium induced self-assembly forming an antiparallel dimeric structure whereas con-T does not[4,5]. However, previous studies have shown that replacing the Lys at position 7 with a Gla (con-T[K7γ]) allows the peptide to form an antiparallel helix dimer in the presence of calcium[4]. Interestingly, neither con-G nor con-T[K7γ] forms a dimeric structure in the presence of Mg^{2+} even though both peptides have a higher affinity for Mg^{2+} than Ca^{2+}[1,4,5]. Con-T also has a 10-fold greater affinity for Ca^{2+} than con-G[4]. This is consistent with the notion of a pre-formed cation binding site in con-T versus an induced cation binding site in con-G. A model for the Ca^{2+}-bridged con-T[K7γ] dimer, similar to that proposed for con-G (Figure 1-18), is presented in Figure 1-19.

	b c d e f g a b c d e f g a b c d e f g a
con-G	**G-E-γ-γ-L-Q- γ- N-Q-γ-L-I- R-γ-K-S- N-NH2**
con-T	**G-E-γ-γ-Y-Q-K-M-L-γ-N-L-R-γ-A-E-V-K-K-N-A-NH2**
con-T[K7γ]	**G-E-γ-γ-Y-Q- γ-M-L-γ-N-L-R-γ-A-E-V-K-K-N-A-NH2**

Figure 1-17 The primary sequences of con-G, con-T, and con-T[K7γ]. The α-helical heptad repeat assignments are shown above.

1.2.5 NMDA receptor

The N-methyl-D-aspartate (NMDA) receptor is a subclass of ionotropic glutamate-dependent receptors that is widely distributed in the mammalian CNS[108]. The NMDA receptor forms a cell membrane ion channel that controls the flux of Ca^{2+} and Na^{2+} into nerve cells. The receptor is activated only if glycine and glutamate bind in concert and polyamines assist in the opening or closing of the ion channel, depending on concentration. Protons, Zn^{2+}, and redox agents further modulate the opening of the

Table 1.2 Metal ion effects secondary structure and M_{app} of con-T and analogs[4]. The percentage of helicity was derived from CD measurements, the M_{app} was derived from ultracentrifugation experiments, and Kd was derived from ITC experiments.

Peptide	Metal Ion	% Helix	M_{app} (Da)	Kd (mM)	Number of Molecules
Con-T	None	55	3160		
	Ca^{2+}	72	3240	428	1
	Mg^{2+}	82	3290	10	1
Con-T[K7γ]	None	12	3390		
	Ca^{2+}	67	6660	120	2-3
	Mg^{2+}	74	3330	18	1-2
	Zn^{2+}	72	3750		
	Mn^{2+}	70	3700		
	Ba^{2+}	50	10080		
	Sr^{2+}	52	7060		

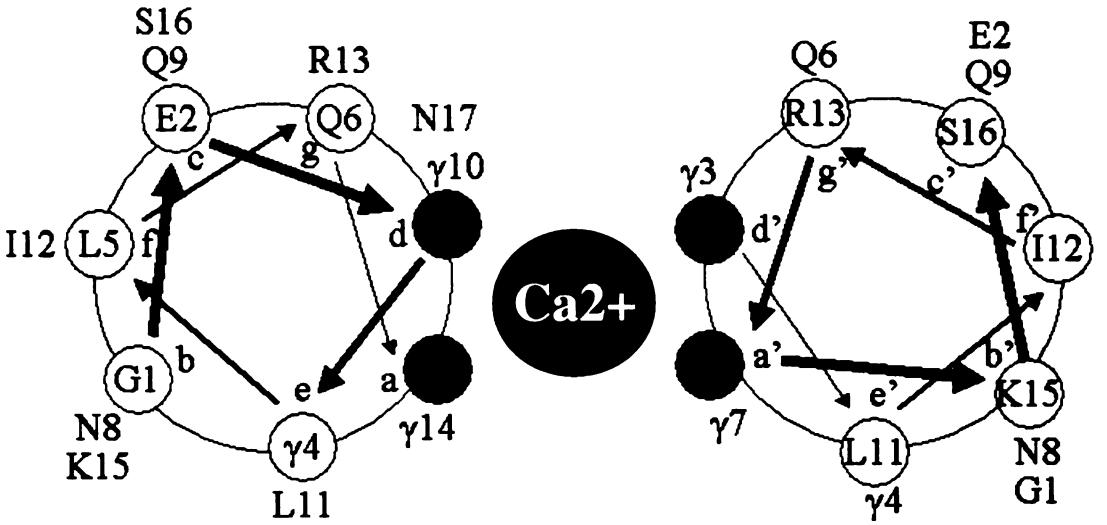

Figure 1-18 Helical wheel representation of the cross-sectional heptad repeat of the antiparallel dimer of con-G in the presence of Ca^{2+5}. For the left chain, the N-terminus is closest to the viewer and for the right chain the C-terminus is closest to the viewer. Gla7 and Gla14 of the left chain occupy position a, while Gla3 and Gla10 occupy position d. Gla chelation pairs include Gla3-Gla14', Gla7-Gla10', Gla10-Gla7', Gla14-Gla3'. The lighter arrows are at the bottom of the wheel, progressively thickening towards the top.

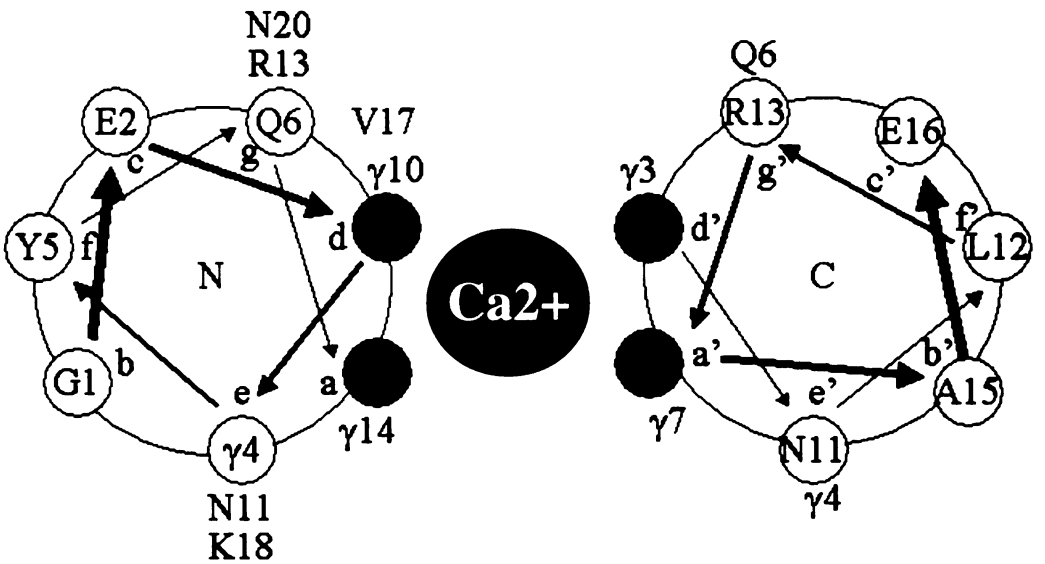

Figure 1-19 Helical wheel representation of the cross-sectional heptad repeat of the antiparallel dimer of con-T[K7γ] in the presence of Ca^{2+4}. For the left chain, the N-terminus is closest to the viewer and for the right chain the C-terminus is closest to the viewer. Gla7 and Gla14 of the left chain occupy position a, while Gla3 and Gla10 occupy position d. Gla chelation pairs include Gla3-Gla14', Gla7-Gla10', Gla10-Gla7', Gla14-Gla3'. The lighter arrows are at the bottom of the wheel, progressively thickening towards the top.

NMDA receptor. The general features of the native form of the NMDA receptor are as follows: a) a voltage dependent channel block by Mg^{2+}; b) glycine and glutamate coactivation; c) allosteric agonism by polyamines such as spermine and spermidine; d) voltage-insensitive and voltage-sensitive allosteric blocks by Zn^{2+}; e) when the ion channel is open high Ca^{2+} permeability; f) proton allosteric antagonism, g) sensitivity to oxidizing and reducing agents; and h) phosphorylation regulation[101]. The opening of the ion channel is believed to be due to voltage-dependent relief of the Mg^{2+} channel block which therefore allows Ca^{2+} to enter the cells[101,109]. When a stroke occurs, a detrimental influx of calcium initiates a cascade of events that will eventually lead to neuronal cell death. Therefore, elucidating the effects of the mechanisms of the NMDA antagonists and developing potent molecules to attenuate Ca^{2+} influx are the focus of intense research.

The action of con-G and presumably con-T as non-competitive antagonists of the NMDA receptor has been attributed to the inhibition of the positive modulatory effects of polyamines. They are the only known naturally occurring peptide ligands that specifically antagonize this receptor. These peptides specifically inhibit polyamine-stimulated dizolcipine (MK-801) binding in rat brain membranes with an IC_{50} of about 500 nM[107,110], attenuate the NMDA-stimulated rise in neuronal intracellular Ca^{2+} (IC_{50} = 2 μM)[105], decrease the NMDA-stimulated cGMP levels in rat cerebellar granule neurons (IC_{50} = 77-177 nM)[111], and inhibit glutamate-induced neurotoxicity[110]. The inhibitory mechanism for which con-G and con-T work involve their inhibition of the positive effector roles of spermine[110], glutamate and glycine[112,113].

The glutamate class of receptors are believed to be tetramers consisting of two types of subunits, NR1 and NR2[92]. So far, only one gene has been found to encode NR1 subunits, but several splice variants have been discovered. However, the four NR2 subunits, NR2A, NR2B, NR2C, and NR2D, are encoded by different genes. Based on homology considerations with other receptors, a membrane binding model of NR1 has been developed[101,114]. This model includes an 18-residue signal peptide, followed by a large N-terminal extracellular portion (543-residues) containing ten potential glycosylation sites and regions that influence Zn^{2+} and spermine function[115,116] and glycine[117,118] and possibly glutamate binding sites[119]. However, there is some speculation that the glutamate binding site resides in the NR2 subunit[120]. Downstream of this extracellular region, a transmembrane (TM) domain of about 19-residues is present followed by a 19-residue linker and a second transmembrane-like domain, TM2. TM2 is thought to exist as a hairpin structure and also contains Asn616, a residue responsible for full expression of the Mg^{2+}, Zn^{2+}, and MK-801 blocks as well as Ca^{2+} permeability[121-123]. Next, a relatively short linker region exists followed by a third TM domain that spans the membrane. A large 163-residue linker exists after TM3 and is thought to contain a region of the glycine binding site[124,125] and a site that partially controls responses to spermine and protons[126]. Finally, a fourth TM domain is present along with an intracellular C-terminal region that contains sites that become phosphorylated[127]. It has been suggested that NR1A is the prevalent mRNA in adult rat brain[128]. NR2 only has about 20% homology with NR1A and is known to be longer than NR1[101]. Assuming similar domain distributions of the NR1 and NR2 subunits, the

major difference resides in the C-terminal domain length. NR2 homomers are inactive with each other and must be co-assembled with NR1 for active channels to result.

Activation of the NMDA receptor ion channel leads to an increase in intracellular Ca^{2+} concentration[101]. This has significant ramifications resulting in many neuropathologies related to the NMDA receptor. One result of the Ca^{2+} influx is the activation of the Ca^{2+}/calmodulin-dependent protein kinase II (CaMKII)[101]. It is thought that this influx of Ca^{2+} following NMDA activation results in a sustained CaMKII activation which is known to lead to nitric acid production. It is thought that this may be in-part responsible for neurodegenerative effects.

1.2.6 Structure-Activity Relationships

The mechanism by which conantokins inhibit ion flow in NMDA receptors appears complex. Small linear peptides like the conantokins may be systematically altered easily. This makes the conantokins prime targets for structure-activity studies. The neurochemical activity of these peptides is usually studied through concentration effects on spermine-enhanced [^{3}H]MK-801 binding to rat neuronal membranes. Single Ala-replacement studies were also performed in order to characterize the functional and structural contributions of the side chains in the conantokins[1]. Table 1.3 shows the effects of side chains on the NMDA receptor antagonist and conformational properties of con-G. Substitutions by Ala at Glu2, Gla4, Leu5, Gln9, and Ile12 resulted in 200-fold increases in the IC_{50} value associated with [^{3}H]MK-801 binding. Gly1, Gla3, Leu11, and Arg13 exhibited inhibitory potencies diminished by 8-20-fold. The remainder of the substituted side chain variants did not show any significant deviations in activity versus the wild-type con-G. A similar experiment was performed on con-T. These results are

displayed in Table 1.4. The most drastic deviations occurred for Ala replacements at Glu2, Gla3, Gla4, and Tyr5. Gly1, Met8, and Leu12 displayed intermediate effects. However, Gln6, Asn11, and Gla14 actually amplified the potency slightly. Position 7 occupied by Gla in con-G and Lys in con-T has no bearing on the activity of conantokins when replaced with Ala. Moreover, con-G[□7K] and con-T[K7□] have IC_{50} values in the [^{3}H]MK-801 binding assay that are comparable to the wild-type. The replacement of Gla10 or Gla14 does not alter con-G nor con-T receptor activity. Replacing either Gla10 or Gla14 with Ala reduces con-G and con-T helicity in the presence of Ca^{2+}. Gla10 and Gla14 have been shown to be important for maintaining metal binding. Since these residues are not essential for activity, metal cation binding to these residues as well as the increased helicity cannot be a factor in NMDA receptor binding. This is also consistent with studies using con-T where the IC_{50} values to [^{3}H]MK-801 binding were unchanged by the addition of 1.5 mM Ca^{2+}[129]. It is also interesting to note that the introduction of a helix-breaking Pro residue at position 7 in con-G completely eliminated activity[108]. This supports the importance of a helical structure.

It is possible that con-G and con-T interact with different NMDA receptor subunit combinations. It has been shown that con-G selectively antagonizes the NR1A/NR2B receptor combination without influencing receptors containing NR1A and NR2A subunits[130]. Also electrophysiological studies using recombinant receptors consisting of NR1a, NR1b, NR2A, and NR2B subunits indicate that the subunit selectivity of con-G may be derived from the presence of Leu5 while in con-T Tyr5 confers broad selectivity[131].

Table 1.3 Effects of side chains on the NMDA receptor antagonist and conformational properties of con-G[1].

Peptide	IC_{50} (μM)	Apo % Helix	Ca^{2+} % Helix
Con-G[G1A]	6.5	0	32
Con-G[E2A]	>100	2	42
Con-G[□3A]	9.6	5	31
Con-G[□4A]	>100	0	36
Con-G[L5A]	>100	0	35
Con-G[Q6A]	0.5	0	35
Con-G[□7A]	0.1	5	24
Con-G[N8A]	0.5	2	44
Con-G[Q9A]	>100	0	36
Con-G[□10A]	2.0	7	14
Con-G[L11A]	5.9	0	37
Con-G[I12A]	>100	0	28
Con-G[R13A]	3.4	0	15
Con-G[□14A]	0.2	11	21
Con-G	0.5	2	50
Con-G [1-15]	1.2	0	27
Con-G [1-14]	5.9	0	17
Con-G [1-13]	2.9	0	4
Con-G [1-12]	>100	0	6
Con-G [1-11]	>100	0	5

Table 1.4 Effects of side chains on the NMDA receptor antagonist and conformational properties of con-T[1].

Peptide	IC_{50} (μM)	Apo % Helix	Ca^{2+} % Helix
Con-T[G1A]	4.2	53	61
Con-T[E2A]	38.7	52	65
Con-T[□3A]	13.5	59	75
Con-T[□4A]	>60	22	47
Con-T[Y5A]	>100	53	86
Con-T[Q6A]	0.17	50	60
Con-T[K7A]	1.6	41	65
Con-T[M8A]	5.8	35	51
Con-T[L9A]	0.60	26	39
Con-T[□10A]	0.71	56	22
Con-T[N11A]	0.17	51	58
Con-T[L12A]	5.0	39	50
Con-T[R13A]	0.43	22	36
Con-T[□14A]	0.20	51	35
Con-T	0.40	55	82
Con-T[1-17]	1.5	46	57
Con-T[1-13]	2.7	46	57
Con-T[1-11]	6.9	15	8
Con-T[1-9]	9.2	10	4
Con-T[1-8]	51.3	2	2
Con-T[1-6]	>100	2	2

1.2.7 Relevance of the project

In the realm of protein design, the forced oligomerization of engineered helices has potential uses in a variety of biotechnological applications, including affinity purification, hydrogel formation, biosensor development, and is of general interest for expanding our knowledge of the forces that govern molecular self-organization[5]. For these applications, a reversible control of the association event through a stimulus-sensitive trigger, such as temperature or pH, is desirable. In this contextual framework, metal ions have been employed to direct the self-assembly of designed helices. However, work describing metal ion induced association of designed peptides has been limited until the discovery of the naturally occurring peptide con-G. In previous metal bound complexes, metal ions initiate strand association but rely mostly on interstrand hydrophobic contacts. However, structural characterization of these complexes has proven to be elusive, raising many questions as to the precise nature of stabilizing interhelix interactions. In order to further understand how metal ions induce self-association and helix stabilization, the crystal structure of con-G and con-T[K7□] in the presence of Ca^{2+} has been determined as well as the structure of con-T[K7□] in the presence of Cd^{2+} and Mg^{2+}. The structures of these complexes are further described herein.

1.3 Literature Cited

1. Prorok, M. & Castellino, F.J. Structure-function relationships of the NMDA receptor antagonist conantokin peptides. *Curr Drug Targets* **2**, 313-22 (2001).

2. Mathews, II, Vanderhoff-Hanaver, P., Castellino, F. J., and Tulinsky, A. Crystal structures of the recombinant kringle 1 domain of human plasminogen in complexes with the ligands epsilon-aminocaproic acid and trans-4-(aminomethyl)cyclohexane-1-carboxylic Acid. *Biochemistry* **35**, 2567-76 (1996).

3. Chang, Y., Mochalkin, I., McCance, S. G., Cheng, B., Tulinsky, A., and Castellino, F. J. Structure and ligand binding determinants of the recombinant kringle 5 domain of human plasminogen. *Biochemistry* **37**, 3258-71 (1998).

4. Dai, Q., Castellino, F.J. & Prorok, M. A single amino acid replacement results in the Ca2+-induced self-assembly of a helical conantokin-based peptide. *Biochemistry* **43**, 13225-32 (2004).

5. Dai, Q., Prorok, M. & Castellino, F.J. A new mechanism for metal ion-assisted interchain helix assembly in a naturally occurring peptide mediated by optimally spaced gamma-carboxyglutamic acid residues. *J Mol Biol* **336**, 731-44 (2004).

6. Rios-Steiner, J.L., Schenone, M., Mochalkin, I., Tulinsky, A., and Castellino, F. J. Structure and binding determinants of the recombinant kringle-2 domain of human plasminogen to an internal peptide from a group A Streptococcal surface protein. *J Mol Biol* **308**, 705-19 (2001).

7. Castellino, F.J. & Ploplis, V.A. Structure and function of the plasminogen/plasmin system. *Thromb Haemost* **93**, 647-54 (2005).

8. Sottrup-Jensen, L., Claeys, H., Zajdel, M., Petersen, T.E., and Magnusson, S. *Prog Chem Fibrinolysis Thrombolysis* **3**, 191-209 (1978).

9. Magnusson, S., Petersen, T.E., Sottrup-Jensen, L., and Claeys, H. *in Proteases and Biological Control*, 123-149 (Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, 1975).

10. Pennica, D. et al. Cloning and expression of human tissue-type plasminogen activator cDNA in E. coli. *Nature* **301**, 214-21 (1983).

11. Steffens, G.J., Gunzler, W.A., Otting, F., Frankus, E. & Flohe, L. The complete amino acid sequence of low molecular mass urokinase from human urine. *Hoppe Seylers Z Physiol Chem* **363**, 1043-58 (1982).

12. McLean, J.W. et al. cDNA sequence of human apolipoprotein(a) is homologous to plasminogen. *Nature* **330**, 132-7 (1987).

13. Urano, T., Deserrano, V. S., Chibber, B. A. K., and Castellino, F. J. The Control of the Urokinase-Catalyzed Activation of Human Glutamic-Acid 1-Plasminogen by Positive and Negative Effectors. *J Biol Chem* **262**, 15959-15964 (1987).

14. Urano, T., Chibber, B.A.K. & Castellino, F.J. The Reciprocal Effects of Epsilon-Aminohexanoic Acid and Chloride-Ion on the Activation of Human (Glu1]Plasminogen by Human Urokinase. *J Biol Chem* **84**, 4031-4034 (1987).

15. Winn, E.S., Hu, S. P., Hochschwender, S. M., and Laursen, R. A. Studies on the lysine-binding sites of human plasminogen. The effect of ligand structure on the binding of lysine analogs to plasminogen. *Eur J Biochem* **104**, 579-86 (1980).

16. McCance, S.G., Menhart, N., and Castellino, F. J. Amino-Acid-Residues of the Kringle-4 and Kringle-5 Domains of Human Plasminogen That Stabilize Their Interactions with Omega- Amino Acid Ligands. *J Biol Chem* **269**, 32405-32410 (1994).

17. Wu, T.P., Padmanabhan, K., Tulinsky, A., and Mulichak, A. M. The refined structure of the epsilon-aminocaproic acid complex of human plasminogen kringle 4. *Biochemistry* **30**, 10589-94 (1991).

18. Geiger, J.H. & Cnudde, S.E. What the structure of angiostatin may tell us about its mechanism of action. *J Thromb Haemost* **2**, 23-34 (2004).

19. Clemmensen, I., Petersen, L.C. & Kluft, C. Purification and characterization of a novel, oligomeric, plasminogen kringle 4 binding protein from human plasma: tetranectin. *Eur J Biochem* **156**, 327-33 (1986).

20. Graversen, J.H., Lorentsen, R. H., Jacobsen, C., Moestrup, S. K., Sigurskjold, B. W., Thogersen, H. C., and Etzerodt, M. The plasminogen binding site of the C-type lectin tetranectin is located in the carbohydrate recognition domain, and binding is sensitive to both calcium and lysine. *J Biol Chem* **273**, 29241-6 (1998).

21. Graversen, J.H., Sigurskjold, B. W., Thogersen, H. C., and Etzerodt, M. Tetranectin-binding site on plasminogen kringle 4 involves the lysine-binding pocket and at least one additional amino acid residue. *Biochemistry* **39**, 7414-9 (2000).

22. Schenone, M.M., Warder, S. E., Martin, J. A., Prorok, M., and Castellino, F. J. An internal histidine residue from the bacterial surface protein, PAM, mediates its binding to the kringle-2 domain of human plasminogen. *J Pept Res* **56**, 438-45 (2000).

23. Wistedt, A.C., Ringdahl, U., Muller-Esterl, W., and Sjobring, U. Identification of a plasminogen-binding motif in PAM, a bacterial surface protein. *Mol Microbiol* **18**, 569-78 (1995).

24. Wistedt, A.C., Kotarsky, H., Marti, D., Ringdahl, U., Castellino, F. J., Schaller, J., and Sjobring, U. Kringle 2 mediates high affinity binding of plasminogen to an internal sequence in streptococcal surface protein PAM. *J Biol Chem* **273**, 24420-4 (1998).

25. Kapetanopoulos, A. et al. Direct interaction of the extracellular matrix protein DANCE with apolipoprotein(a) mediated by the kringle IV-type 2 domain. *Mol Genet Genomics* **267**, 440-6 (2002).

26. Deguchi, H., Takeya, H., Gabazza, E.C., Nishioka, J. & Suzuki, K. Prothrombin kringle 1 domain interacts with factor Va during the assembly of prothrombinase complex. *Biochem J* **321 (Pt 3)**, 729-35 (1997).

27. Kotkow, K.J., Deitcher, S.R., Furie, B. & Furie, B.C. The second kringle domain of prothrombin promotes factor Va-mediated prothrombin activation by prothrombinase. *J Biol Chem* **270**, 4551-7 (1995).

28. Chekenya, M. et al. NG2 proteoglycan promotes angiogenesis-dependent tumor growth in CNS by sequestering angiostatin. *Faseb J* **16**, 586-8 (2002).

29. Goretzki, L., Lombardo, C.R. & Stallcup, W.B. Binding of the NG2 proteoglycan to kringle domains modulates the functional properties of angiostatin and plasmin(ogen). *J Biol Chem* **275**, 28625-33 (2000).

30. Suenson, E. & Thorsen, S. Secondary-Site Binding of Glu-Plasmin, Lys-Plasmin and Miniplasmin to Fibrin. *Biochem J* **197**, 619-628 (1981).

31. Christensen, U. & Molgaard, L. Positive co-operative binding at two weak lysine-binding sites governs the Glu-plasminogen conformational change. *Biochem J* **285 (Pt 2)**, 419-25 (1992).

32. Takada, Y., Urano, T. & Takada, A. Conformational change of plasminogen: effects of N-terminal peptides of Glu-plasminogen. *Thromb Res* **70**, 151-9 (1993).

33. Mangel, W.F., Lin, B., and Ramakrishnan, V. Characterization of an Extremely Large, Ligand-Induced Conformational Change in Plasminogen. *Science* **248**, 69-73 (1990).

34. Marshall, J.M., Brown, A. J., and Ponting, C. P. Conformational Studies of Human Plasminogen and Plasminogen Fragments - Evidence for a Novel 3rd Conformation of Plasminogen. *Biochemistry* **33**, 3599-3606 (1994).

35. McCance, S.G. & Castellino, F.J. Contributions of individual kringle domains toward maintenance of the chloride-induced tight conformation of human glutamic acid-1 plasminogen. *Biochemistry* **34**, 9581-6 (1995).

36. Lin, L.F., Houng, A.Y. & Reed, G.L. Epsilon amino caproic acid inhibits streptokinase-plasminogen activator complex formation and substrate binding through kringle-dependent mechanisms. *Biochemistry* **39**, 4740-4745 (2000).

37. Cockell, C.S., Marshall, J.M., Dawson, K.M., Cederholm-Williams, S.A. & Ponting, C.P. Evidence that the conformation of unliganded human plasminogen is maintained via an intramolecular interaction between the lysine-binding site of kringle 5 and the N-terminal peptide. *Biochem J* **333 (Pt 1)**, 99-105 (1998).

38. Berge, A. & Sjobring, U. PAM, a Novel Plasminogen-binding Protein from Streptococcus pyogenes. *J Biol Chem* **268**, 25417-25424 (1993).

39. Fischetti, V.A. Streptococcal M protein: molecular design and biological behavior. *Clin Microbiol Rev* **2**, 285-314 (1989).

40. Sanderson-Smith, M. et al. Divergence in the plasminogen-binding group A streptococcal M protein family: functional conservation of binding site and potential role for immune selection of variants. *J Biol Chem* **281**, 3217-26 (2006).

41. Boxrud, P.D., and P.E. Bock. Streptokinase binds preferentially to the extended conformation of plasminogen through lysine binding site and catalytic domain interactions. *Biochemistry* **39**, 13974-13981 (2000).

42. Phillips, G.J., Flicker, P., Cohen, C., Manjula, B. & Fischetti, V.A. Streptococcal M protein: alpha-helical coiled-coil structure and arrangement on the cell surface. *Proc Natl Acad Sci U S A* **78**, 4689-93 (1981).

43. O'Reilly, M.S., Holmgren, L., Shing, Y., Chen, C., Rosenthal, R. A., Moses, M., Lane, W. S., Cao, Y., Sage, E. H., and Folkman, J. Angiostatin: a novel angiogenesis inhibitor that mediates the suppression of metastases by a Lewis lung carcinoma. *Cell* **79**, 315-28 (1994).

44. O'Reilly, M.S., Holmgren, L., Shing, Y., Chen, C., Rosenthal, R. A., Cao, Y., Moses, M., Lane, W. S., Sage, E. H., and Folkman, J. Angiostatin: a circulating endothelial cell inhibitor that suppresses angiogenesis and tumor growth. *Cold Spring Harb Symp Quant Biol* **59**, 471-82 (1994).

45. Folkman, J. Angiogenesis in cancer, vascular, rheumatoid and other disease. *Nat Med* **1**, 27-31 (1995).

46. Folkman, J. & Shing, Y. Angiogenesis. *J Biol Chem* **267**, 10931-4 (1992).

47. MacDonald, N.J., Murad, A. C., Fogler, W. E., Lu, Y. Y., and Sim, B. K. L. The tumor-suppressing activity of angiostatin protein resides within kringles 1 to 3. *Biochem Biophys Res Commun* **264**, 469-477 (1999).

48. Stathakis, P., Fitzgerald, M., Matthias, L.J., Chesterman, C.N. & Hogg, P.J. Generation of angiostatin by reduction and proteolysis of plasmin. Catalysis by a plasmin reductase secreted by cultured cells. *J Biol Chem* **272**, 20641-5 (1997).

49. Stathakis, P. et al. Angiostatin formation involves disulfide bond reduction and proteolysis in kringle 5 of plasmin. *J Biol Chem* **274**, 8910-6 (1999).

50. Lay, A.J., Jiang, X.M., Daly, E., Sun, L. & Hogg, P.J. Plasmin reduction by phosphoglycerate kinase is a thiol-independent process. *J Biol Chem* **277**, 9062-8 (2002).

51. Lay, A.J. & Hogg, P.J. Measurement of reduction of disulfide bonds in plasmin by phosphoglycerate kinase. *Methods Enzymol* **348**, 87-92 (2002).

52. Lay, A.J. et al. Phosphoglycerate kinase acts in tumour angiogenesis as a disulphide reductase. *Nature* **408**, 869-73 (2000).

53. Dong, Z.Y., Kumar, R., Yang, X.L. & Fidler, I.J. Macrophage-derived metalloelastase is responsible for the generation of angiostatin in Lewis lung carcinoma. *Cell* **88**, 801-810 (1997).

54. Patterson, B.C. & Sang, Q.X.A. Angiostatin-converting enzyme activities of human matrilysin (MMP-7) and gelatinase B type IV collagenase (MMP-9). *J Biol Chem* **272**, 28823-28825 (1997).

55. Lijnen, H.R., Ugwu, F., Bini, A. & Collen, D. Generation of an angiostatin-like fragment from plasminogen by stromelysin-1 (MMP-3). *Biochemistry* **37**, 4699-4702 (1998).

56. Cornelius, L.A. et al. Matrix metalloproteinases generate angiostatin: Effects on neovascularization. *J. Immunol.* **161**, 6845-6852 (1998).

57. O'Reilly, M.S., Wiederschain, D., Stetler-Stevenson, W.G., Folkman, J. & Moses, M.A. Regulation of angiostatin production by matrix metalloproteinase-2 in a model of concomitant resistance. *J Biol Chem* **274**, 29568-29571 (1999).

58. Gonzalez-Gronow, M., Grenett, H.E., Weber, M.R., Gawdi, G. & Pizzo, S.V. Interaction of plasminogen with dipeptidyl peptidase IV initiates a signal transduction mechanism which regulates expression of matrix metalloproteinase-9 by prostate cancer cells. *Biochem J.* **355**, 397-407 (2001).

59. Cao, Y.H., Ji, R. W., Davidson, D., Schaller, J., Marti, D., Sohndel, S., McCance, S. G., Oreilly, M. S., Llinas, M., and Folkman, J. Kringle domains of human angiostatin - Characterization of the anti-proliferative activity on endothelial cells. *J Biol Chem* **271**, 29461-29467 (1996).

60. Gately, S. et al. The mechanism of cancer-mediated conversion of plasminogen to the angiogenesis inhibitor angiostatin. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 10868-10872 (1997).

61. Heidtmann, H.H. et al. Generation of angiostatin-like fragments from plasminogen by prostate-specific antigen. *Br J Cancer* **81**, 1269-73 (1999).

62. Fortier, A.H. et al. Recombinant prostate specific antigen inhibits angiogenesis in vitro and in vivo. *Prostate* **56**, 212-219 (2003).

63. Abad, M.C. et al. The X-ray crystallographic structure of the angiogenesis inhibitor angiostatin. *J Mol Biol* **318**, 1009-17 (2002).

64. Lee, H., Kim, H. K., Lee, J. H., You, W. K., Chung, S. I., Chang, S. I., Park, M. H., Hong, Y. K., and Joe, Y. A. Disruption of interkringle disulfide bond of plasminogen kringle 1-3 changes the lysine binding capability of kringle 2, but not its antiangiogenic activity. *Arch Biochem Biophys* **375**, 359-63 (2000).

65. Risau, W. Differentiation of endothelium. *Faseb J* **9**, 926-933 (1995).

66. Pepper, M.S. Extracellular proteolysis and angiogenesis. *Thromb Haemost* **86**, 346-55 (2001).

67. O'Reilly, M.S., Holmgren, L., Chen, C. & Folkman, J. Angiostatin induces and sustains dormancy of human primary tumors in mice. *Nat Med* **2**, 689-92 (1996).

68. Moser, T.L. et al. Angiostatin binds ATP synthase on the surface of human endothelial cells. *Proc Natl Acad Sci U S A* **96**, 2811-6 (1999).

69. Moser, T.L. et al. Endothelial cell surface F1-F0 ATP synthase is active in ATP synthesis and is inhibited by angiostatin. *Proc Natl Acad Sci U S A* **98**, 6656-61 (2001).

70. Moser, T.L., Stack, M.S., Wahl, M.L. & Pizzo, S.V. The mechanism of action of angiostatin: can you teach an old dog new tricks? *Thromb Haemost* **87**, 394-401 (2002).

71. Tarui, T., Miles, L.A. & Takada, Y. Specific interaction of angiostatin with integrin alpha(v)beta(3) in endothelial cells. *J Biol Chem* **276**, 39562-8 (2001).

72. Tarui, T., Majumdar, M., Miles, L.A., Ruf, W. & Takada, Y. Plasmin-induced migration of endothelial cells. A potential target for the anti-angiogenic action of angiostatin. *J Biol Chem* **277**, 33564-70 (2002).

73. Reynolds, L.E. et al. Enhanced pathological angiogenesis in mice lacking beta3 integrin or beta3 and beta5 integrins. *Nat Med* **8**, 27-34 (2002).

74. Troyanovsky, B., Levchenko, T., Mansson, G., Matvijenko, O. & Holmgren, L. Angiomotin: an angiostatin binding protein that regulates endothelial cell migration and tube formation. *J Cell Biol* **152**, 1247-54 (2001).

75. Bratt, A. et al. Angiomotin belongs to a novel protein family with conserved coiled-coil and PDZ binding domains. *Gene* **298**, 69-77 (2002).

76. Akasaka, T., van Leeuwen, R.L., Yoshinaga, I.G., Mihm, M.C., Jr. & Byers, H.R. Focal adhesion kinase (p125FAK) expression correlates with motility of human melanoma cell lines. *J Invest Dermatol* **105**, 104-8 (1995).

77. Plopper, G.E., McNamee, H.P., Dike, L.E., Bojanowski, K. & Ingber, D.E. Convergence of integrin and growth factor receptor signaling pathways within the focal adhesion complex. *Mol Biol Cell* **6**, 1349-65 (1995).

78. Claesson-Welsh, L. et al. Angiostatin induces endothelial cell apoptosis and activation of focal adhesion kinase independently of the integrin-binding motif RGD. *Proc Natl Acad Sci U S A* **95**, 5579-83 (1998).

79. Xiong, J.P. et al. Crystal structure of the extracellular segment of integrin alpha Vbeta3. *Science* **294**, 339-45 (2001).

80. Abrahams, J.P., Leslie, A.G., Lutter, R. & Walker, J.E. Structure at 2.8 A resolution of F1-ATPase from bovine heart mitochondria. *Nature* **370**, 621-8 (1994).

81. Brown, J.E. & Klee, W.A. Helix-coil transition of the isolated amino terminus of ribonuclease. *Biochemistry* **10**, 470-6 (1971).

82. Siedlecka, M., Goch, G., Ejchart, A., Sticht, H. & Bierzyski, A. Alpha-helix nucleation by a calcium-binding peptide loop. *Proc Natl Acad Sci U S A* **96**, 903-8 (1999).

83. Ghadiri, M.R. & Choi, C. Secondary structure nucleation in peptides. Transition metal ion stabilized .alpha.-helices. *J Am Chem Soc* **112**, 1630-2 (1990).

84. Ghadiri, M.R. & Fernholz, A.K. Peptide architecture. Design of stable .alpha.-helical metallopeptides via a novel exchange-inert ruthenium(III) complex . *J Am Chem Soc* **112**, 9633-5 (1990).

85. Ruan, F., Chen, Y. & Hopkins, P.B. Metal ion-enhanced helicity in synthetic peptides containing unnatural, metal-ligating residues. *J Am Chem Soc* **112**, 9403-4 (1990).

86. Ghosh, D. & Pecoraro, V.L. Understanding metalloprotein folding using a de novo design strategy. *Inorg Chem* **43**, 7902-15 (2004).

87. Doerr, A.J. & McLendon, G.L. Design, folding, and activities of metal-assembled coiled coil proteins. *Inorg Chem* **43**, 7916-25 (2004).

88. Li, X.Q. et al. Soft metal ions, Cd(II) and Hg(II), induce triple-stranded alpha-helical assembly and folding of a de novo designed peptide in their trigonal geometries. *Protein Sci* **9**, 1327-33 (2000).

89. Kiyokawa, T. et al. Binding of Cu(II) or Zn(II) in a de novo designed triple-stranded alpha-helical coiled-coil toward a prototype for a metalloenzyme. *J Pept Res* **63**, 347-53 (2004).

90. Handel, T., Williams, S.A. & Degrado, W.F. Metal ion-dependent modulation of the dynamics of a designed protein. *Science* **261**, 879-85 (1993).

91. Kohn, W.D., Kay, C.M., Sykes, B.D. & Hodges, R.S. Metal Ion Induced Folding of a de Novo Designed Coiled-Coil Peptide. *J Am Chem Soc* **120**, 1124-32 (1998).

92. Terlau, H. & Olivera, B.M. Conus venoms: a rich source of novel ion channel-targeted peptides. *Physiol Rev* **84**, 41-68 (2004).

93. Jimenez, E.C., Olivera, B.M., Gray, W.R. & Cruz, L.J. Contryhan is a D-tryptophan-containing *Conus* peptide. *J Biol Chem* **281**, 28002-28005 (1996).

94. Loughnan, M. et al. α-Conotoxin, EpI, a novel sulfated peptide from *Conus episcopatus* that selectively targets neuronal nicotinic acetylcholine receptors. *J Biol Chem* **273**, 15667-15674 (1998).

95. Jimenez, E.E. et al. Bromocontryphan: post-translational bromination of tryptophan. *Biochemistry* **36**, 989-994 (1997).

96. Craig, A.G. et al. An O-glycosylated neuroexcitatory *Conus* peptide. *Biochemistry* **37**, 16019-16025 (1998).

97. McIntosh, J.M., Olivera, B.M., Cruz, L.J. & Gray, W.R. y-Carboxy-glutamate in a neuroactive toxin. *J Biol Chem* **259**, 14343-14346 (1984).

98. McIntosh, J., Olivera, B.M., Cruz, L. & Gray, W. Gamma-carboxyglutamate in a neuroactive toxin. *J Biol Chem* **259**, 14343-6 (1984).

99. Olivera, B.M. et al. A sleep-inducing peptide from *Conus geographus* venom. *Toxicon* **23**(1985).

100. Olivera, B.M., Cruz, L.J. & Yoshikami, D. Effects of *Conus* peptides on the behavior or mice. *Curr Opin Neurobiol* **9**, 772-777 (1999).

101. Castellino, F.J. & Prorok, M. Conantokins: inhibitors of ion flow through the N-methyl-D-aspartate receptor channels. *Curr Drug Targets* **1**, 219-35 (2000).

102. Warder, S.E., Blandl, T., Klein, R.C., Castellino, F.J. & Prorok, M. Amino acid determinants for NMDA receptor inhibition by conantokin-T. *J Neurochem* **77**, 812-22 (2001).

103. Mcintosh, J.M., Olivera, B.M., Cruz, L.J. & Gray, W.R. γ-carboxy-glutamate in a neuroactive toxin. *J Biol Chem* **259**, 14343-14346 (1984).

104. Prorok, M., Warder, S.E., Blandl, T. & Castellino, F.J. Calcium binding properties of synthetic gamma-carboxyglutamic acid-containing marine cone snail "sleeper" peptides, conantokin-G and conantokin-T. *Biochemistry* **35**, 16528-34 (1996).

105. Haack, J.A. et al. Conantokin-T. A gamma-carboxyglutamate containing peptide with N-methyl-d-aspartate antagonist activity. *J Biol Chem* **265**, 6025-9 (1990).

106. Skjaerbaek, N., Nielsen, K.J., Lewis, R.J., Alewood, P. & Craik, D.J. Determination of the solution structures of conantokin-G and conantokin-T by CD and NMR spectroscopy. *J Biol Chem* **272**, 2291-9 (1997).

107. Warder, S.E. et al. The roles of individual gamma-carboxyglutamate residues in the solution structure and cation-dependent properties of conantokin-T. *J Biol Chem* **273**, 7512-22 (1998).

108. Blandl, T., Prorok, M. & Castellino, F.J. NMDA-receptor antagonist requirements in conantokin-G. *FEBS Lett* **435**, 257-62 (1998).

109. Ault, B., Evans, R.H., Francis, A.A., Oakes, D.J. & Watkins, J.C. Selective depression of excitatory amino acid induced depolarizations by magnesium ions in isolated spinal cord preparations. *J Physiol* **307**, 413-428 (1980).

110. Skolnick, P., Boje, K., Miller, R., Pennington, M. & Maccecchini, M.L. Noncompetitive inhibition of N-methyl-D-aspartate by conantokin-G: evidence for an allosteric interaction at polyamine sites. *J Neurochem* **59**, 1516-21 (1992).

111. Mena, E.E. et al. Conantokin-G: a novel peptide antagonist to the N-methyl-D-aspartic acid (NMDA) receptor. *Neurosci Lett* **118**, 241-4 (1990).

112. Hammerland, L.G., Olivera, B.M. & Yoshikami, D. Conantokin-G selectively inhibits N-methyl-D-aspartate-induced currents in Xenopus oocytes injected with mouse brain mRNA. *Eur J Pharmacol* **226**, 239-44 (1992).

113. Huang, C.C., Lyu, P.C., Lin, C.H. & Hsu, K.S. Conantokin-T selectively antagonizes N-methyl-D-aspartate-evoked responses in rat hippocampal slice. *Toxicon* **35**, 355-63 (1997).

114. Wo, Z.G. & Oswald, R.E. Transmembrane topology of two kainate receptor subunits revealed by N-glycosylation. *Proc Natl Acad Sci U S A* **91**, 7154-7158 (1994).

115. Williams, K., Kashiwagi, K. & Fukuchi, J. An acidic amino acid in the N-methyl-D-aspartate receptor that is important for spermine stimulation. *Mol Pharmacol* **48**, 1087-1098 (1995).

116. Zheng, X., Zhang, L., Durand, G.M., Bennett, M.V. & Zukin, R.S. Mutagenesis rescues spermine and Zn2+ potentiation of recombinant NMDA receptors. *Neuron* **12**, 811-818 (1994).

117. Kuryatov, A., Laube, B., Betz, H. & Kuhse, J. Mutational analysis of the glycine-binding site of the NMDA receptor: structural similarity with bacterial amino acid-binding proteins. *Neuron* **12**, 1291-1300 (1994).

118. Laube, B., Kuryatov, A., Kuhse, J. & Betz, H. Glycine-glutamate interactions at the NMDA receptor: role of cysteine residues. *FEBS Lett* **335**, 331-334 (1993).

119. Stern-Bach, Y. et al. Agonist selectivity of glutamate receptors is specified by two domains structurally related to bacterial amino acid-binding proteins. *Neuron* **13**, 1345-1357 (1994).

120. Kendrick, S.J., Lynch, D.R. & Pritchett, D.B. Characterization of glutamate binding sites in receptors assembled from transfected NMDA receptor subunits. *J Neurochem* **67**, 608-616 (1996).

121. Sakurada, K., Masu, M. & Nakanishi, S. Alteration of Ca2+ permeability and sensitivity to Mg2+ and channel blockers by a single amino acid substitution in the N-methyl-D-aspartate receptor. *J Biol Chem* **268**(1993).

122. Mori, H., Masaki, H., Yamakura, T. & Mishina, M. Identification by mutagenesis of a Mg(2+)-block site of the NMDA receptor channel. *Nature* **358**, 673-675 (1992).

123. Kawajiri, S. & Dingledine, R. Multiple structural determinants of voltage-dependent magnesium block in recombinant NMDA receptors. *Neuropharmacology* **32**, 1203-1211 (1993).

124. Wood, M.W., VanDongen, H.M. & VanDongen, A.M. An alanine residue in the M3-M4 linker lines the glycine binding pocket of the N-methyl-D-aspartate receptor. *J Biol Chem* **272**, 3532-3537 (1997).

125. Hirai, H., Kirsch, J., Laube, B., Betz, H. & Kuhse, J. The glycine binding site of the N-methyl-D-aspartate receptor subunit NR1: identification of novel determinants of co-agonist potentiation in the extracellular M3-M4 loop region. *Proc Natl Acad Sci U S A* **93**, 6031-6036 (1996).

126. Kashiwagi, K., Fukuchi, J., Chao, J., Igarashi, K. & Williams, K. An aspartate residue in the extracellular loop of the N-methyl-D-aspartate receptor controls sensitivity to spermine and protons. *Mol Pharmacol* **49**, 1131-1141 (1996).

127. Tingley, W.G., Roche, K.W., Thompson, A.K. & Huganir, R.L. Regulation of NMDA receptor phosphorylation by alternative splicing of the C-terminal domain. *Nature* **364**, 70-73 (1993).

128. Sugihara, H., Moriyoshi, K., Ishii, T., Masu, M. & Nakanishi, S. Structures and properties of seven isoforms of the NMDA receptor generated by alternative splicing. *Biochem Biophys Res Commun* **185**, 826-832 (1992).

129. Nielsen, K.J. et al. Structure-activity studies of conantokins as human N-methyl-D-aspartate receptor modulators. *J Med Chem* **42**, 415-26 (1999).

130. Donevan, S.D. & McCabe, R.T. Conantokin G is an NR2B-selective competitive antagonist of N-methyl-D-aspartate receptors. *Mol Pharmacol* **58**, 614-23 (2000).

131. Klein, R.C., Prorok, M., Galdzicki, Z. & Castellino, F.J. The amino acid residue at sequence position 5 in the conantokin peptides partially governs subunit-selective antagonism of recombinant N-methyl-D-aspartate receptors. *J Biol Chem* **276**, 26860-7 (2001).

Chapter II

X-ray Structure Determination

2.1 Structure Determination from X-ray Diffraction Data

2.1.1 Theory

When a crystal is exposed to X-rays, constructive interferences between rays scattered from successive planes in the crystal will only take place if the path differences between the rays are equal to an integral number of wavelengths. This is known as Bragg's law:

$$2d \sin\theta = n\lambda \tag{2.1}$$

The Bragg equation gives the condition for diffraction so that if a crystal is rotated in a beam of X-rays, the scattering pattern is a series of intensity maxima. In a crystal, electrons in atoms are the scatters, and each atom has a different effectiveness as a scatter. Consequently, when an experiment is carried out, a set of diffraction maxima of different intensities are observed. The crystal is rotated to obtain the scattering intensity at various angles. The scattering intensity depends on the scattering effectiveness of the individual atoms and the phase of the wave from each scattering source, known as individual structure factor (fj). The structure factor, $\mathrm{F}(hkl)$, for each plane (hkl) can be defined as the sum of the structure factors for individual atoms, fj, times a phase factor, $\alpha(hkl)$, for each atom. In other words, the structure factor can be represented by its amplitude and phase:

$$F(hkl) = \sum f_j e^{2\pi i(hxj+kyj+lzj)} = F(hkl)e^{i\alpha(hkl)} \quad (2.2)$$

where, $F(hkl)$ is the amplitude and $\alpha(hkl)$ is the phase.

When the diffracted X-ray is recorded, all information on the phase is lost and only a measurement of the intensity of the diffracted beam is recorded. The intensity in each spot of the diffraction pattern is given by:

$$I(hkl) = \left[F(hkl)\right]^2 \quad (2.3)$$

The electron density, $\rho(r)$, is a function of the coordinates of the scattering centers (the atoms) and has a maximum around the position of each atom. It is desirable to convert the measured structure factors into atomic coordinates. This is done by taking the Fourier transform of equation 2.2. In this case, the Fourier transform takes the structure factors, which are functions of the electron density, and inverts the functional dependence so that the electron density is expressed as a

$$\rho(r) = \int_{diffraction\ space} F(s)e^{-2\pi ir\bullet s} dv_s \quad (2.4)$$

function of the structure factors:

where, dv_s is a small unit of volume in diffraction space. The integration can be replaced by a summation since $F(s)$ is not continuous and is non-zero only at the reciprocal lattice points. Therefore:

where, $\rho(xyz)$ is the electron density at any point x,y,z and $F(hkl)$ is the amplitude, which is proportional to the square root of the measured intensity of each reflection, labeled *hkl*.

$$\rho(\mathrm{xyz}) = \frac{1}{V}\sum_h\sum_k\sum_l \mathrm{F}(hkl)e^{-2\pi i(hx+ky+lz)}$$

$$= \frac{1}{V}\sum_h\sum_k\sum_l \mathrm{F}(hkl)e^{i\alpha(hkl)}e^{-2\pi i(hx+ky+lz)} \qquad (2.5)$$

The problem is that only the intensity, which is the square of the amplitude, can be directly derived from the measured intensity of the diffracted beam. Since only intensities but not phases are measured in the recorded diffraction pattern it is impossible to determine the electron density, and therefore a structure, directly from a recorded diffraction pattern. This is a major problem in crystallography and is referred to as the "*phase problem*"[1,2].

There are few methods by which the phase problem can be solved. (1) The Patterson summation: This is a Fourier summation based on the experimentally observed $\left[\mathrm{F}(hkl)\right]^2$. It is basically a vector map of the structure and is applied for relatively small molecules. (2) Direct methods: In this method mathematical relationships between the reflections can be used to provide phase information. (3) Heavy atom isomorphous replacement: In this method a heavy atom is introduced to a structure to provide phase information. (4) Anomalous scattering: In this method phase information is obtained from the information contained in the scattering by an atom, when its natural absorption frequency is close to the frequency of the incident radiation. (5) Molecular replacement method: A known structure will be used to find the phase of the unknown structure. However molecular replacement is limited because to have a good chance of success in finding a solution, the search and target molecules must have a reasonable sequence identity (>25%). Likewise, having data

with high completeness can be crucial. Most of the time, but not always, molecular replacement seems to be relatively easier than the other methods and is the first choice in solving the phase problem.

Since many crystal structures of kringles have already been solved including the multiple kringle containing compound angiostatin[3], we could easily use molecular replacement to determine the initial phases of the angiostatin/VEK-30 complex structure using angiostatin and/or its individual kringle components as models. Since the structures of metal bound conantokins-G and -T[K7γ] are known to adopt α-helical conformations[4,5], any α-helix can be used as a model. This will be further discussed in this chapter.

Generally there are two steps in molecular replacement and these are known as the rotation and translation functions.

Rotation Function

The rotation function should allow the orientation of the search molecule, which produces a maximal overlap with the target structure to be determined in the absence of any phases for the unknown structure. To do this, it compares the Patterson self-vectors of the known and unknown structures at different orientations of the search model. It should be noted that Patterson functions can be calculated from the amplitudes only and using Patterson space so that the translation vector is irrelevant, since all intramolcular vectors are shifted to the origin. The rotation function is usually calculated as a function of Eulerian angles, α, β and γ. The molecule is placed in an orthogonal coordinate system with the axis of highest symmetry along Z (about α) to reduce the amount of computation.

Translation Function

Having determined the angles, α, β and γ, from the rotation search the rotation matrix can be used and applied to the coordinates of the search molecule. The shift vector, which is required to position the search molecule correctly relative to the symmetry elements of the space group, can be determined by one of a number of translation searches. Patterson methods can be used to measure the overlap of the target Patterson cross-vectors with those calculated for the oriented search molecule as it ranges through the target cell. The simpler way to solve the translation problem is the reliability factor (R-factor) search. It involves the calculation of an R-factor as the search molecule and its symmetry mates are moved through the unit cell of the target crystal. The correct position should give the lowest R-factor defined as:

$$R=\frac{\sum\left||F_{obs}|-|F_{cal}|\right|}{\sum|F_{obs}|} \qquad (2.6)$$

where, F_{obs} and F_{cal} are the observed and calculated structure factors, respectively. Basically when a crystal structure is known, values of $F(hkl)$ can be calculated and a way to test the correctness of the structure is how well the calculated values of $F(hkl)$ agree with the observed ones. Any random collection of atoms in the cell has been shown to result in an R-factor of 83% and 59% for centric and acentric space groups, respectively. Therefore any model that gives R-factors approaching these values is just a little better than a collection of atoms randomly placed in a cell. An R-factor around 45% tells us that the solution is not hopeless but major changes are

needed to fit the model. An R-factor of ~35% is likely to be a correct solution. An R-factor of 25% and below indicates that the model is most probably correct except for small (1.0 Å) atomic shifts and changes in temperature factors (B-factors).

Other parameters such as the Correlation Coefficient (CC) can be used to measure the agreement between the F_{obs} and F_{cal} as the search model is moved around. CC can be expressed as:

$$CC=\frac{\sum\left(|F_{obs}|^2-\overline{|F_{obs}|^2}\right)\left(|F_{cal}|^2-\overline{|F_{cal}|^2}\right)}{\left[\sum\left(|F_{obs}|^2-\overline{|F_{obs}|^2}\right)^2\sum\left(|F_{cal}|^2-\overline{|F_{cal}|^2}\right)\right]^{\frac{1}{2}}} \tag{2.7}$$

The CC runs from –1 (perfect inverse correlation), through 0 (no correlation), to +1 (perfect correlation). It has the advantages of being almost independent of scaling between F_{obs} and F_{cal}, and is much more sensitive than the R-factor in the region where the R-factor approaches its random limits. Conversely, the coefficient approaches 1.0 closely as the R-factor goes below 20%, and thus becomes of limited value[6].

If successful, a preliminary model of the target structure will be obtained by correctly orienting and positioning the search molecule in the target cell. Subsequently, this model can be refined and optimized by rigid body refinement, restrained refinements, as well as manual refitting of the models.

A number of programs exist for performing molecular replacement. These include AMoRe[7], MolRep within CCP4[8], CaspR[9], X-PLOR[10], and Phaser[11]. All of these programs except Phaser[11] output R-factors and Correlation Coefficients.

Instead Phaser[11] outputs Z-scores and the log-likelihood gain (LLG). The likelihood is the probability that the data would have been measured, given the model, so it allows us to compare how well different models agree with the data. The LLG is the difference between the likelihood of the model and the likelihood calculated from a Wilson distribution, so it measures how much better the data can be predicted with your model than with a random distribution of the same atoms. The LLG allows us to compare different models against the same data set, but the LLG values for different data sets should not be compared with each other. If the best LLG is negative, then the model is worse than a collection of random atoms. The LLG should always be positive, and it should increase as the solution progresses. By default, Phaser[11] selects solutions over 75% of the difference between the top solution and the mean. Ideally, only the number of solutions you are expecting should be selected by this criterion, but if the signal-to-noise of your search is low, there will be noise peaks in this selection also. For a translation function, the correct solution will generally have a Z-score (number of standard deviations above the mean value) over 5 and be well separated from the rest of the solutions. For a rotation function, the correct solution may be in the list with a Z-score under 4, and will not be found until a translation function is performed and picks out the correct solution.

Phaser[11] proved to be the only molecular replacement program capable of providing a correct solution for the metal bound structure of conantokin-G. This is because the other programs rely on a spherically shaped protein whereas in the case of conantokin-G, the protein was not spherically shaped but helical. So, the

intermolecular vectors are extremely limited whereas the intramolecular vectors are not.

As mentioned above, heavy atom isomorphous replacement is another method of solving the phase problem. In this technique, differences in the diffraction pattern are measured after the introduction of a heavy atom. In order to be able to record these differences reliably, the atom or group of atoms must have many electrons. If hard ions such as lanthanides or actinides are used as derivatives, the interaction between the protein and the heavy atom is usually ionic. On the other hand, soft ions such as mercury, gold, platinum, etc. tend to react with sulfurs on cysteines, deprotonated nitrogen on histidines or even with the sulfur from methionine. These interactions are covalent, so they tend to bind more specifically. For this technique to work, it is important that the protein molecules within the crystal be identically bound to the heavy atom compound with essentially no change in the structure of the native crystal lattice.

The coordinates and the diffraction pattern of the heavy atom alone can be determined by calculating the difference between the diffraction patterns of the native and heavy atom replaced crystal. This is because the contributions from every atom to a reflection are combined in an additive way. With the diffraction pattern of just the heavy atom, a small number of atoms exist and the structure is easier to solve. This technique was used in order to provide phases for K3 in the angiostatin/VEK-30 complex structure. This technique was also used in order to determine the structure of conantokin-G bound to Ca^{2+}. This will be further discussed below in this chapter.

Anomalous dispersion is another method mentioned above that was tried in order to determine the structure of Ca^{2+} bound conantokin-G and the Cd^{2+}/Mg^{2+}/con-T[K7γ] structure. Anomalous dispersion involves the use of anomalous scatterers in the protein. In many cases, the sulfur in methionine is substituted with selenium by expressing the protein in the presence of SeMet. Advantage is taken of the anomalous differences by irradiating the crystal with X-ray radiation at the absorption edge of the scatterer (0.98 Å for Se). However, in the case of the conantokin structures, this method was used in order to locate Ca^{2+} sites in the conantokin-G structure and Cd^{2+} sites in the Cd^{2+}/Mg^{2+}/con-T[K7γ] structure. This will be further described below.

Structure Refinement

Once one has overcome the phase problem and a solution is found for the structure, $2F_{obs} - F_{cal}$ and $F_{obs} - F_{cal}$ Fourier maps are calculated, and the atomic positions are taken as the locations of the electron density function. The advent of high-speed computers has led to widespread use of the method of least squares, which automatically adjusts the parameters so as to minimize some functions such as $\sum(F_{obs}-F_{cal})^2$. Each residue will be fitted manually in the structure until the bias introduced by the starting model is reduced considerably. A correct structure should have a satisfactory R-factor with no major unexplained discrepancies. A satisfactory value for the R-factor depends on the resolution of the data. The higher the resolution the lower the R-factor must be. Before refining the model, a fraction of the reflections, usually 5-10%, are randomly chosen and put aside for cross-validation. This set of randomly selected reflections is called the test set. All the

fitting of the residues and refinement will be done on the working set (95% or 90% of the data). Then two separate R-factors are calculated. The R-factor calculated for the working set is called "R_{work}" or "R_{real}" and that of the test set is called "R_{free}". As the R_{real} decreases upon model fitting, the R_{free} should concurrently decrease. Generally R_{free} is higher than R_{real}, however they should not differ by more than 6-7%[1,6].

Values of R_{work} and R_{free} are measures of the agreement between the values of the observed structure factors, given by these equations:

$$R_{work} = \frac{\sum \| |F_{obs}| - |F_{cal}| \|}{\sum |F_{obs}|} \tag{2.8}$$

$$R_{free} = \frac{\sum \| |F_{obs}| - |F_{cal}| \|}{\sum |F_{obs}|} \tag{2.9}$$

2.1.2 Overview of Crystallography

The first step in every structure determination is the production of single and well diffracting crystals. For this purpose it is essential to have access to large quantities of highly pure material. The protein is set up for crystallization using sparse matrices of precipitating solutions that have been proven successful in crystallizing other proteins. The most common method for macromolecular crystallization is the hanging drop vapor diffusion method (Figure 2-1). In this method, a drop containing a mixture of protein and precipitating solution is equilibrated over a reservoir containing the precipitating solution. The protein is slowly precipitated over time to where the molecules adopt identical orientations in order to form a three dimensional array of molecules that are held together by non-

covalent interactions. A well-formed crystalline lattice aids in better diffracting crystals. The crystallization process involves setting up thousands of protein/precipitant drops as well as monitoring of each of these drops. When monitoring these drops, it is important to look for precipitation behavior, relative solubility, and the appearance of crystals. Based on the observations from the initial screens, new optimized screens can be made. This can turn into an iterative process that will hopefully yield crystals suitable for X-ray diffraction data collection. This strategy was followed in the crystallization of the angiostatin/VEK-30 complex and metal bound structures of con-G and con-T[K7γ].

2.2 Angiostatin/VEK-30 Complex

2.2.1 Crystallization and Data Collection of the $P6_122$ Structure

The human angiostatin mutant N289E (this mutant lacks N-linked glycosylation) containing K1-3 was expressed in *Pichia pastoris* and purified as previously described[12]. The purified protein was provided by EntreMed Inc. (Rockville, MD) and was buffer exchanged into saline buffer (0.15 M NaCl/100 mM Tris pH 7.5) and concentrated to 15 mg/ml. The VEK-30 peptide was synthesized and purified as previously described and provided by our collaborators Frank Castellino and co-workers[13]. An angiostatin/VEK-30 complex was made using a five molar excess of VEK-30 compared to that of angiostatin. The protein was extensively screened for crystallization by using the hanging drop vapor diffusion method (Figure 2-1). The search for well diffracting crystals was done by using several crystallization screens at two different temperatures, 25°C and 4°C. The best crystals of the angiostatin/VEK-30 complex were grown at 25°C with a precipitant

solution containing 20% PEG 8000/0.1 M potassium dihydrogen phosphate after optimizing the crystallization condition from Hampton Research Crystal Screen #42 (20% PEG 8000/0.05 M potassium dihydrogen phosphate). However, these crystals were not single (Figure 2-2). In order to grow single crystals of the angiostatin/VEK-30 complex, multiple crystal screen additives were tried. The best single crystals were grown in 20% PEG 8000/0.1 M potassium dihydrogen phosphate/5% 1,4-dioxane. Diamond-like crystals appeared in 1 day and continued to increase in size for 2-3 weeks (Figure 2-3).

The crystals were briefly soaked in a cryoprotectant solution of 22% PEG 8000/0.1 M potassium dihydrogen phosphate/5% 1,4-dioxane/30% glycerol at 298 K and flash frozen by immersion in liquid N_2. Data were collected at the Advanced Photon Source IMCA-CAT 17-ID at Argonne National Laboratory to a resolution of 2.0 Å and data were processed and scaled using the HKL suite of programs[14]. Crystals were in the hexagonal space group $P6_122$ with unit cell parameters a = b = 58.4 Å and c = 391.0 Å. The crystal to detector distance was 270 mm and 200° of data were collected with an oscillation range of 0.5°. Assuming one molecule of angiostatin/VEK-30 complex per asymmetric unit, the crystal volume per protein mass is 2.9, which corresponds to approximately 57.5% solvent content in the crystal. This value is within the range observed for protein crystals. The crystal parameters of the angiostatin/VEK-30 crystal are listed in Table 2.1. A synchrotron X-ray diffraction data set to a resolution of 2.0 Å, with an overall I/σ of 45.12, was obtained. The data was 80.2% complete with an R_{merge} of 9.1% for 28,416 unique

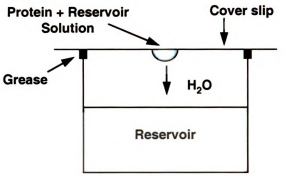

Figure 2-1 The hanging drop vapor diffusion method for crystallizing proteins.

Figure 2-2 A non-single crystal of the angiostatin/VEK-30 complex grown in 20% PEG 8000/0.1 M potassium dihydrogen phosphate.

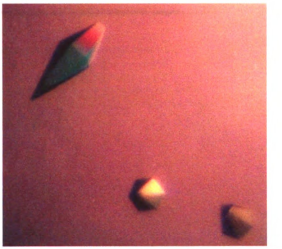

Figure 2-3 A single crystal of the angiostatin/VEK-30 complex grown in 20% PEG 8000/0.1 M potassium dihydrogen phosphate/5% 1,4-dioxane.

Table 2.1 Crystal parameters for the angiostatin/VEK-30 crystals in the space groups $P6_122$ and $P6_1$.

Crystal Form	Hexagonal	Hexagonal
Space Group	$P6_122$	$P6_1$
Unit Cell	a = b = 58.4 Å c = 391.0 Å α = β = 90.0° γ = 120.0°	a = b = 58.8 Å c = 389.2 Å α = β = 90.0° γ = 120.0°
Solvent Content	57.5%	57.4%
Mol. Per Asymmetric Unit	1	2

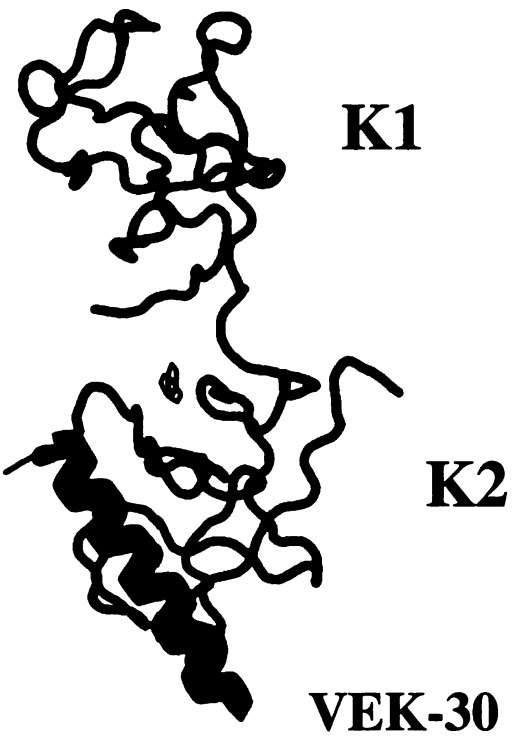

Figure 2-4 The overall structure of the angiostatin/VEK-30 complex. Angiostatin is shown in red and VEK-30 is shown in blue. One molecule of dioxane is shown in green with two different conformations.

Table 2.2 Data statistics for the angiostatin/VEK-30 diffraction data collection[a]

Space Group	$P6_122$	$P6_1$
Wavelength, Å	1.0000	1.0719
Resolution, Å	2.0 (2.07-2.00)	3.0 (3.20-3.04)
Completeness, %	80.2 (78.7)	78.5 (97.1)
I/σ	45.12 (4.8)	12.38 (3.02)
R_{merge}, %	9.1 (30.8)	8.4 (31.2)
Unique Reflections	28,416	18,600
Measured Reflections	763,409	281,739

[a]The parentheses denote those values for the last resolution shell.

Table 2.3 Refinement statistics for the angiostatin/VEK-30 complex

Space Group	$P6_122$	$P6_1$
R_{work}	20.85 %	20.17 %
R_{free}	25.45 %	29.56 %
Resolution	15-2.3 Å	20-3.0 Å
rmsd Bonds	0.01 Å	0.02 Å
rmsd Angles	1.58°	2.59°
Mean B value	25.6 $Å^2$	35.5 $Å^2$

reflections from a total of 763,409 measured reflections. Detailed data collection statistics are found in Table 2.2.

2.2.2 Structure Determination and Refinement of the $P6_122$ Structure

The structure was solved by molecular replacement using the AMoRe[7] program and the structures of K1 and K2 from human angiostatin as search models (PDB id 1KIO)[7]. A translation search with K2 gave one solution and a translation search with K1 after fixing the K2 translation solution also yielded one solution. The crystal packing of K2 and K1 was consistent with the location of K2 and K1 in the structure. This solution had a correlation factor of 41.9 and an R value of 44.2%, after rigid body refinement. Fixing the positions of angiostatin K1 and K2 and calculating an electron density map revealed density corresponding to the VEK-30 helix. The same electron density map revealed density connecting K1 and K2. Even after fixing K1, K2, and VEK-30, rotation and translation searches performed using human angiostatin K3 as a search model (PDB id 1KIO[3]) were unsuccessful. After refinement of K1, K2, and VEK-30, an electron density map revealed density for two residues in the K2-K3 linker peptide, residues T244 and T245, but no electron density was ever seen past residue T245. No electron density was seen for C297 of K3 at the inter-kringle K2-K3 disulfide bond. When C297 was built into the structure, negative density was calculated for residue C297 and an increase in R and R_{free} were observed. This indicated that residues 246 to 333 encompassing K3 are disordered in this structure. CNS[10] refinements of the angiostatin/VEK-30 structure yielded extremely high B-factors averaging 68 $Å^2$. In order to finish refining the structure and lower the average B-factors, TLS restrained refinements were carried

out within the CCP4[8] programs suite. Three TLS groups were defined using K1 in group 1, K2 in group 2, and VEK-30 in group 3. The refinement parameters are shown in Table 2.3. The final model (Figure 2-4) includes 165 residues from angiostatin (containing amino acids 81-245), 23 residues from VEK-30 (containing amino acids 88-110), and 321 water molecules. In addition to the water molecules, electron density for a single dioxane molecule, the additive used in crystallization, was located in two different conformations. A Ramachandran diagram is a plot of ϕ (angle between N and Cα) and φ (angle between C and Cα). According to geometry and steric restrictions the ϕ and φ angles must be within certain values. All residues with the exception of glycines and prolines, that don't have a side chain, must lie within these allowed regions. The red region is the most favored, bright yellow is the additionally allowed region, dull yellow is the generously allowed region, and white is the disallowed region. The Ramachandran plot of angiostatin in the complex structure contained 138 non-glycine, non-proline residues (80.4%) in the most favored regions and 15.9% in the additionally allowed regions with C169 and E165 in the disallowed region (Figure 2-5). The Ramachandran plot of VEK-30 in the complex contained 100% in the most favored regions (Figure 2-6). An example of the $2F_o$-F_c map contoured at 1σ is shown in Figure 2-7.

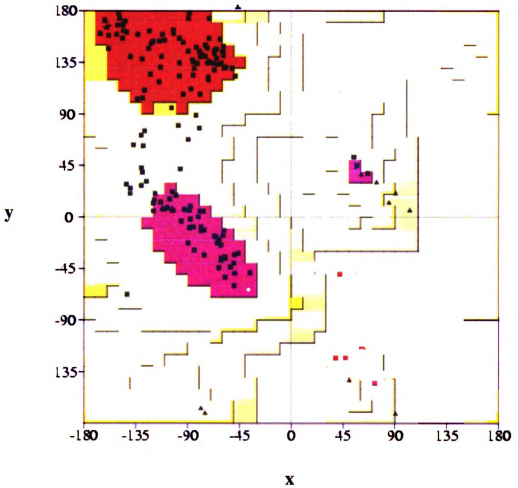

Figure 2-5 Ramachandran plot of angiostatin in the angiostatin/VEK-30 complex in the $P6_122$ space group. Phi (degrees) is x and Psi (Degrees) is y.

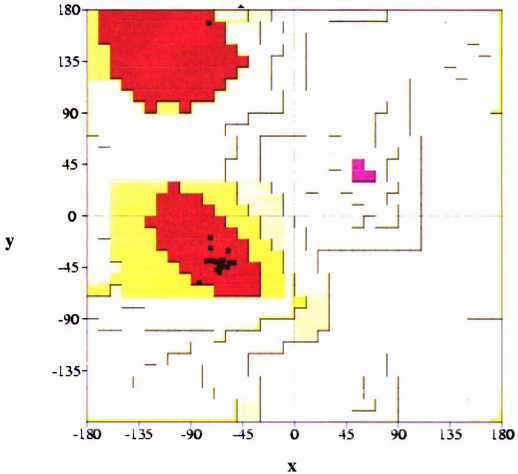

Figure 2-6 Ramachandran plot of VEK-30 in the angiostatin/VEK-30 complex in the $P6_122$ space group. Phi (degrees) is x and Psi (Degrees) is y.

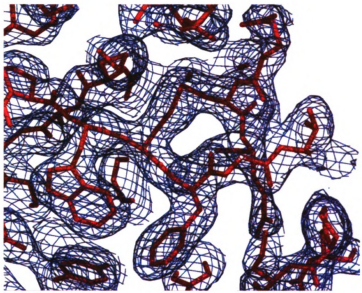

Figure 2-7 An example of the $2F_o$-F_c map contoured at 1σ of the angiostatin/VEK-30 complex in the $P6_122$ space group.

2.2.3 Crystallization and Data Collection of the $P6_1$ Structure

In order to find K3, the heavy atom isomorphous replacement was used. Heavy atom soaks were performed on angiostatin/VEK-30 complex crystals grown in the native crystal condition using 5 mM $Pt(C_5H_5N)_2Cl_2$ (*cis*-dichloro-bis(pyridine)-platinumII). After soaking for 24 hr, the crystal was back-soaked in its cryoprotectant solution. The crystal was flash frozen by immersion in liquid N_2. Data were collected at the DND-CAT 5-ID beamline at the APS to a resolution of 2.8 Å. The crystal to detector distance was 200 mm and 70° of data were collected with an oscillation of 0.5°. The data were processed and scaled using The HKL suite[15] in the space group $P6_1$ with unit cell parameters a = b = 58.8 Å and c = 389.2 Å. Assuming two molecules of angiostatin/VEK-30 complex per asymmetric unit, the crystal volume per protein mass is 2.9, which corresponds to approximately

57.4% solvent content in the crystal. This value is within the range observed for protein crystals. The crystal parameters of the angiostatin/VEK-30 $P6_1$ crystal are listed in Table 2.1. A synchrotron X-ray diffraction data set to a resolution of 3.0 Å, with an overall I/σ of 12.38, was obtained. The data was 78.5% complete with an R_{merge} of 8.4% for 18,600 unique reflections from a total of 281,739 measured reflections. Detailed data statistics are tabulated in Table 2.2.

2.2.4 Structure Determination and Refinement of the $P6_1$ Structure

Automated heavy atom Patterson searches using the program SOLVE[16] failed to locate heavy atom positions. Since the space group changed from the native crystal, molecular replacement was performed in order to find K3. Molecular replacement using MolRep with CCP4[8] and the $P6_122$ structure as a model yielded two translation solutions with a correlation of 46.6 and 73.6 and an R of 43.2% and 31.2%, respectively[8]. These two solutions represented the two molecules of K1, K2 and VEK-30 in the asymmetric unit of the $P6_1$ crystal form. Molecular replacement using K3 of angiostatin (PDB id 1KIO) as a model produced no solution. An electron density map was calculated and density for residues T244 and T245 of the K2-K3 linker was seen as in the $P6_122$ structure. No density was seen after residue T245. However, density was seen at the inter-kringle K2-K3 disulfide bond (C169-C297) for C297 of K3 for only one of the molecules in the asymmetric unit (Figure 2-8). Residues P296 and K298 were then built into corresponding density in the K3 disulfide region. However, no density was seen beyond K298 or before P296. Further refinement of the structure was carried out using TLS restrained refinement in CCP4[8] similar to that of the $P6_122$ structure. The final model includes 332

residues from angiostatin (containing amino acids 81-245 of two molecules A and B and 296-298 of molecule A), and 48 residues from VEK-30 (containing amino acids 88-110). The Ramachandran plots of the $P6_1$ structure are shown in Figure 2-9 and 2-10. Molecule A contained 139 non-glycine, non-proline residues (59%) in the most favored regions and 36% in the additionally allowed regions with E165, C169, and K212 in the disallowed region (Figure 2-9). Molecule B contained 137 non-glycine, non-proline residues (55.5%) in the most favored regions and 38.7% in the additionally allowed regions with E130, C169, and K212 in the disallowed region (Figure 2-10). The Ramachandran plots of the corresponding VEK-30 peptides are shown in Figure 2-11 and 2-12. Molecule A contained 22 non-glycine, non-proline residues (59.1%) in the most favored regions and 31.8% in the additionally allowed regions with T89 in the disallowed region (Figure 2-11). Molecule B contained 22 non-glycine, non-proline residues (72.7%) in the most favored regions and 22.7% in the additionally allowed regions (Figure 2-12). Refinement statistics are shown in Table 2.3. All model building was done using TURBO FRODO[17] and the refinement and map calculations were carried out using CNS[10] and CCP4[8,18]. The details of this structure will be discussed in Chapter III.

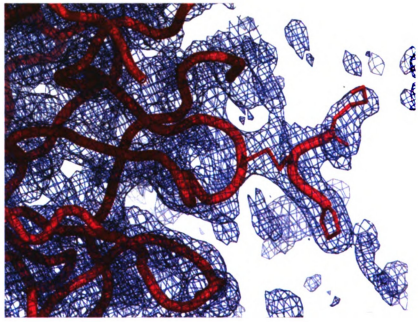

Figure 2-8 An example of the $2F_o$-F_c map contoured at 1.2σ of the angiostatin/VEK-30 complex $P6_1$ structure. The map is centered on the interkringle K2-K3 disulfide bond (C169-C297).

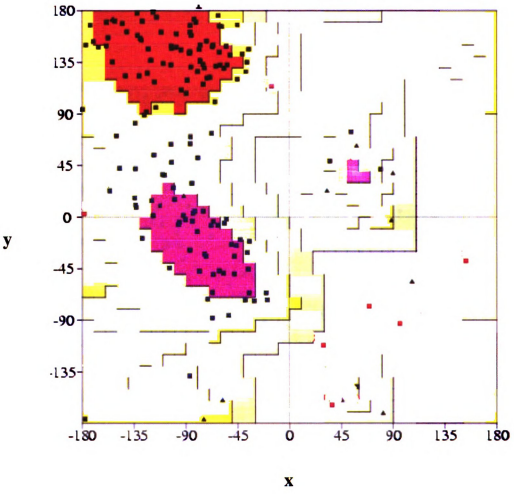

Figure 2-9 Ramachandran plot of angiostatin molecule A in the angiostatin/VEK-30 complex $P6_1$ structure. Phi (degrees) is x and Psi (Degrees) is y.

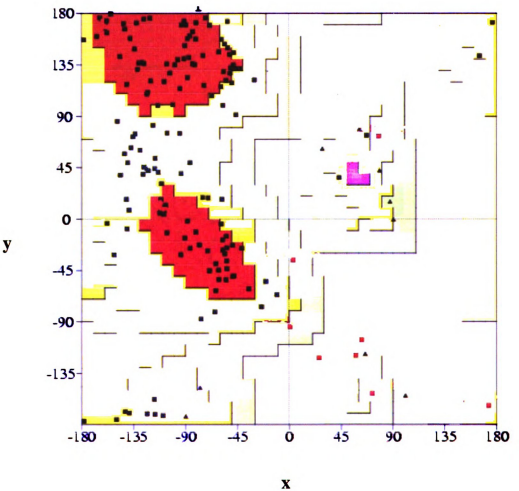

Figure 2-10 Ramachandran plot of angiostatin molecule B in the angiostatin/VEK-30 complex $P6_1$ structure. Phi (degrees) is x and Psi (Degrees) is y.

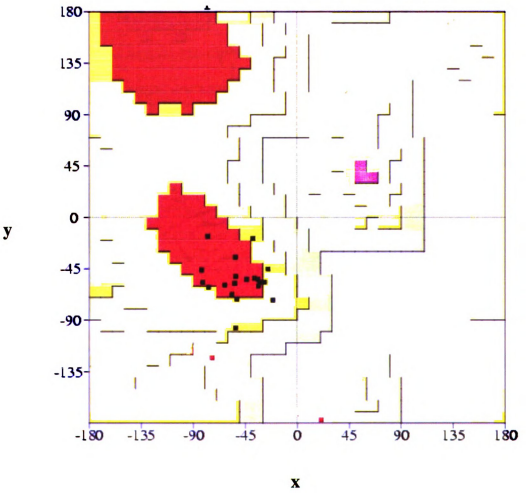

Figure 2-11 Ramachandran plot of VEK-30 molecule A in the angiostatin/VEK-30 complex $P6_1$ structure. Phi (degrees) is x and Psi (Degrees) is y.

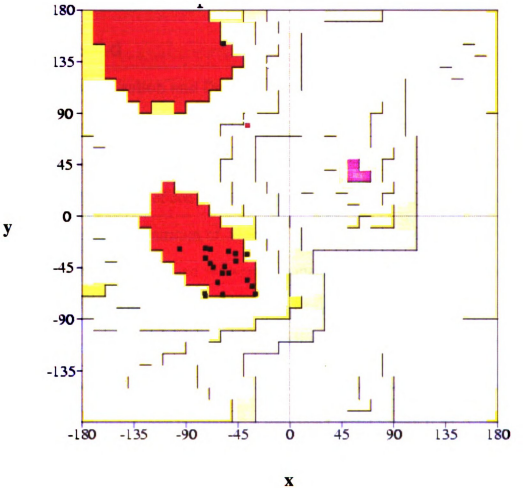

Figure 2-12 Ramachandran plot of VEK-30 molecule B in the angiostatin/VEK-30 complex $P6_1$ structure. Phi (degrees) is x and Psi (Degrees) is y.

2.3 Ca^{2+}/con-G

2.3.1 Crystallization and Data Collection

The con-G peptide was synthesized, purified, and characterized as previously described[5] and provided by Frank Castellino and co-workers in lyophilized form. The lyopholized solid of con-G was dissolved in 50 mM $CaCl_2$ and 100 mM Tris-HCl, pH 8, to a concentration of 10 mg/ml. Previous experiments had shown that the concentration of Ca^{2+} and pH were important in forming a dimeric structure of Ca^{2+}/con-G[5]. The search for well diffracting crystals was performed using several crystallization screens at two different temperatures, 25°C and 4°C. The best crystals of con-G were grown at 25°C by the hanging drop vapor diffusion method in 35% v/v 1,4-dioxane from the Hampton Research Crystal Screen II #4. Crystals first appeared within 24 hr and were highly variable in size (Figure 2-13).

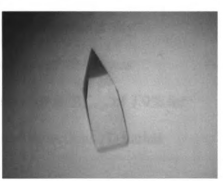

Figure 2-13 Crystal of Ca^{2+}/con-G.

Crystals of con-G were briefly soaked in a cryoprotectant solution containing 35% 1,4-dioxane and 30% MPD at 25°C and flash frozen by immersion in liquid N_2. X-ray diffraction data were collected at the Advanced Photon Source COM-CAT 32-ID at the Argonne National Laboratory to a resolution of 1.2 Å and were processed and scaled using the HKL suite of programs[14,15] in the tetragonal spacegroup $P4_222$ with unit cell parameters a = b = 29.3 Å and c = 46.9 Å. The crystal to detector distance was 60 mm, and 200° of data were collected with an

Table 2.4 Crystal parameters for the metal bound conantokin structures

	Ca^{2+}/con-G	Ca^{2+}/con-T[K7γ]	Cd^{2+}/Mg^{2+}/con-T[K7γ]
Crystal Form	Tetragonal	Cubic	Hexagonal
Space Group	$P4_222$	$P4_332$	$P6_322$
Unit Cell (a,b,c) (α, β, γ)	29.3, 29.3, 46.9 Å 90.0, 90.0, 90.0°	89.0, 89.0, 89.0 Å 90.0, 90.0, 90.0°	57.1, 57.1, 32.8 Å 90.0, 90.0, 120.0°
Solvent Content	46%	71.3%	43.8%
Mol. Per Asymu	1	2	1

oscillation of 1°. The crystal parameters of con-G are listed in Table 2.4. Assuming one molecule of con-G per asymmetric unit, the crystal volume per protein mass is 2.3, which corresponds to approximately 46% solvent. A synchrotron X-ray diffraction data set to a resolution of 1.2 Å, with an overall I/σ of 20.91, was obtained. The X-ray diffraction data was 97.9% complete with an R_{merge} of 7.9% for 12,427 unique reflections from a total of 266,169 measured reflections. Detailed data collection statistics are found in Table 2.5.

2.3.2 Structure Determination and Refinement

In order to determine the structure of Ca^{2+}/con-G molecular replacement using the program MolRep within CCP4[8] was utilized. Many different models were tried in order to obtain a correct solution. Models using known NMR structures of con-G were not helpful in molecular replacement attempts as well as polyalanine helices of many different lengths from previously determined crystal structures. The

Table 2.5 Data statistics for the metal bound conantokin structures[a]

	Ca^{2+}/con-G	Ca^{2+}/con-T[K7γ]	Cd^{2+}/Mg^{2+}/con-T[K7γ]
Wavelength, Å	1.0000	1.0000	1.0332
Resolution, Å	1.2 (1.29-1.24)	1.6 (1.66-1.60)	1.2 (1.29-1.27)
Completeness, %	97.9 (97.7)	99.5 (99.5)	84.2 (62.7)
I/σ	20.91 (1.31)	21.31 (1.88)	29.34 (2.70)
R_{merge}, %	7.9	9.8	6.7
Unique Reflections	12,427	16,539	10,347
Measured Reflections	266,169	215,476	132,076

[a]The parentheses denote the values for the last resolution shell

Table 2.6 Refinement statistics for the metal bound conantokin structures

	Ca^{2+}/con-G	Ca^{2+}/con-T[K7γ]	Cd^{2+}/Mg^{2+}/con-T[K7γ]
R_{work}, %	11.99	13.35	11.03
R_{free}, %	16.15	18.63	14.40
Resolution, Å	8-1.25	10-1.7	8-1.3
rmsd Bond Length, Å	0.01	0.01	0.01
rmsd Angle Length, °	0.02	0.03	0.02
Mean B-factors, $Å^2$	28.6	34.6	19.1

same was true for the programs AMoRe[7] and CaspR[9]. In order to determine the structure, data sets were collected near the Ca^{2+} edge (5000 kEV and 6000 kEV) in order to determine the positions of Ca^{2+} sites. However, the program SOLVE[16] failed to find any sites.

The isomorphous replacement method using natively grown crystals soaked with $BaCl_2$ was also done. Previous experiments had shown that con-G in the presence of Ba^{2+} formed a dimeric structure[5]. Ba^{2+} has more electrons than Ca^{2+} and the Ba^{2+} edge is easier to achieve than Ca^{2+}. Soaks were performed on natively grown Ca^{2+}/con-G crystals. Different concentrations as well as different time spans were tried. Crystals that showed no evidence of degradation were taken to the synchrotron. A number of data sets were collected and compared with the native for intensity differences and isomorphism. However, the program SOLVE[16] failed to find any Ba^{2+} sites.

The structure of Ca^{2+}/con-G was finally solved by molecular replacement using the recently developed program, Phaser[11], and a 14-mer polyalanine helix as a model. For con-G, brute force rotation and translation functions were performed, followed by refinement. The translation Z-score and LLG were 6.83 and 48.46, respectively indicating a correct solution. The LLG score after refinement improved to 63.82 further implicating that a correct solution had been obtained. Phaser[11] proved to be the only molecular replacement program capable of locating a correct solution, as many other programs such as AMoRe[7], MolRep within CCP4[8], and CaspR[9] failed. Examination of the crystal packing of the solution revealed that no

collisions occur with any of the symmetry-related molecules further supporting that the molecular replacement solution is correct.

A rigid body refinement was performed using Refmac5[19] within the CCP4 suite of programs[8,20], producing R_{work} 53.7% and R_{free} 54.6%. Side-chains were built into the structure from the electron density map. Subsequent refinements of all 17 residues of con-G led to an R_{work} of 35.9% and R_{free} 45.4%. In order to determine the positions of the Ca^{2+} ions, a F_o-F_c map was calculated and contoured at 8σ (Figure 2-14). This map showed that three calcium ions were present within the structure. After these ions were added, a refinement was performed where the R_{work} and R_{free} dropped to 27.16% and 32.29%, respectively. Arp/warp[19,21] was then run in order to add water molecules. Further cycles of refinement using the Refmac5[19] program in CCP4[8,20] produced a final R_{work} of 18.22% and R_{free} 21.37%. In order to finish refining the structure, the program SHELXL-97[22] was employed. Further refinement and additional water molecules were added, as well as adding side-chain secondary conformations of residues L5, Q9, and S16. SHELXL produced a final R_{work} and R_{free} of 11.99% and 16.15% respectively. The final model includes all 17 residues of con-G, 3 Ca^{2+}, and 68 water molecules. Electron density was also seen for the C-terminal NH_2 atom. The Ramachandran plot of the structure contained 15 non-glycine, non-proline residues with 100% in the most favored region (Figure 2-15). Refinement statistics are shown in Table 2.6. All model building was done using TURBO FRODO[17].

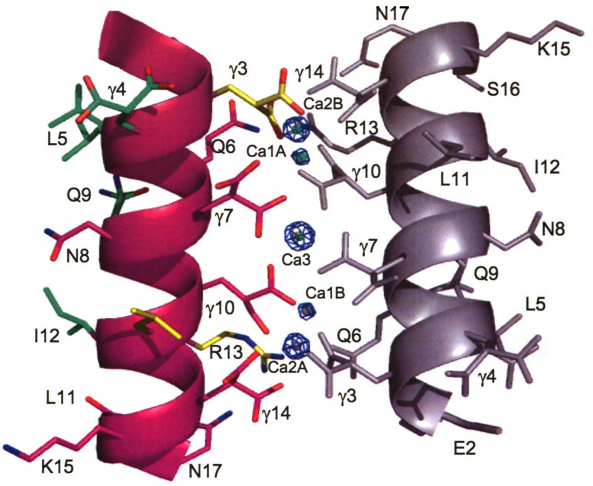

Figure 2-14. Overall structure of the con-G antiparallel dimer with Fo-Fc map calculated in the absence of calcium and contoured at 8σ. The gray molecule is related by crystallographic two-fold symmetry that forms the dimer. The side-chains colored green (E2, γ4, L5, Q9, I12) have been shown to decrease potency of the NMDA receptor when mutated to Ala by at least 100-fold. The side-chains colored yellow (G1, γ3, L11, R13) have been shown to decrease the potency of the NMDA receptor when mutated to Ala by at least 10-60-fold. Ca3 is positioned directly on the crystallographic two-fold axis. Ca1A and Ca2A are crystallographically related to Ca1B and Ca2B, respectively. The peptide side-chains are colored by atom type.

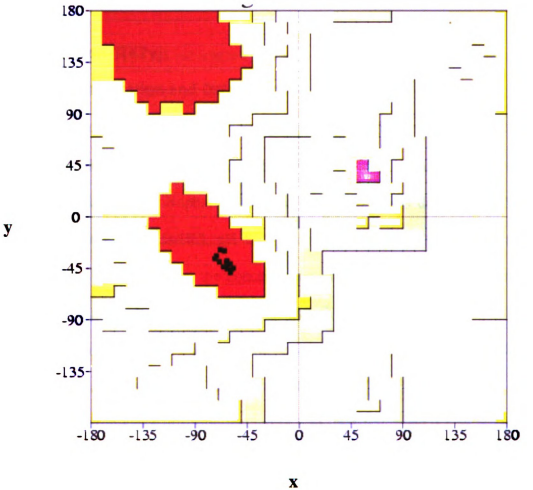

Figure 2-15 Ramachandran plot of the Ca^{2+}/con-G structure. Phi (Degrees) is x and Psi (Degrees) is y.

2.4 Ca^{2+}/con-T[K7γ]

2.4.1 Crystallization and Data Collection

The con-T[K7γ] peptide was synthesized, purified, and characterized as previously described[5] and provided by Frank Castellino and co-workers in a lyophilized form. The lyopholized solid of con-T[K7γ] was dissolved in 50 mM $CaCl_2$ and 100 mM Tris-HCl, pH 8, to a concentration of 10 mg/ml. Previous experiments had shown that the concentration of Ca^{2+} and pH were important in forming a dimeric structure of Ca^{2+}/con-T[K7γ][4]. The search for well diffracting crystals was performed using several crystallization screens at two different temperatures, 25°C and 4°C. The best crystals of con-T[K7γ] were grown at 4°C by the hanging drop vapor diffusion method in 3 M $(NH_4)_2SO_4$ and 0.1 M NaOAc, pH 5.5. Crystals first appeared after 1 year.

Crystals of con-T[K7γ] were briefly soaked in a cryoprotectant solution containing 4 M sodium dihydrogen phosphate, pH 5.5/ 20% glycerol at 4° C, and flash frozen by immersion in liquid N_2. The salt component was changed in order to cryoprotect the crystal since the glycerol concentration would not be high enough in the original crystal condition from above. Data were collected at the Advanced Photon Source COM-CAT 32-ID at the Argonne National Laboratory to a resolution of 1.6 Å, and data were processed and scaled using AUTOMAR in the cubic spacegroup $P4_332$ with unit cell parameters a = b = c = 89.0 Å. The crystal to detector distance was 100 mm and 40° of data were collected with an oscillation of 1°. The crystal parameters of con-T[K7γ] are listed in Table 2.4. Assuming two

molecules of con-T[K7γ] per asymmetric unit, the crystal volume per protein mass is 4.3, which corresponds to approximately 71.3% solvent in the crystal. A synchrotron X-ray diffraction data set to a resolution of 1.6 Å, with an overall I/σ of 21.31, was obtained. The X-ray diffraction data was 99.5% complete with an R_{merge} of 9.8% for 16,539 unique reflections from a total of 215,476 measured reflections. Detailed data statistics are found in Table 2.5.

2.4.2 Structure Determination and Refinement

The structure of Ca^{2+}/con-T[K7γ] was solved by molecular replacement using the program, Phaser[11], and a 14-mer polyalanine helix as a model. Automated searches using were performed in order to find the correct solutions. The translation Z-score and LLG for the first molecule were 8.6 and 57, respectively, and 19.1 and 216 for the second molecule indicating that two distinct correct solutions had been found. Examination of the crystal packing of the solution revealed that no collisions occur with any of the symmetry-related molecules, further supporting that the molecular replacement solution is correct.

A rigid body refinement was performed as with con-G, producing R_{work} 52.5% and R_{free} 56.5%. After building in side-chains and adding Ca^{2+} and water molecules, an R_{work} of 20.36% and R_{free} 23.08% was obtained. All model building was done using the program TURBO FRODO[17] and refinement and map calculations were carried out using CCP4[8]. In order to finalize the refinement of the structure, SHELXL-97[22] was employed, producing a final R_{work} and R_{free} of 13.35% and 18.63%, respectively. The final model includes all 21 residues of con-T[K7γ], 4 Ca^{2+}, and 142 water molecules (Figure 2-16). No alternative amino acid side-chain

confirmations are observed for any of the residues. The NH_2 moiety of the C-terminal amide was not seen in the electron density map of either con-T[K7γ] helix. As with con-G, the positions of the Ca^{2+} ions was determined by calculating a F_o-F_c map in the absence of calcium and contoured at 8σ (Figure 2-16).

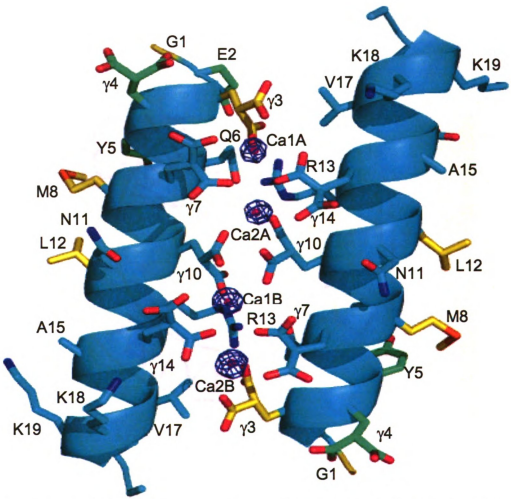

Figure 2-16. The overall structure of the antiparallel con-T[K7γ] dimer with a Fo-Fc map calculated in the absence of calcium and contoured at 8σ. The side-chains colored green (E2, γ4, Y5) have been shown to decrease potency of the NMDA receptor when mutated to Ala by at least 100-fold. The side-chains colored yellow (G1, γ3, M8, L12) have been shown to decrease the potency of the NMDA receptor when mutated to Ala by at least 10-60-fold. In all cases, the peptide side-chains are colored by atom type.

The Ramachandran plot of helix A of the structure contained 19 non-glycine, non-proline residues (94.7%) in the most favored region and 5.3% in additional allowed region (Figure 2-17). The Ramachandran plot of helix B of the structure contained 19 non-glycine, non-proline residues 100% in the most favored region (Figure 2-18). Refinement statistics are shown in Table 2.6.

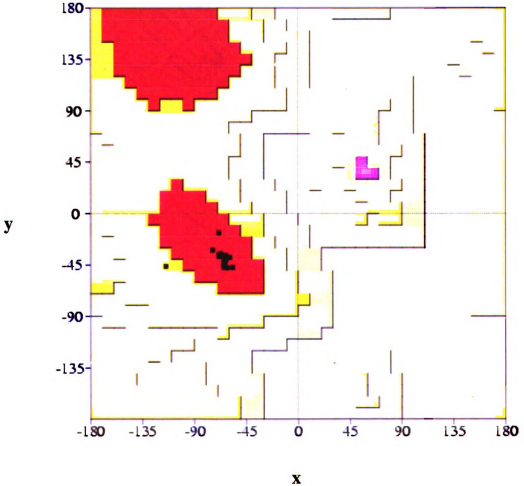

Figure 2-17 Ramachandran plot of helix A in the Ca^{2+}/con-T[K7γ] structure. Phi (Degrees) is x and Psi (Degrees) is y.

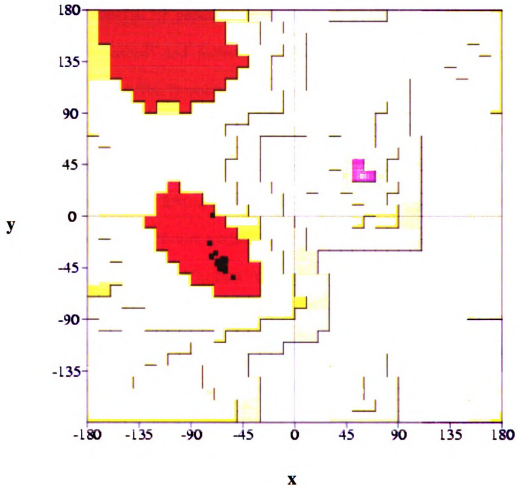

Figure 2-18 Ramachandran plot of helix B in the Ca^{2+}/con-T[K7γ] structure. Phi (Degrees) is x and Psi (Degrees) is y.

2.5 Cd^{2+}/Mg^{2+}/con-T[K7γ]

2.5.1 Crystallization and Data Collection

The con-T[K7γ] peptide was synthesized, purified, and characterized as previously described[5] and provided by Frank Castellino and co-workers in a lyophilized form. The lyopholized solid of con-T[K7γ] was dissolved in 50 mM $MgCl_2$ and 100 mM Tris-HCl pH 8 to a concentration of 10 mg/ml. The search for well diffracting crystals was performed using several crystallization screens at two different temperatures, 25°C and 4°C. The best crystals were grown at 25°C by the hanging drop vapor diffusion method in 1 M sodium acetate, 0.1 M Hepes pH 7.8, and 0.05 M $CdSO_4$ from Hampton Research Crystal Screen II #34. The crystals appeared after 3 days. The crystals were briefly soaked in a cryoprotectant solution containing 1 M sodium acetate, 0.1 M Hepes pH 7.8, 0.05 M $CdSO_4$, and 30% glycerol at 25°C and flash frozen by immersion in liquid nitrogen. X-ray diffraction data was collected at the Advanced Photon Source SBC 19-BM at Argonne National Laboratory to a resolution of 1.2 Å and data were processed and scaled using the HKL suite of programs to the hexagonal space group $P6_322$ with unit cell parameters a = b = 57.1 Å and c = 32.8 Å. The crystal to detector distance was 125 mm and 120° of data were collected with an oscillation of 1°. The crystal parameters are listed in Table 2.4. Assuming one molecule of con-T[K7γ] per asymmetric unit, the crystal volume per protein mass is 2.2, which corresponds to approximately 43.8% solvent in the crystal. A synchrotron X-ray diffraction data set to a resolution of 1.2 Å, with an overall I/σ of 29.34, was obtained. The X-ray diffraction data was 84.2%

complete with an R_{merge} of 6.7% for 10,347 unique reflections from a total of 132,076 measured reflections. Detailed data statistics are found in Table 2.5.

2.5.2 Structure Determination and Refinement

The structure was solved by molecular replacement using the program Phaser[11] and the Ca^{2+}/con-T[K7γ] (PDB id 2DPR) as a model. The rotational and translational Z-scores were 9.4 and 10.4, and the LLG was 150 indicating a correct solution. Examination of the crystal packing of the solution revealed that no collisions occur with any of the symmetry-related molecules further supporting that the molecular replacement solution is correct.

A rigid body refinement was performed using Refmac5[19] within the CCP4[8] suite of programs producing R_{work} and R_{free} of 53.2%. An anomalous difference map was calculated using CCP4[8] in order to determine which sites were Cd^{2+} and which sites were Mg^{2+} since Cd^{2+} existed in the crystallization condition (Figure 2-19). A total of four Cd^{2+} sites were seen in the electron density of the anomalous difference map at peaks of 24σ (Cd3), 14σ (Cd1), and 11σ (Cd2 and Cd4). The program SOLVE[16] was used to determine actual Cd^{2+} site occupancy. Only Cd3 is 100% occupied whereas Cd1 is 80% occupied and Cd2 and Cd4 are both 50% occupied. We were unable to determine whether these partially occupied Cd^{2+} sites were also partially occupied by magnesium since magnesium does not have an anomalous signal. One Mg^{2+} site was located from the F_o-F_c map contoured at 8σ (Figure 2-20). Figure 2-21 also shows electron density for the Cd^{2+} sites. Subsequent refinements using Refmac5[19] within CCP4[20] of all 21 residues of con-T[K7γ] after adding 4 Cd^{2+}, 1 Mg^{2+}, and 28 waters produced an R_{work} of 18.61%

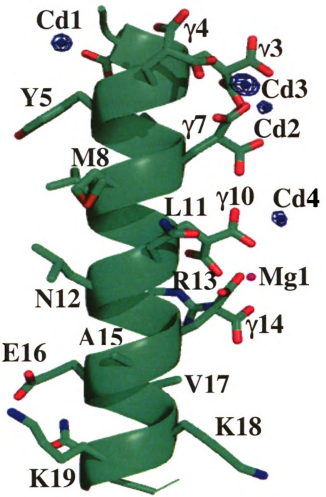

Figure 2-19 The overall structure of Cd^{2+}/Mg^{2+}/con-T[K7γ] with an anomalous difference map contoured at 8σ. The four Cd^{2+} are shown as blue spheres and the Mg^{2+} is shown as a magenta sphere. The side-chains are colored by atom.

and R_{free} 20.47%. In order to finish refining the structure, the program SHELXL-97[22] was employed. Further refinement was done by adding secondary side-chain conformations of Met8 and Val17 and adding more waters. SHELXL produced a final R_{work} and R_{free} of 11.03% and 14.40% respectively. The final model includes all 21 residues of con-T[K7γ], 4 Cd^{2+}, 1 Mg^{2+} and 61 water molecules. The NH_2 moiety of the C-terminal amide was seen in the electron density map of the con-T[K7γ] helix. The Ramachandran plot of the structure contained 19 non-glycine, non-proline residues (100%) in the most favored region (Figure 2-21). Refinement statistics are shown in Table 2.6.

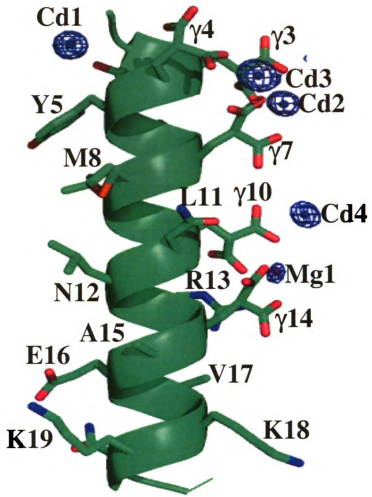

Figure 2-20 The overall structure of Cd^{2+}/Mg^{2+}/con-T[K7γ] with a F_o-F_c map calculated in the absence of metal cations and contoured at 5σ. The four Cd^{2+} are shown as blue spheres and the Mg^{2+} is shown as a magenta sphere. The side-chains are colored by atom.

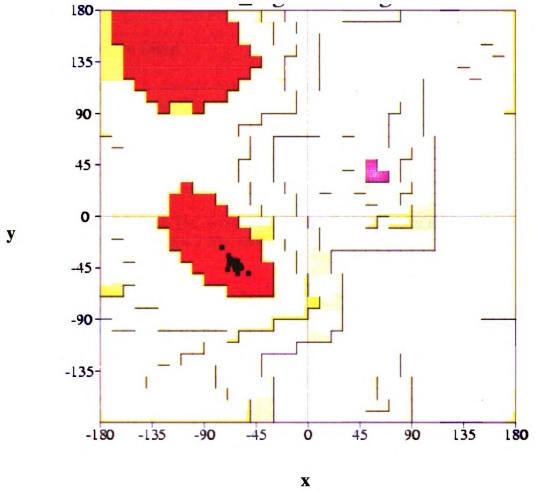

Figure 2-21 Ramachandran plot of the Cd^{2+}/Mg^{2+}/con-T[K7γ].
Phi (Degrees) is x and Psi (Degrees) is y.

2.6 Literature Cited

1. McPherson, A. *Introduction to Macromolecular Crystallography*, John Wiley & Sons Inc (Hoboken, NJ, 2003).

2. Hammes, G.G. *Spectroscopy for the Biological Sciences*, John Wiley & Sons Inc (Hoboken, NJ, 2005).

3. Abad, M.C. et al. The X-ray crystallographic structure of the angiogenesis inhibitor angiostatin. *J Mol Biol* **318**, 1009-17 (2002).

4. Dai, Q., Castellino, F.J. & Prorok, M. A single amino acid replacement results in the Ca2+-induced self-assembly of a helical conantokin-based peptide. *Biochemistry* **43**, 13225-32 (2004).

5. Dai, Q., Prorok, M. & Castellino, F.J. A new mechanism for metal ion-assisted interchain helix assembly in a naturally occurring peptide mediated by optimally spaced gamma-carboxyglutamic acid residues. *J Mol Biol* **336**, 731-44 (2004).

6. Stout, G.H. & Jensen, L.H. *X-ray Structure Determination*, John Wiley & Sons Inc (New York, N. Y., 1989).

7. Navaza, J. AMoRe, an automated program for molecular replacement. *Acta Crystallog. sect. A* **50**, 157-163 (1994).

8. Collaborative Computational Project, N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst D50*, 760-763 (1994).

9. Claude, J.-B., Suhre, K., Notredame, C., Claverie, J.-M. & Abergel, C. CaspR: a web-server for automated molecular replacement using homology modelling. *Nucleic Acids Research* **32**, W606-W609 (2004).

10. Brunger, A.T. *X-PLOR, version 3.1, a system for X-ray Crystallography and NMR*, (New Haven, CT, 1992).

11. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. & Read, R.J. Likelihood-enhanced fast translation functions. *Acta Cryst* **D61**, 458-64 (2005).

12. Shepard, S.R., Boucher, R., Johnston, J., Boerner, R., Koch, G., and Madsen, J. Large-scale purification of recombinant human angiostatin. *Protein Expr Purif* **20**, 216-227 (2000).

13. Rios-Steiner, J.L., Schenone, M., Mochalkin, I., Tulinsky, A., and Castellino, F. J. Structure and binding determinants of the recombinant kringle-2 domain of human plasminogen to an internal peptide from a group A Streptococcal surface protein. *J Mol Biol* **308**, 705-19 (2001).

14. Otwinowski, Z. *Oscillation data reduction program.*, 56-62 (SERC Daresbury Laboratory, Daresbury, UK, 1993).

15. Otwinowski, Z.a.M., W. Processing of x-ray diffraction data collected in oscillation mode. *Methods Enzymol* **276**, 307-326 (1997).

16. Terwilliger, T.C. & Berendzen, J. Automated MAD and MIR structure solution. *Acta Crystallographica D55*, 849-861 (1999).

17. Jones, T.A., Zou, J.Y., Cowan, S.W. & Kjeldgaard, M. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Cryst* **A47**, 110-119 (1991).

18. Brunger, A.T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., and Grosse-Kuntsleve, R. W. Crystallography & NMR System: a new software suite for macromolecular structure determination. *Acta Crystallog. sect. D* **54**, 905-921 (1998).

19. Murshudov, G.N., Vagin, A.A. & Dodson, E.J. Refinement of Macromolecular Structures by the Maximum-Likelihood Method. *Acta Cryst* **D53**, 240-55 (1997).

20. COLLABORATIVE COMPUTATIONAL PROJECT, N. The CCP4 Suite: Programs for Protein Crystallography. *Acta Cryst* **D50**, 760-3 (1994).

21. Perrakis, A., Morris, R.J.H. & Lamzin, V.S. Automated protein model building combined with iterative structure refinement. *Nature Struct Biol*, 458-463 (1999).

22. Sheldrick, G.M. & Schneider, T.R. SHELXL: High resolution refinement. in *Methods in Enzymology*, Vol. 277 (eds. Sweet, R.M. & Carter Jr, C.W.) 319-43 (Academic Press, Orlando, Florida, 1997).

Chapter III

The Three Dimensional Structure of Angiostatin/VEK-30

The angiostatin/VEK-30 complex structure has been solved and further described in this chapter (Cnudde SE, *et al.* 2006. *Biochemistry.* 45(37): 11052-60). The methods used in order to determine this structure have been thoroughly described in Chapter II. The complex of angiostatin/VEK-30 was crystallized and its structure determined in two different space groups, $P6_122$ and $P6_1$. The $P6_1$ structure shows the interkringle K2-K3 disulfide and two residues on either side of the disulfide region in K3. This and other comparisons discussed herein prove that plasminogen kringle domains are capable of significant structural rearrangement relative to one another. This provides insight into the mechanism of plasminogen during streptococcal infection and further characterizes plasminogen's importance as a critical serine protease for bacteria. Also, VEK-30 binds specifically to angiostatin K2 through a pseudo-lysine moiety. This pseudo-lysine motif seen in VEK-30 may help to identify unknown angiostatin ligands.

3.1 Overall Structure of Angiostatin/VEK-30 Complex

The resolved portion of the overall structure of the angiostatin/VEK-30 complex is shown in Figure 2-4. In comparison to the unbound form of angiostatin (Figure 1-6 shows the overall structure of angiostatin), no electron density was detected for any of the residues of Pg-K3 (residues C256-C333) in the angiostatin/VEK-30 complex. Additionally, electron density is absent for most of the residues in the K2-K3 interkringle peptide (residues P246-Q255). However, the remaining residues of the Pg-K1 and Pg-K2 domains are well ordered and exist in a relatively extended orientation

with essentially no interactions between them. There are also very few interactions between either Pg-K1 or Pg-K2 and the linker peptide that connects the two domains. Electron density for the VEK-30 peptide is absent for N-terminal residues Val85-Lys87 and C-terminal residues K111-Y114 (Figure 1-4 shows the primary sequence of VEK-30). However, of the remaining 23 residues, 20 correspond to about 5 turns of a well-defined α-helix that is approximately 30 Å long (Figure 2-4).

3.2 Interactions Between VEK-30 and Angiostatin

Figure 3-1 and 3-2 shows the interactions occurring between VEK-30 and angiostatin. Table 3.1 lists distances involved in the interaction between VEK-30 and angiostatin K2. VEK-30 interacts with K2 of angiostatin from VEK-30 residues Glu93 to Glu104. This region encompasses most of the first direct repeat and the first residue of the second (Figure 1-4 shows the VEK-30 primary sequence). Most of these interactions occur between a single face of the helix and the K2 motif. One side of the helix consisting of residues K98-E104 makes numerous contacts with the angiostatin K2 LBS. Hydrogen bond and salt bridge electrostatic interactions occur between angiostatin K2 and VEK-30, both within and outside of the K2 LBS. There are no interactions between angiostatin K1 and VEK-30. The total number of contacts between angiostatin K2 and VEK-30 with a distance less than 3.8 Å is about 61. The LBS of the K2 domain contains consensus anionic (Asp219 and Glu221) and cationic (Arg234) centers (Figure 3-2 and 3-3). VEK-30 forms a pseudo-lysine by inserting the residues Arg101 and Glu104 located on one face of the α-helix into the LBS of angiostatin K2. The principal interactions that occur between the anionic loci of the K2 LBS and VEK-30 involve residues Asp219 and Glu221 of K2 and Lys98, Arg101, and His102 of VEK-30.

Critical salt bridge interactions occur between angiostatin K2 LBS residue Asp219 and VEK-30 residues Arg101 and Lys98. In addition, Glu221 makes a tight (distance 2.57 Å) salt bridge contact with VEK-30 residue Arg101, and also makes contacts with His102 of VEK-30. Figure 3-2 shows a π-cation interaction. This interaction occurs between the guanidino group of Arg101 of VEK30 and one of the two K2 Trp residues (Trp235) that make up the hydrophobic portion of the LBS (distances shown in Table 3.1). At the cationic site, a salt bridge exists between residues Arg234 of K2 and Glu104 of VEK-30. An interaction between residues Arg220 of angiostatin and Asp91 of VEK-30 is also observed. VEK-30 Arg101 has proved to be vital for angiostatin K2 binding, since mutating this residue to Ala results in no measurable affinity for K2[2]. Mutating VEK-30 residues Lys98 and His102 to Ala and Glu104 to Gln results in a decreased affinity for the angiostatin K2 domain. These results suggest that the overall binding relies on interactions mediated by VEK-30 residues Lys98, Arg101, His102 and are consistent with our crystallographic results. There are also significant interactions between VEK-30 and K2 outside the LBS. A salt bridge interaction occurs between the angiostatin K2 residue Lys204 and VEK-30 residue Glu93 (Figure 3-1). This interaction was also first observed in the structure of the isolated K2/VEK-30 (Pg-K2) complex[2]. It is unknown whether this exosite interaction is critical for binding. The exosite salt bridge interaction between angiostatin K2 residue Lys204 and VEK-30 residue Glu93 is further stabilized by hydrogen bonds and hydrophobic contacts of VEK-30 residues Leu94, Leu97, and Lys98 and angiostatin K2 residues Tyr200, Phe205, Asp219, and Arg220.

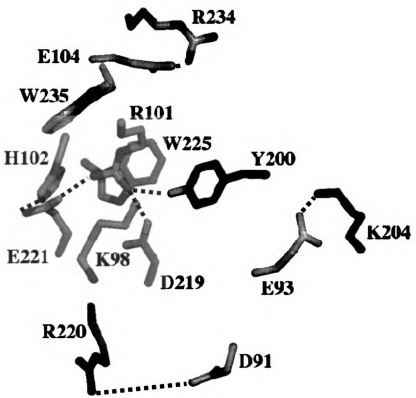

Figure 3-1 Interactions between the angiostatin K2 LBS and VEK-30. R101 and E104 are spaced by almost one helical turn and form the pseudo-lysine residue. Angiostatin is colored green and VEK-30 is shown in magenta. K2 residues are labeled using plasminogen numbering and VEK-30 residues are labeled with PAM numbering. All atoms are colored by atom type (nitrogen, blue; oxygen, red).

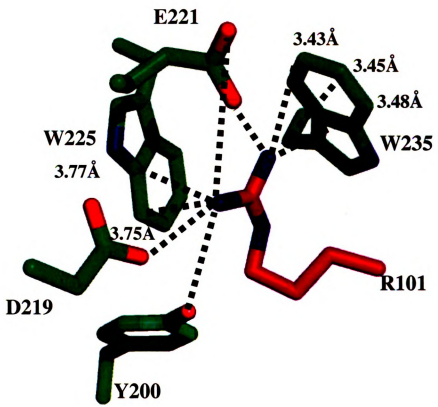

Figure 3-2 A π-cation interaction occurs between the guanidino group of VEK-30 R101 and angiostatin K2 LBS residue W235. Angiostatin is shown in green and VEK-30 is shown in magenta. All atoms are colored by atom type (nitrogen, blue; oxygen, red). Angiostatin residues are labeled with plasminogen numbering. VEK-30 residues are labeled with PAM numbering.

Most of the interactions discussed above are also seen in the structure of a mutated Plg K2 domain (C169G, E221D, L237Y) bound to VEK-30[2]. However, two of the mutated residues, D221 and Y237, made interactions with VEK-30 in the LBS, calling into question whether the wild-type LBS would provide a similar interface. Our structure confirms that in spite of these mutations, most of the interface is similar in the two structures.

3.3 Kringle Domain Rotation

Comparison of the structures of unbound angiostatin to the VEK-30 bound form reveals that K1 has rotated significantly from its unbound conformation when the K2 regions are overlaid (Figure 3-3). K1 of the angiostatin/VEK-30 complex rotates 48.1° and translates about 0.5 Å from its position in the unbound angiostatin crystal structure (relative motion determined using the program DynDom[1]). Residues encompassing the angiostatin K1-K2 linker peptide are the bending residues primarily responsible for such a large rotation. More specifically, the changes in psi- and phi-dihedral angles of residues Glu163 and Glu165 contribute significantly to most of the interdomain rotation and are likely responsible for the large rotation of the angiostatin K1 domain seen in the angiostatin/VEK-30 structure (Figure 3-4). Cys166 also contributes to the rotation as well. VEK-30 binding does not appear to be directly responsible for the rotation of K1 since there are no interactions between K1 and VEK-30. However, crystal packing dictates the K1 position in each structure indicating relative motion in solution. Angiostatin K1 clashes into a symmetry-related molecule of angiostatin/VEK-30 when both K2 regions are overlaid. The same phenomenon is seen

Table 3.1 Interactions between angiostatin and VEK-30

Angiostatin	VEK-30	$P6_122$ (Å)
Y200-OH	L97-O	3.36
Y200-OH	K98-N	3.19
Y200-OH	R101-NH1	3.20
K204-NZ	E93-OE1	3.61
K204-NZ	E93-OE2	3.48
D219-OD1	R101-NH1	3.69
D219-OD2	R101-NH1	2.58
D219-O	K98-NZ	2.66
R220-NH1	D91-OD2	2.97
E221-OE1	R101-NH2	2.57
E221-OE1	H102-NE2	3.48
W225-NE1	R101-NH1	3.67
W225-CE2	R101-NH1	3.77
W225-CH2	R101-NE	3.80
W225-CH2	R101-CD	3.80
W225-CZ2	R101-NH1	3.75
W235-CD2	R101-NH2	3.43
W235-CE3	R101-NH2	3.41
W235-CZ3	R101-NH2	3.43
W235-CH2	R101-NH2	3.45
W235-CZ2	R101-NH2	3.48
R234-NH1	E104-OE1	3.15
W235-NE1	R101-O	3.15

when K2 of the angiostatin/VEK-30 complex is overlaid onto angiostatin K2. This motion demonstrates that the kringle domains of angiostatin are not rigidly positioned as previously thought but are in fact mobile relative to each other. The LBS of Pg-K1 and Pg-K4 are known to play important roles in the maintenance of the closed conformation of Pg[3] although the involvement of the NTD with these binding sites is unclear. The fact that kringle domains of Pg are capable of significant structural rearrangement relative to one another indicates that transitions between the open and closed Pg conformations may also involve significant motion of K1-K3.

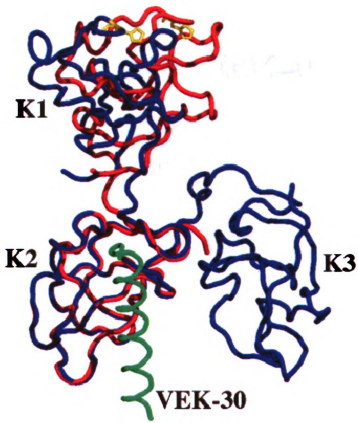

Figure 3-3 Overlay of the K2 domains of angiostatin and angiostatin/VEK-30. The structure of angiostatin is shown in blue and the angiostatin/VEK-30 complex is shown in red (angiostatin) and green (VEK-30). H114 of angiostatin is colored yellow in both structures.

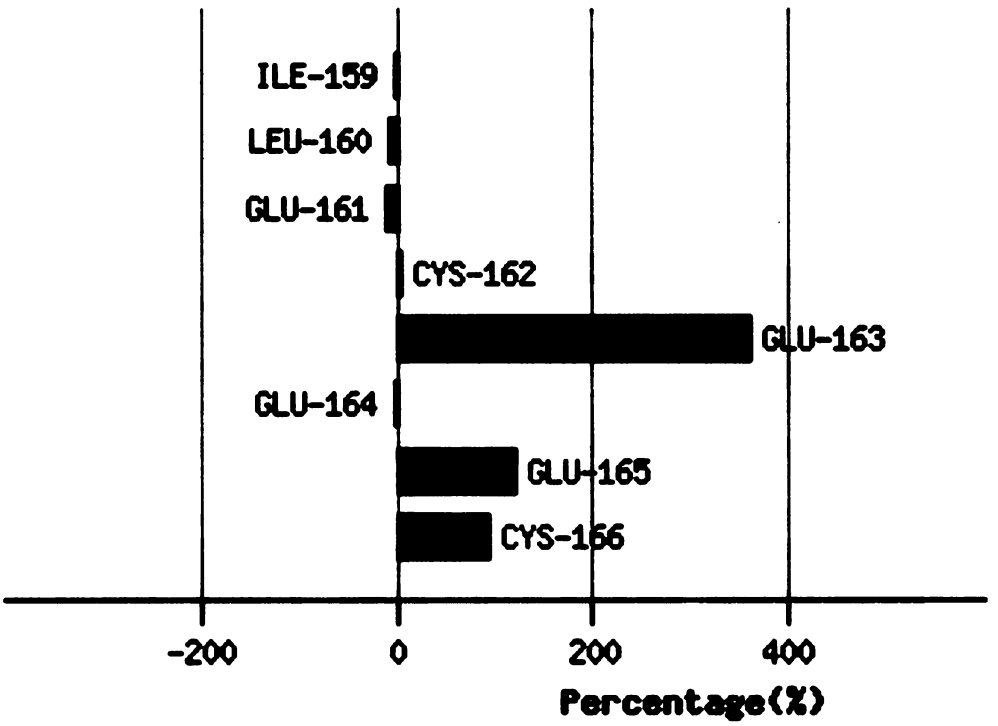

Figure 3-4 Residues responsible for interdomain rotation as determined by the program DynDom[1].

3.4 The Kringle 3 Domain

A second, $P6_1$ crystal form of the angiostatin/VEK-30 complex was produced by soaking crystals grown in the first, $P6_122$, form with Pt(II)-derivatizing molecules. This second crystal form preserves most of the crystal packing of the $P6_122$ form, although it loses the 2-fold axis perpendicular to the six-fold axis. This results in a crystal form with two molecules in the asymmetric unit, instead of one. Though the structure of the complex is quite similar, several important differences are seen between the $P6_1$ and $P6_122$ crystal forms of the angiostatin/VEK-30 complex. In one molecule of the $P6_1$ structure, the inter-kringle disulfide bond is ordered. Residues Pro296, Cys297, and Lys298 of K3 are seen in the electron density though no other residues of K3 were identified (Figure 2-8). This shows that the K2-K3 interkringle disulfide bond remains intact in these structures, even though K3 appears to be disordered in the structures. Overlaying K2 of angiostatin with K2 of the angiostatin/VEK-30 complex in the $P6_1$ structure displays a significant motion of about 3.3 Å at the inter-kringle disulfide bond (Figure 3-5). When residues Pro296-Lys298 of the VEK-30-bound form are overlaid with the corresponding residues in the crystal structure of free angiostatin, it can be seen that angiostatin K3 encroaches upon a crystallographic symmetry related molecule of K2 (Figure 3-6). This indicates that there may be some structural differences within the angiostatin K3 domain to avoid steric clashes with a crystallographically related molecule. It should be noted that the conformation of the tripeptide alone fits in either structure. The motion of angiostatin K3 is limited due to the inter-kringle disulfide bond

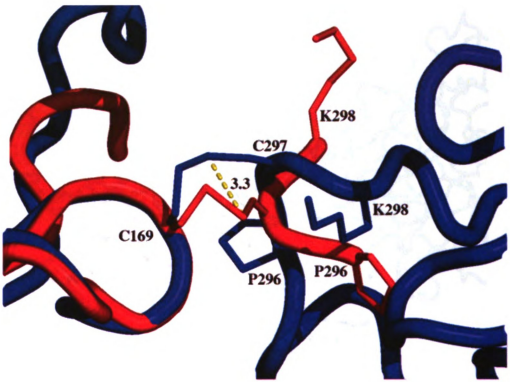

Figure 3-5 Overlay of the K2 of angiostatin with K2 of the angiostatin/VEK-30 complex from the $P6_1$ crystal form where 3 residues of angiostatin K3 are ordered. The $P6_1$ structure is shown in red and the structure of angiostatin is shown in blue. The residues are labeled with Pg numbering. The inter-kringle disulfide bond in the $P6_1$ complex structure has rotated 3.3Å from its position in the angiostatin structure.

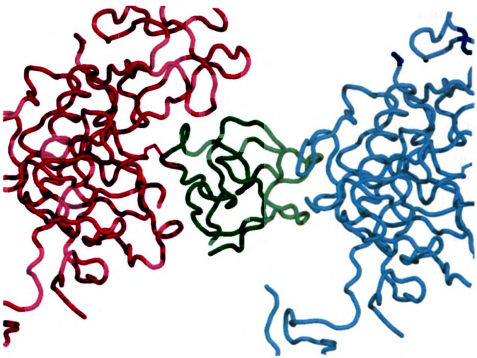

Figure 3-6 An overlay of residues Pro296-Lys298 of the VEK-30-bound with the corresponding residues in the crystal structure of free angiostatin. Angiostatin in the angiostatin/VEK-30 complex is colored magenta, K3 of angiostatin is shown in green, and the angiostatin/VEK-30 symmetry related molecule is shown in cyan.

between residues Cys169 of K2 and Cys297 of K3 (Figure 1-6). It has been suggested that the disruption of the K2-K3 inter-kringle disulfide bond is required for maximum angiogenic inhibition[4]. However, the angiostatin double mutant (C169S, C297S), which eliminates the inter-kringle disulfide bond, has little effect on angiogenic activity, but resulted in the loss of EACA binding by K2 leading to the supposition that lysine binding by K2 was unimportant for anti-angiogenic activity[5]. However, this loss of EACA binding is not in agreement with the binding of a series of α,ω –amino acids and VEK-30 to the C169G mutant of K2,[6] and the observation that the C169D/C297R double mutant retains the chloride- and EACA-induced hydrodynamic properties of wild-type plasminogen[3]. Similar conclusions regarding the irrelevance of lysine binding to angiostatin were drawn from comparisons of lysine-binding affinity and anti-angiogenic potency[7].

3.5 Dimerization

Inspection of the crystal packing indicated that dimerization between two angiostatin/VEK-30 complex molecules occurs along a crystallographic 2-fold axis (Figure 3-7). The same dimerization was seen in the K2/VEK-30 structure between the two molecules of K2/VEK-30 in the asymmetric unit[2] (Figure 3-8). This indicates that dimerization is not an artifact of crystal packing since there is obviously no relationship between the crystal packing in the K2/VEK-30 complex and the angiostatin/VEK-30 complex. Figure 3-9 shows that the Cα positions of the dimeric structure of K2/VEK-30 superimpose well with the two angiostatin/VEK-30 complexes that are related by crystallographic 2-fold symmetry (rmsd for all atoms ~0.4 Å). Dimerization results in

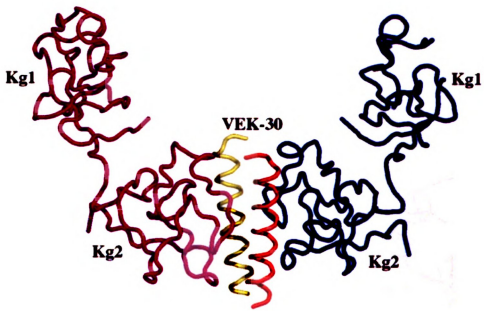

Figure 3-7 Dimerization of the angiostatin/VEK-30 complex. One molecule has angiostatin colored blue and VEK-30 colored red, while the other molecule has angiostatin colored magenta and VEK-30 colored yellow.

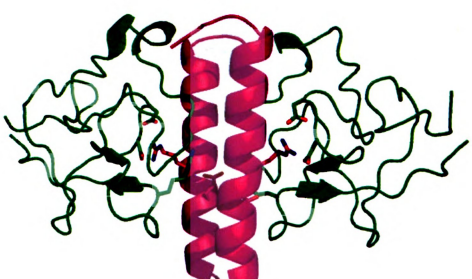

Figure 3-8 Dimerization of the K2/VEK-30 structure[2]. K2 is shown in green and VEK-30 is colored magenta.

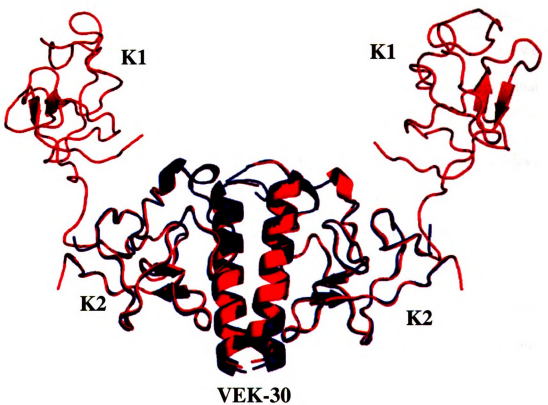

Figure 3-9 Overlay of the K2/VEK-30 dimer onto the angiostatin/VEK-30 dimer. The angiostatin/VEK-30 dimer is shown in red and the K2/VEK-30 dimer is shown in blue. The Cα positions of the dimeric structure of K2/VEK-30 superimpose well with the two angiostatin/VEK-30 complexes that are related by crystallographic 2-fold symmetry (rmsd for all atoms ~0.4 Å).

two molecules of α–helical VEK-30 packing parallel and side-by-side in the center of the dimer, and two K2 domains located on either side of the parallel helices. Interestingly, full-length PAM is predicted to homo-dimerize as a coiled-coil that extends on either side of the VEK-30 region[8]. In fact, several homology modeling programs predict the PAM structure based on the extended coiled-coil structure of tropomyosin[8,9]. The only region of PAM (between amino acids 60-310) that is not well-fitted by the coiled-coil structural prediction is the Pg-binding direct repeat region that encompasses the VEK-30 peptide (W. Wedemeyer, unpublished results). This is consistent with the structure of the angiostatin/VEK-30 dimer, in that the two helices do not form a classical coiled coil, though they are parallel and stacked side-to-side. Numerous contacts between the two molecules at the dimerization interface are observed. As shown in Figure 3-10, water-mediated interactions occur at the dimerization interface between angiostatin K2 residue Gln193 of the symmetry-related molecule and VEK-30 Asn99. Another water-mediated interaction occurs at angiostatin K204 and VEK-30 Glu94. Numerous hydrogen bonds also play a role at the dimerization interface. Hydrogen bonds occur between VEK-30 residue Glu103 and angiostatin K2 symmetry-related molecule residues His196 and Ala197. Another hydrogen bond occurs between VEK-30 residue E100 and angiostatin symmetry related molecule Arg234, the cationic site within the K2 LBS. The calculated total buried surface area of a dimer of angiostatin/VEK-30 is 1627 $Å^2$ suggesting a relatively strong interaction. Furthermore, because a dimeric structure is exhibited by both K2/VEK-30 and angiostatin/VEK-30, it is plausible that such higher order structures may exist in

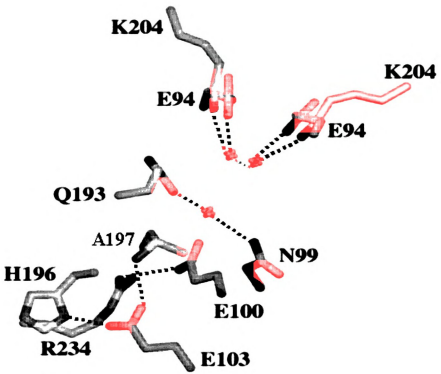

Figure 3-10 Residues involved in dimerization. VEK-30 residues are shown in green and angiostatin K2 residues are shown in yellow. The symmetry-related molecule of angiostatin K2 is shown in cyan with its VEK-30 residue shown in magenta. The atoms are colored by type. The water molecules are designated as W1 and the symmetry-related waters W2a and W2b. All residues are labeled with Pg and PAM numbering.

solution. To address this possibility, sedimentation equilibrium analysis was conducted on angiostatin in the absence and presence of VEK-30 by our collaborators Frank Castellino and co-workers. The apparent molecular weight of angiostatin (at a concentration of 13 μM) was determined to be 27,700±100 (calculated sequence-based weight = 29,000). In the presence of a 5-fold molar excess of VEK-30, (allowing for virtually all angiostatin to exist in VEK-30-bound form assuming a K_d of 460 nM for the angiostatin/VEK-30 interaction[6]), an apparent molecular weight of 28,400±800 was obtained. These data fail to support a model of VEK-30-mediated dimerization of angiostatin at angiostatin concentrations that are physiologically feasible based on circulating plasma levels (ca., 2 μM) of the Pg parent[10]. However, this does not rule out the possibility of a PAM-induced angiostatin (or Pg) dimer on the bacterial surface, where high effective concentrations of both binding partners can be encountered and where full-length PAM is strongly predicted to exist as a parallel coiled-coil dimer, as are all proteins in the M protein family[8,9].

3.6 Kringle 2 Domain Specificity

The Pg kringles have high sequence and structural homology. However, the Pg kringles are very different with respect to affinity for C-terminal lysine mimics and in their biological functions in proliferation and migration assays. A closer examination of the residues of K2 involved in VEK-30 interactions reveals that many of these residues are not conserved in Pg kringle domains. Figure 3-11 shows a sequence alignment for all five Pg kringles. For instance Gln193 is responsible for mediating interactions at the angiostatin/VEK-30 dimerization interface. In K1 and K4, this sequence position is occupied by Thr and Met, respectively. Ala197 of K2 is not conserved in Pg kringle

domains and also plays a role at the dimerization interface (K1, K4, and K5 all have Arg residues while K3 has a Thr). K2 residue Lys204 is also not conserved and is involved in mediating interactions at the dimerization interface, as well as directly interacting with VEK-30 through an exosite region. Focusing on residues within the LBS, Y200 is not conserved, where K1 and K5 have Phe and K3 and K4 have Arg and Lys, respectively. Tyr200 of K2 makes numerous interactions with VEK-30. Specifically, the hydroxyl group of Tyr200 makes a hydrogen bond with Arg101-NH1. Finally, Arg220 makes contacts with VEK-30, and this site is also not conserved in Pg kringle domains. Specifically, Arg220 forms a salt bridge with Asp219 in the angiostatin structure (Figure 1-9). Because of the interaction with Arg220, Asp219 is flipped out of the LBS and is incapable of interacting with the C-terminal group of EACA, possibly explaining the poor EACA binding affinity of K2[7,11-14]. However, VEK-30 abrogates the salt bridge interaction between Arg220 and Asp219 so that Asp219 flips into the LBS. This rearrangement recapitulates the canonical LBS architecture, permitting extensive LBS/VEK-30 interactions (Figure 3-12). Mouse Pg has substantially lower affinity for PAM relative to human Pg, although the two domains are 86% identical in sequence. Arg220 is the only residue that both interacts, either directly or indirectly, with VEK-30 and is not identically conserved in mouse Pg. This strongly implicates Arg220 as a residue that is both important for binding and critical to the species-specificity of PAM. In summary, many of the K2 residues that interact with VEK-30 are not conserved among Pg-K1, -K3, -K4 and -K5, likely explaining why only human Pg-K2 has affinity for VEK-30. The sequence analysis results identify potential targets for further mutagenesis studies. For example, it would be interesting to learn whether

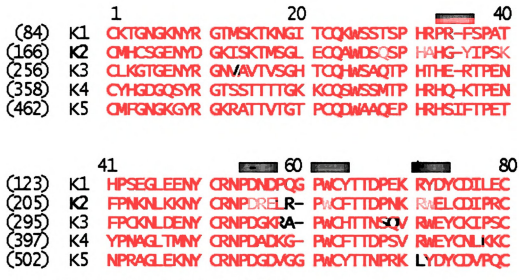

Figure 3-11 A sequence alignment of all five Pg kringle domains. Yellow highlighted residues are conserved.

mutating K1 residue N138 to R (same R220 in K2) and determine if this mutated K1 binds VEK-30 or mutate the mouse K2 to contain R220 to see if it now binds VEK-30. These studies would help characterize kringle domain specificity and identify why only plasminogen K2 binds VEK-30. These mutagenesis studies may also help to identify

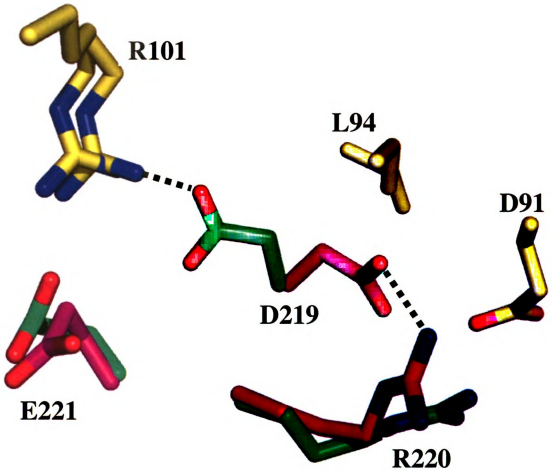

Figure 3-12 An overlay of angiostatin K2 onto angiostatin/VEK-30 K2. Residues from angiostatin alone are colored magenta, residues from the Angiostatin/VEK-30 complex are colored green, and VEK-30 residues are colored yellow. The atoms are colored by type. All residues are labeled with Pg and PAM numbering. D219 is flipped out of the LBS, resulting in a tight salt bridge contact with R220.

other potential kringle domain ligands. Ligands for plasminogen K1 and K3 have not been identified.

3.7 Further Studies of VEK-30

As previously mentioned in Chapter I, a region of PAM, spanning amino acids 91 - 116, contains two direct repeat sequences and is responsible for Pg binding by PAM. VEK-30, an α-helical peptide derived from residues 85-113 of PAM and containing the first and most of the second direct repeat, possesses a high affinity binding site for Pg-K2 (K_d = 460 nM),[2,15-17] even though it does not contain a C-terminal lysine residue. This led us to explore whether a peptide of PAM containing two full direct repeats (VEK-32L) might bind two molecules of K2 since it appears that PAM may have two distinct Pg binding regions from the analysis of the PAM sequence. Figure 3-13 shows both primary sequences of VEK-30 and the wild-type sequence further named VEK-32L. A model was made showing that binding of two K2 domains to a longer peptide of PAM was possible (Figure 3-14). However, Figure 3-15 shows that two angiostatin K2 binding to a longer peptide of PAM would disrupt the dimer seen in the angiostatin/VEK-30 and mK2/VEK-30 crystal structures. Gel filtration and sedimentation velocity experiments using analytical ultracentrifugation were performed to determine whether this postulate is true.

Figure 3-13 The primary sequences of VEK-30 (top) and VEK-32L (bottom). They are labeled with PAM numbering above. The green box indicates the direct repeat sequence.

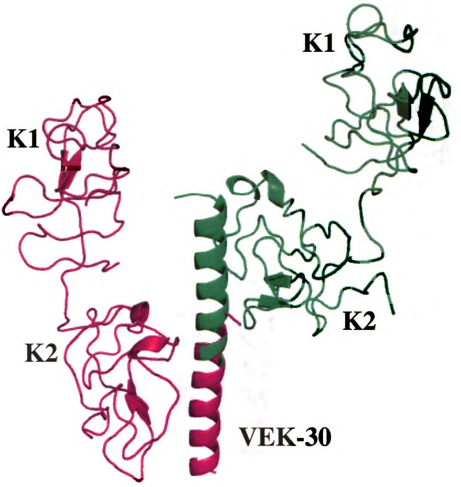

Figure 3-14 A model showing that two angiostatin K2s can bind to a longer PAM peptide. The figure was made by overlaying the direct repeat sequences from the angiostatin/VEK-30 structure. The angiostatin/VEK-30 structure is colored green and the second molecule magenta.

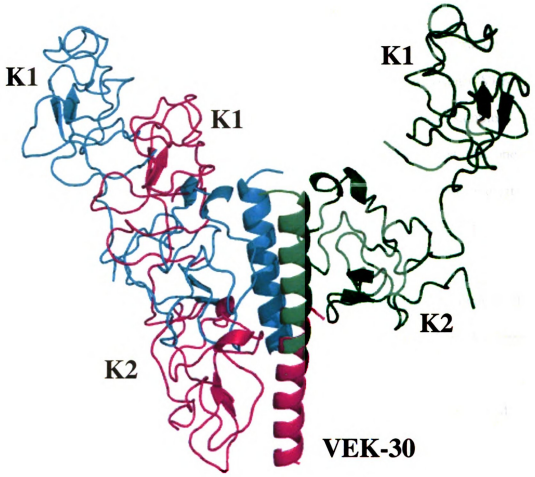

Figure 3-15 A model showing that two angiostatin K2s can bind to a longer PAM peptide but disrupts the dimer. The figure was made by overlaying the direct repeat sequences from the angiostatin/VEK-30 structure. The angiostatin/VEK-30 structure is colored green and the dimer colored cyan whereas the second molecule magenta.

Gel filtration was first performed to determine whether two molecules of angiostatin interact with a PAM peptide containing both direct repeats. Gel filtration analysis was performed on unbound angiostatin (~29 kDa) and on the angiostatin/VEK-32L (~33 kDa) complex using a Superdex 200 10/300 GL column. Different sample volumes as well as varying concentrations were tried. The results showed that only one molecule of angiostatin was bound to the VEK-32L peptide (Figure 3-16). However, it was unclear whether the complex dissociated over the column since the VEK-32L peptide cannot be seen using SDS-PAGE.

Sedimentation velocity experiments were then performed using a Beckman XL-I analytical ultracentrifuge operated in absorbance mode (280 nm) at 20°C. Samples were rotated at speeds of 50,000 rpm and 45,000 rpm. For experiments conducted on unbound angiostatin, angiostatin dissolved in saline buffer (0.15 M NaCl and 100 mM Tris, pH 7.5) at concentrations of 0.5 mM and 1 mM were used. For experiments conducted on angiostatin/VEK-32L, a five molar excess of VEK-32L was used in the saline buffer previously used with unbound angiostatin. The data was analyzed through programs designed by W. Wedemeyer. The results showed that only one molecule of angiostatin binds to VEK-32L. However, a binding constant of VEK-32L for angiostatin was not determined. Because of this, it was unclear whether angiostatin was in fact binding to the VEK-32L peptide.

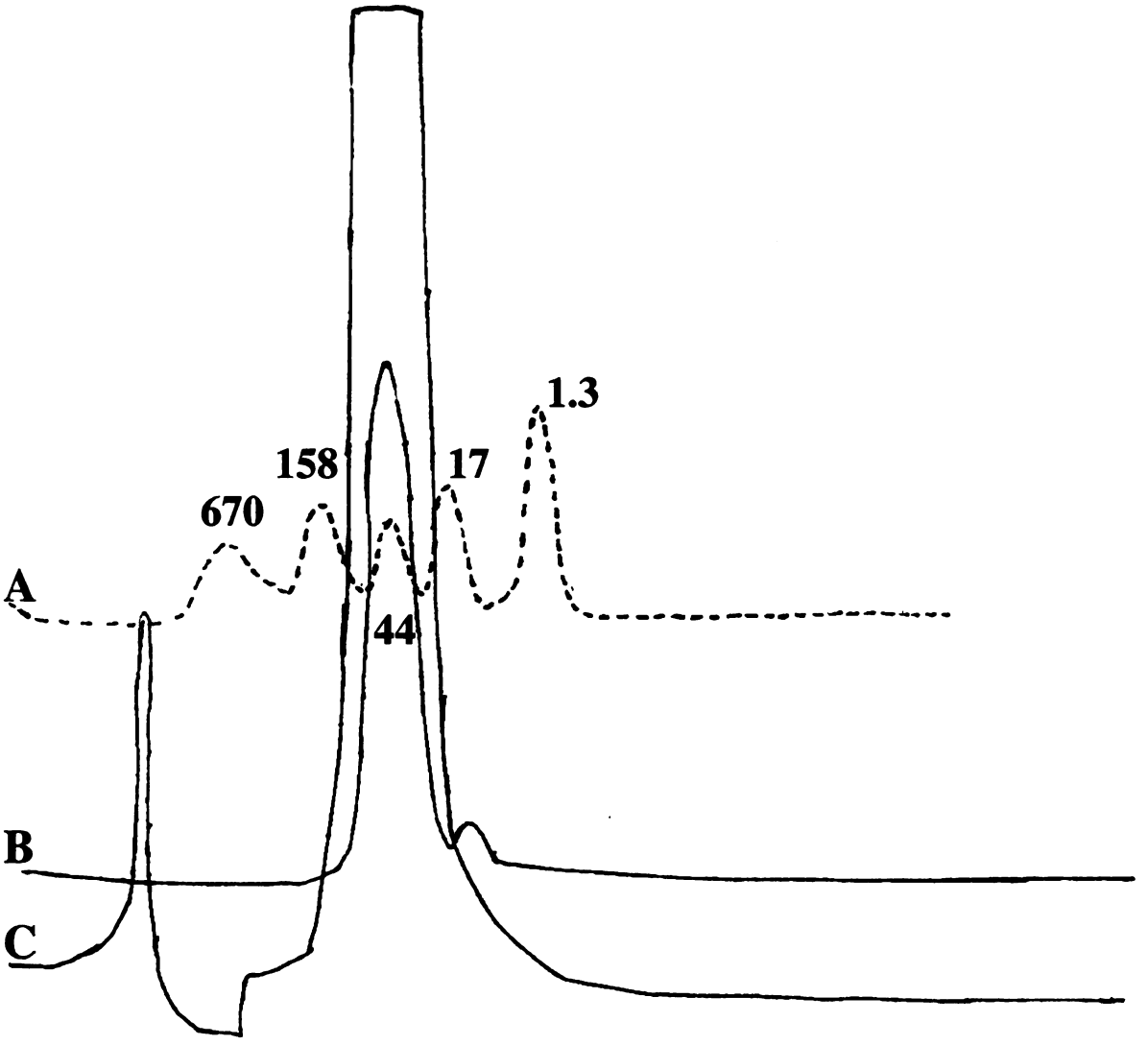

Figure 3-16 Gel filtration chromatogram results from Superdex 200 10/300 GL column. A) The dotted line is Biorad molecular weight standards (670 kDa thyroglobulin, 158 kDa bovine gamma-globulin, 44 kDa chicken ovalbumin, 17 kDa equine myoglobin, and 1.3 kDa Vitamin B12). B) Angiostatin/VEK-32L complex approximately 33 kDa. C) Unbound angiostatin approximately 30 kDa. All experiments were performed at a flow rate of 1 mL/min and the fractions were 5 mL.

3.8 Angiostatin Binding to Protein Domains

It has been determined that kringle domains not only bind six carbon zwitterions such as C-terminal lysines and EACA, but also protein domains[2,18-20]. The structures of the VEK-30 bound forms of K2 and angiostatin presented a model for protein binding at the surface of bacteria. The VEK-30 peptide has an arginine and a glutamate separated by almost one turn of the helix. This resembles a C-terminal lysine with a positive (R101) and a negative (E104) end that interacts with the K2 anionic and cationic residues in the LBS. This arrangement of residues in VEK-30 has been called a "pseudo-lysine." Known angiostatin ligands such as $\alpha_v\beta_3$ integrin and F(1)-F(0) ATP synthase contain similar "pseudo-lysine" moieties as seen in VEK-30 (Figure 1-13 and 1-15, respectively). This led us to explore other possible molecules relevant to angiostatin's mode of action. The X-ray structure of the angiogenesis inhibitor endostatin reveals that it contains an α-helix with the RGAD sequence[21]. R158 and D161 appear to form a pseudo lysyl site similar to the "pseudo-lysine" within the VEK-30 peptide (Figure 3-18). Figure 3-17 shows endostatin modeled in the angiostatin cavity between K2 and K3. Endostatin fills the cavity with very few collisions. Moreover, inspection of the angiostatin K3 LBS shows that endostatin's E272 may possibly interact with the K3 residues R290 and R324 (Figure 3-18). This validates the previously suggested observation that K3 is suited to bind short carboxylate ligands. Although there is no biochemical data indicating that angiostatin binds endostatin, there have been observations of an increase in tumor reduction when both inhibitors are given in combination with cancer patients[22].

Marneros and Olsen proposed a mechanism to explain the role of endostatin as an angiogenesis inhibitor[23]. This mechanism is based on the binding affinity of the endostatin domain of collagen XVIII. Collagen XVIII is involved in the activation of cell migration by interactions with extracellular components through its endostatin domain. It is proposed that endostatin binding competes for the binding of matrix components, inhibiting endothelial cell migration and ultimately angiogenesis. We propose that in the same way, angiostatin interacts with the endostatin domain of collagen XVIII by the interactions previously described thus providing an inhibitory mechanism of angiostatin.

Figure 3-17 Endostatin modeled into the angiostatin K2-K3 cavity. This was done by overlaying the helices of endostatin and VEK-30. Angiostatin is colored blue and endostatin is colored orange. R158 and D161 of endostatin are colored green with atom coloring.

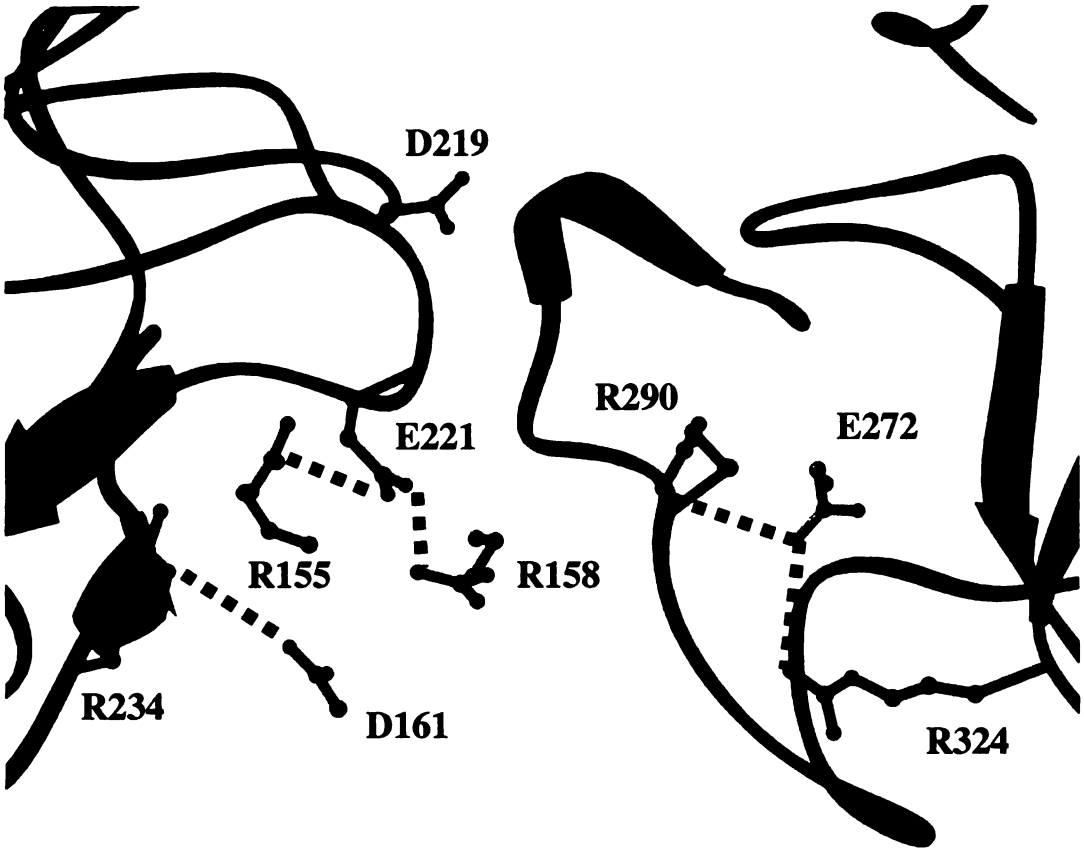

Figure 3-18 The residues possibly involved in endostatin/angiostatin binding are shown. Endostatin was modeled into the angiostatin K2-K3 cavity by overlaying the helices of endostatin and VEK-30. Angiostatin is colored green and endostatin is colored lavender. The residues are colored by atom.

3.9 Discussion

The angiostatin/VEK-30 structure is the only interaction between a multiple kringle-containing protein and a ligand to be characterized, and is the sole model for angiostatin complexation with its cognate protein-binding partners. Interestingly, other known angiostatin protein ligands, $\alpha_v\beta_3$ integrin and F(1)-F(0) ATP synthase, both contain a possible pseudo-lysine arrangement similar to that of VEK-30[7] (Figure 1-13 and 1-15, respectively). However, data defining the interface between angiostatin and these targets has not yet been reported.

The complex between VEK-30 and angiostatin suggests a possible mechanism for Pg activation during group *A. streptococci* infection. SK is secreted by group A streptococci and activates human Pg. It does so by binding to the catalytic domain of Pg, causing a conformational change in Pg within the complex that results in formation of an active site in the bound Pg (SK-Pg'). This SK-Pg' complex activates other molecules of Pg or is converted to SK-Pm, another potent Pg activator. Previous studies have shown that both SK and PAM contribute to virulence during group A streptococci impetigo[24]. Specifically, "molecular cooperation" between SK and PAM occurs, resulting in the bacterial acquisition of the host Pm so that the pathogen is able to invade host tissues[25]. Previous experiments suggest that streptokinase activates PAM-bound Pg and also preferentially binds the open form of Pg[26]. Our structure of the angiostatin/VEK-30 complex suggests that the PAM region might induce the open conformation of Pg on the cell surface by altering the relative orientation of the Pg kringle domains and by steric interference between its extended helical structure and Pg. Although VEK-30-mediated dimerization was not observed under physiologically-

relevant solution conditions, this does not rule out dimerization of Pg by full-length PAM on the bacterial cell surface. This dimerization may also promote a Pg conformation that is more amenable to SK binding and would also bring two molecules of Pg together on the cell surface. SK binding to one Pg could then result in proximity-accelerated activation of the second.

3.8 Literature Cited

1. Hayward, S. & Berendsen, H.J.C. Systematic Analysis of Domain Motions in Proteins from Conformational Change; New Results on Citrate Synthase and T4 Lysozyme. *Proteins, Structure, Function, and Genetics* **30, 144**(1998).

2. Rios-Steiner, J.L., Schenone, M., Mochalkin, I., Tulinsky, A., and Castellino, F. J. Structure and binding determinants of the recombinant kringle-2 domain of human plasminogen to an internal peptide from a group A Streptococcal surface protein. *J Mol Biol* **308**, 705-19 (2001).

3. McCance, S.G. & Castellino, F.J. Contributions of individual kringle domains toward maintenance of the chloride-induced tight conformation of human glutamic acid-1 plasminogen. *Biochemistry* **34**, 9581-6 (1995).

4. Cao, Y., Ji, R. W., Davidson, D., Schaller, J., Marti, D., Sohndel, S., McCance, S. G., O'Reilly, M. S., Llinas, M., and Folkman, J. Kringle domains of human angiostatin. Characterization of the anti-proliferative activity on endothelial cells. *J Biol Chem* **271**, 29461-7 (1996).

5. Lee, H., Kim, H. K., Lee, J. H., You, W. K., Chung, S. I., Chang, S. I., Park, M. H., Hong, Y. K., and Joe, Y. A. Disruption of interkringle disulfide bond of plasminogen kringle 1-3 changes the lysine binding capability of kringle 2, but not its antiangiogenic activity. *Arch Biochem Biophys* **375**, 359-63 (2000).

6. Nilsen, S.L., Prorok, M. & Castellino, F.J. Enhancement through mutagenesis of the binding of the isolated kringle 2 domain of human plasminogen to omega-amino acid ligands and to an internal sequence of a Streptococcal surface protein. *J Biol Chem* **274**, 22380-6 (1999).

7. Geiger, J.H. & Cnudde, S.E. What the structure of angiostatin may tell us about its mechanism of action. *J Thromb Haemost* **2**, 23-34 (2004).

8. Jones, D. THREADER: Protein Sequence Threading by Double Dynamic Programming. in *Computational Methods in Molecular Biology*, Vol. Chapter 13 (eds. Salzberg, S., Searls, D. & Kasif, S.) (Elsevier Science, 1998).

9. Whitby, F. & Phillips Jr., G. Crystal structure of tropomyosin at 7 Angstroms resolution. *Proteins* **38**, 49-59 (2000).

10. Collen, D. & Lijnen, H.R. Basic and clinical aspects of fibrinolysis and thrombolysis. *Blood* **78**, 3114-3124 (1991).

11. Winn, E.S., Hu, S. P., Hochschwender, S. M., and Laursen, R. A. Studies on the lysine-binding sites of human plasminogen. The effect of ligand structure on the binding of lysine analogs to plasminogen. *Eur J Biochem* **104**, 579-86 (1980).

12. Urano, T., Chibber, B.A.K. & Castellino, F.J. The Reciprocal Effects of Epsilon-Aminohexanoic Acid and Chloride-Ion on the Activation of Human (Glu1]Plasminogen by Human Urokinase. *J Biol Chem* **84**, 4031-4034 (1987).

13. Urano, T., Deserrano, V. S., Chibber, B. A. K., and Castellino, F. J. The Control of the Urokinase-Catalyzed Activation of Human Glutamic-Acid 1-Plasminogen by Positive and Negative Effectors. *J Biol Chem* **262**, 15959-15964 (1987).

14. Abad, M.C. et al. The X-ray crystallographic structure of the angiogenesis inhibitor angiostatin. *J Mol Biol* **318**, 1009-17 (2002).

15. Wistedt, A.C., Ringdahl, U., Muller-Esterl, W., and Sjobring, U. Identification of a plasminogen-binding motif in PAM, a bacterial surface protein. *Mol Microbiol* **18**, 569-78 (1995).

16. Wistedt, A.C., Kotarsky, H., Marti, D., Ringdahl, U., Castellino, F. J., Schaller, J., and Sjobring, U. Kringle 2 mediates high affinity binding of plasminogen to an internal sequence in streptococcal surface protein PAM. *J Biol Chem* **273**, 24420-4 (1998).

17. Berge, A. & Sjobring, U. PAM, a Novel Plasminogen-binding Protein from Streptococcus pyogenes. *J Biol Chem* **268**, 25417-25424 (1993).

18. Moser, T.L. et al. Angiostatin binds ATP synthase on the surface of human endothelial cells. *Proc Natl Acad Sci U S A* **96**, 2811-6 (1999).

19. Troyanovsky, B., Levchenko, T., Mansson, G., Matvijenko, O. & Holmgren, L. Angiomotin: an angiostatin binding protein that regulates endothelial cell migration and tube formation. *J Cell Biol* **152**, 1247-54 (2001).

20. Tarui, T., Miles, L.A. & Takada, Y. Specific interaction of angiostatin with integrin alpha(v)beta(3) in endothelial cells. *J Biol Chem* **276**, 39562-8 (2001).

21. Hohenester, E., Sasaki, T., Olsen, B.R. & Timpl, R. Crystal structure of the angiogenesis inhibitor endostatin at 1.5A resolution. *EMBO J* **17**, 1656-1664 (1998).

22. Yokoyama, Y., Dhanabal, M., Griffioen, A.W., Sukhatme, V.P. & Ramakrishnan, V. Synergy between angiostatin and endostatin: inhibition of ovarian cancer growth. *Cancer Res* **60**, 2190-2196 (2000).

23. Marneros, A.G. & Olsen, B.R. *Matrix Biol* **20**, 337-345 (2001).

24. Svensson, M.D., Sjobring, U., Luo, F. & Bessen, D.E. Roles of the plasminogen activator streptokinase and the plasminogen-associated M protein in an experimental model for streptococcal impetigo. *Microbiology* **148**, 3933-3945 (2002).

25. Ringdahl, U. et al. Molecular co-operation between protein PAM and streptokinase for plasmin acquisition by Streptococcus pyogenes. *J Biol Chem* **272**, 6424-6230 (1998).

26. Boxrud, P.D., and P.E. Bock. Streptokinase binds preferentially to the extended conformation of plasminogen through lysine binding site and catalytic domain interactions. *Biochemistry* **39**, 13974-13981 (2000).

Chapter IV

The Three Dimensional Structures of Calcium-Bound con-G and con-T[K7γ], and Cd^{2+}/Mg^{2+}/con-T[K7γ]

The physiological importance of the metal bound structures of con-G and con-T[K7γ] has been discussed in Chapter I. In order to get a better understanding of conantokin helix self-assembly, calcium-bound con-G and con-T[K7γ] (Cnudde, SE *et al.* 2007. *J Am Chem Soc. 129(6):* 1586-93) as well as Cd^{2+}/Mg^{2+}/con-T[K7γ] were crystallized and their structures determined. Structure determination including the crystallization conditions, data collection and refinement statistics are discussed in Chapter II. The structural characterization of these complexes is the first example of helix stabilization and self-association that is mediated by metal ion coordination and electrostatics with absolutely no component due to the hydrophobic effect. The details are discussed herein.

4.1 Overall Structure of Ca^{2+}/con-G and Ca2+/con-T[K7γ]

The X-ray crystal structures of the Ca^{2+}/con-G and Ca^{2+}/con-T[K7γ] have been determined at high resolution. Their overall structures are shown in Figure 2-13 and 2-15, respectively. The con-G α-helix is approximately 26 Å in length. A total of three Ca^{2+}, one Cl^-, and 68 water molecules were included in the model. Alternative side-chain conformations exist for con-G residues, Leu5, Gln9, and Ser16. The con-T[K7γ] α-helix is approximately 30 Å in length with a total of four Ca^{2+} and 142 water molecules. No alternative side-chain conformations exist in the Ca^{2+}/con-T[K7γ] structure.

4.2 The Structure of Ca^{2+}/con-G

In the presence of Ca^{2+}, con-G forms a helical antiparallel dimeric structure (Figure 2-13). Four of the Gla residues (3,7,10,14) are present on one face of the helix, specifically the Ca^{2+} coordination interface, while Gla4 resides on the opposite side of the helix and does not coordinate Ca^{2+}. A crystallographic two-fold axis passes through the central Ca^{2+} ion, Ca3, relating the two helices to form a dimer. Therefore, the con-G dimer contains a total of 5 Ca^{2+}, two associated with each of the crystallographically equivalent helices, and one shared between the two that is bisected by a crystallographic two-fold axis. The two helices are further referred to as Helix A and Helix B, for the unique and crystallographically equivalent helix, respectively. Ca^{2+} ions are further designated as Ca1A, Ca2A, Ca1B, Ca2B and Ca3, for those associated with Helix A, Helix B, and shared, respectively. Ca1A and Ca1B, and Ca2A and Ca2B, are related by crystallographic symmetry, as are Helix A and Helix B. Ca1A is coordinated by both carboxylates of two Gla residues, Gla3 and Gla7, accounting for a total of four ligands to Ca1A (Figure 4-1 top left). Similarly, Ca2A is chelated by Gla10 and Gla14, with all four carboxylates of the two Gla residues acting as ligands (Figure 4-1 top right). Both Ca1A and Ca2A have octahedral geometry with the two carboxylates of a single Gla residue occupying coordination sites separated by about 90°. The central Ca^{2+} ion, Ca3, is also coordinated by two Gla residues, Gla7 and Gla10, but each Gla residue donates only one carboxylate (Figure 4-1 bottom). Therefore, Helix A donates a total of only two ligands to Ca3. There is an identical arrangement of chelated Ca^{2+} ions on Helix B, as required by the crystallographic symmetry relating the two helices. It is clear that

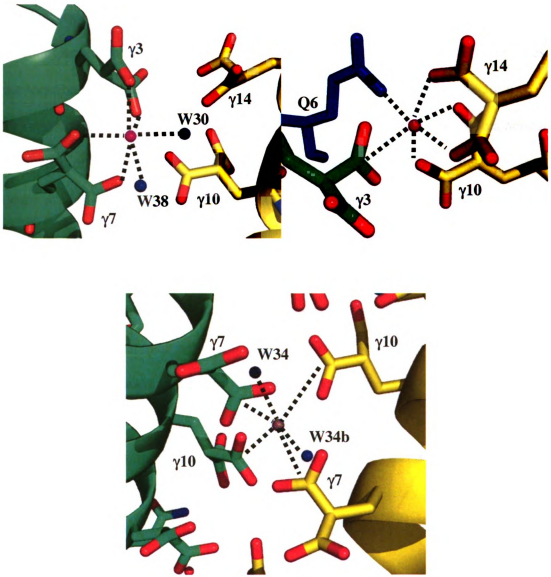

Figure 4-1 Ca^{2+} *coordination in con-G.* (Top left) Ca1A (magenta) coordinates with γ3 and γ7 and waters W30 and W38 (blue). (Top right) Ca2B (magenta) coordinates with γ3 and Q6 of helix A and γ14 and γ10 of the crystallographically-related helix B forming the dimer. Q6 is displayed in blue. (Bottom) Ca3 (gray) lies on the crystallographic twofold and coordinates with γ7 and γ10 of each helix as well as waters W34 and W34b (blue), a symmetry-related water molecule. The symmetry-related helix forming the dimer is colored yellow. In all cases, the peptide side-chain atoms are colored by atom type, except as otherwise indicated.

the helical conformation is stabilized by the metal-mediated interactions of Gla residues located one helical turn apart. Metal chelation enforces the positioning of all four Gla residues and, by nature of the Ca3 chelation of Gla7 and Gla10, enforces the location of all four Gla residues on a single helical face. A similar chelation arrangement is likely for other metal ions, such as Mg^{2+}, which also leads to a helical conformation in con-G. Absent metal ions, the electrostatic repulsion caused by the concentration of negative charges from the Gla residues on a single helical face would destabilize a helical conformation, consistent with all biophysical measurements that show that con-G is not helical in the absence of metal ions [1-5].

The dimerization interface between Helix A and Helix B is defined almost exclusively by intermolecular Ca^{2+} coordination between the two helices. There are no direct contacts between residues of the two helices, although there are some water-mediated interactions (*vide infra*). Three of the 5 Ca^{2+}, Ca2A, Ca2B, and Ca3, are involved in intermolecular coordination. Helix B donates two ligands to Ca2A, the side-chain carbonyl oxygen of Gln6 and a single carboxylate group of Gla3 (Figure 4-1 top right). This creates an octahedral coordination geometry about Gla3, with two of the six ligands originating from Helix B. Since Ca3 resides on a crystallographic two-fold axis that relates Helix A with Helix B, the coordination of Ca3 to Helix B is identical to that of Ca3 to Helix A. This results in four carboxylate ligands, one each from Gla7 and Gla10 from each of the two helices acting as ligands for Ca3. The other two coordination sites of Ca3 are occupied by water molecules to complete the octahedral geometry. The other two Ca^{2+}, Ca1A and Ca1B, do not participate in interhelical contacts (Figure 4-1 top left), though

mutational studies indicate that their ligands, Gla3 and Gla7, are both important for Ca^{2+}-mediated helix dimerization [6]. This is most likely due to the induction of a stable monomeric helix upon occupation of the Ca1A/Ca1B site formed by Gla3 and Gla7.

4.3 The Structure of Ca^{2+}/con-T[K7γ]

The overall structure of Ca^{2+}/con-T[K7γ] is shown in Figure 2-15. Unlike con-G, there is no crystallographic two-fold symmetry relating the helices of the con-T[K7γ] dimer, even though the two con-T[K7γ] helices are virtually identical (overlay in Figure 4-2). Con-G contained one molecule in the asymmetric unit whereas the con-T variant con-T[K7γ] contains two molecules in the asymmetric unit. Also, the Ca^{2+} coordination within each con-T[K7γ] helix is also identical. There are a total of four Ca^{2+} bound to the con-T[K7γ] helix, Ca1A and Ca2A, which are predominantly associated with Helix A, and Ca1B and Ca2B, which are associated with Helix B of the dimer. As occurs with Ca1A and Helix A of con-G, Ca1A coordinates Helix A of con-T[K7γ], *via* a tetravalent interaction between all of the carboxylate residues of Gla3 and Gla7. The analogous scheme exists for Ca2A, which chelates all four carboxylates of Gla10 and Gla14 of Helix A, resulting in an essentially identical interaction to that seen for Ca2A of con-G. Ca1B and Ca2B of con-T[K7γ] are correspondingly identical to Ca1B and Ca2B of con-G. Taken together, it is clear that when a single helix is viewed in isolation, the con-G and con-T[K7γ] structures are very similar, except that Ca3 appears to be missing in the con-T[K7γ] structure (Figure 4-3 top and 4-3 bottom).

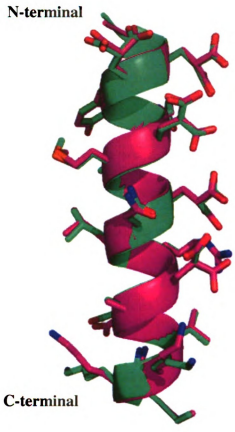

Figure 4-2 Comparison of the con-T[K7γ] helices. Overlay of the two con-T[K7γ] helices show their *identity*.

However, the dimerization interface of con-T[K7γ] is not similar to that of con-G. Given the structural similarity between con-G and con-T[K7γ], it was surprising to see a difference in the dimer interface between the two structures. While both peptides yield antiparallel helical dimers, and both are largely stabilized by Ca^{2+}-mediated chelation, the relative orientation of the helices as well as the Ca^{2+} coordination in the interface, is radically different in the con-T variant (overlays in Figure 4-3).

The two helices of con-T[K7γ] are almost parallel and all four Ca^{2+} are involved in the dimerization interface of con-T[K7γ]. As seen in Figure 4-4 a, Ca1A, chelated by Gla3 and Gla7 of Helix A, coordinates one carboxylate of Gla14 of Helix B, while Ca2A, chelated by Gla10 and Gla14 of Helix A, coordinates one carboxylate each of Gla7 and Gla10 of Helix B (Figure 4-4 b). The coordination sphere of Ca1B is identical to that of Ca1A (Figure 4-4 c), and the coordination sphere of Ca2B (Figure 4-4 a) is identical to that of Ca2A, except that the residues of Helix A are transposed for the residues of Helix B. The result of this chelation pattern is that the two helices are staggered by approximately one helical turn along the colinear helix axes and the Gla residues account for all of the Ca^{2+} chelation in con-T[K7γ]. In addition to the Ca^{2+} interactions, a short salt bridge exists between Arg13 of Helix A and Gla3 of Helix B and an identical salt bridge is formed between Arg13 of Helix B and Gla3 of Helix A to complete the symmetric interaction (Figure 4-4 d).

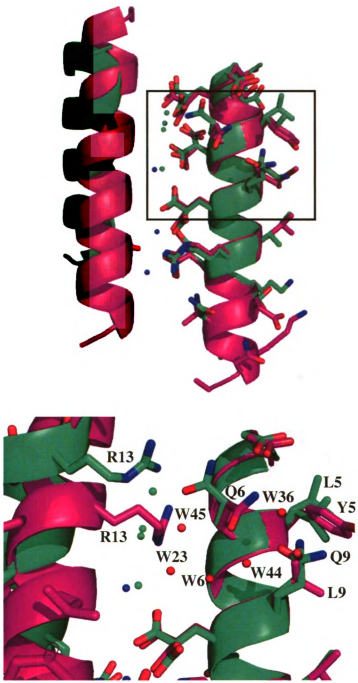

Figure 4-3 Comparison of the con-T[K7γ] and con-G helices. (Top) Overlay of the con-T[K7γ] (magenta) and con-G(green) helices demonstrates that *these helices are nearly identical to each other*. Blue spheres represent Ca^{2+} ions of con-G and green spheres designate Ca^{2+} ions of con-T[K7γ]. (Bottom) The boxed area is magnified and includes waters (red spheres). In all cases, the peptide side-chain hetero-atoms are colored by atom type.

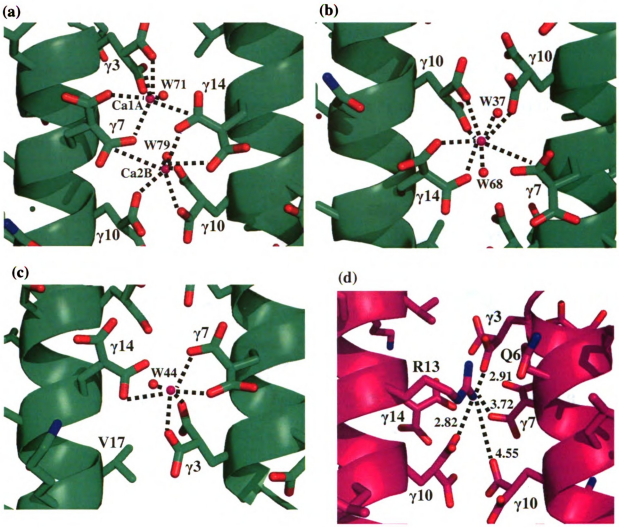

Figure 4-4 Ca^{2+} coordination in the con-T[K7γ] dimer. (a) Ca1A (magenta) coordinates with γ3 and γ7 of helix A and γ14 of helix B, as well as with one water molecule, W71 (red sphere). Ca2B (magenta) coordinates with γ7 and γ10 of helix A and γ14 and γ10 of helix B, along with water molecule W79 (red sphere). (b) Ca2A (magenta) coordinates with γ10 and γ14 of helix A and γ10 and γ7 of helix B. Ca2A also coordinates with two water molecules, W37 and W68 (red spheres). (c) Ca1B (magenta) coordinates with γ14 of helix A and γ7 and γ3 of helix B. Ca1B (magenta) also coordinates with one water molecule, W44 (red sphere). (d) The salt bridge interactions in con-T[K7γ]. All distances are shown in Angstroms. In each case, the peptide side-chain hetero-atoms are colored by atom type. Helix A is always displayed on the left whereas Helix B is displayed on the right.

A further stabilization of the con-T[K7γ] dimeric structure involves a spine of water molecules that runs between the Ca^{2+}, (Figure 4-5), all of which are chelated by Ca^{2+}. This continuous spine of hydration is not seen in the con-G structure.

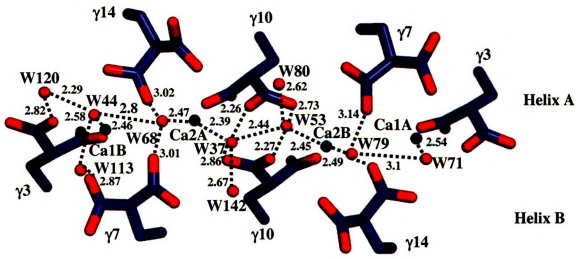

Figure 4-5 The con-T[K7γ] dimer interface is significantly different from con-G. A water network exists at the Ca^{2+}/dimer interface of Ca^{2+}/con-T[K7γ]. All distances are shown in Angstroms. The Ca^{2+} ions are shown as blue spheres and the waters are shown as red spheres. The top four Gla residues belong to Helix A and the bottom four Gla residues belong to Helix B. The peptide side-chain hetero-atoms are colored by atom.

The fundamental differences between the dimer interface in con-G versus con-T[K7γ] are illustrated by the schematic "wiring diagram" presented in Figure 4-6. First, there is a clear correspondence between Ca1A, Ca1B, Ca2A and Ca2B in all of the structures, since they have identical chelation from one of the two helices. For example, all Ca1-type ions are chelated to Gla3 and Gla7 of a single helix by both of their carboxylate moieties, while Ca2-type ions are chelated to Gla10 and Gla14 of a single helix by both of their carboxylate moieties. In addition, all four helices coordinate a single Ca^{2+} *via* Gla7 and Gla10 in bidentate fashion, using only one carboxylate from each Gla. Therefore, if any of the four helices are viewed in isolation, the Ca^{2+} coordination appears to be identical. The dramatic difference occurs because each Ca^{2+} interacts differently with the adjacent helix in the two structures. For example, in con-T[K7γ] all Ca1-type ions chelate with a single carboxylate of Gla14 on the partner helix, thus creating a symmetric interaction. On the other hand, in con-G Ca1-type ions do not chelate with their partner helix. Furthermore, the Ca2-type ions of con-T[K7γ] chelate Gla7 and Gla10 of their partner helix, while in con-G, the Ca2-type ions chelate the side-chain of Gln6, while Ca3 coordinates Gla7 and Gla10 residues from both helices. The result of this coordination difference is to cause the con-G helices to be twisted out of plane to accommodate the octahedral coordination about Ca3, while the helices in con-T[K7γ] are parallel, but are shifted along the helix axis by one turn to accommodate the chelation of both Ca1 and Ca2 (Figure 4-7).

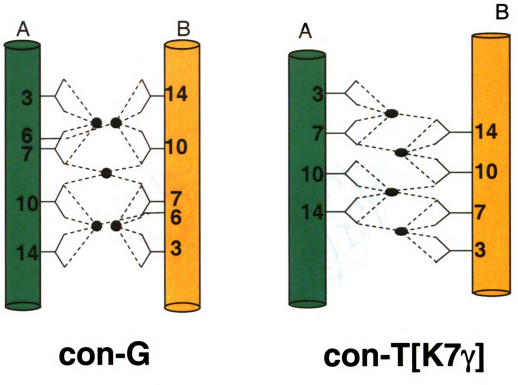

Figure 4-6 The con-T[K7γ] dimer interface is significantly different from con-G. The fundamental differences between the Ca^{2+}/con-G structure and Ca^{2+}/con-T[K7γ] structure are shown in a schematic wiring diagram. Each dotted line represents a coordination between one Gla carboxylate and a Ca^{2+}. Helix A is colored green and Helix B is colored yellow. The black spheres are Ca^{2+}.

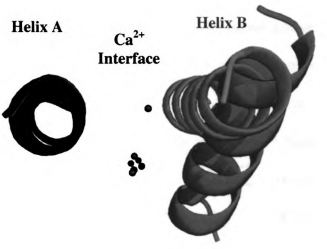

Figure 4-7 The con-T[K7γ] dimer interface is significantly different from con-G dimer. Overlay of the con-G and con-T[K7γ] structures showing difference in helix alignment. The con-G dimer is colored green and its Ca^{2+} ions colored red. The con-T[K7γ] dimer is colored magenta and its Ca^{2+} ions colored blue. In all cases, peptide side-chain hetero-atoms are colored by atom.

4.4 Discussion

While both con-G and the con-T variant (con-T[K7γ]) provide an identical Gla pattern that allows dimerization in the presence of Ca^{2+} in an antiparallel orientation and would therefore be expected to yield similar Ca^{2+}-bridged dimeric interfaces, this was not the case. There is a substantial difference in dimerization and Ca^{2+} affinity between these two peptides. In fact, con-T[K7γ] exhibits a 10-fold higher association constant for dimerization and 10-fold higher affinity for Ca^{2+} than con-G [7]. This is consistent with the present structures as there are four Ca^{2+} in the dimer interface of con-T[K7γ] and only three in con-G. There are also more bridging interactions in con-T[K7γ], with six cross-helix ligand interactions compared to only four in con-G (these are Ca^{2+}-mediated interactions where the metal ion is assumed to be bound to one of the helices and coordination sites to the partner helix are counted). There are also two tight salt-bridge interactions (involving Arg13 and Gla3) between helices in con-T[K7γ] that are not seen in con-G. A spine of water molecules exists that span all four Ca^{2+} in con-T[K7γ]. Each of the six water molecules that constitute this spine are coordinated by one of the four Ca^{2+} ions. This spine of hydration is absent in the con-G structure, at least partly due to the fact that the Ca^{2+} ions in con-G have fewer coordinated water molecules. Taken together, these observations are consistent with the higher binding and dimerization affinity of con-T[K7γ] compared to that of con-G.

Our structure is also consistent with the fact that native con-T binds Ca^{2+} without dimerization [6]. The absence of Gla7 destabilizes, if not eliminates, the Ca1A

and Ca1B binding sites, leading to loss of two helix bridging interactions to Gla14. Gla7 is also involved in helix bridging interactions to Ca1A and Ca2B. Therefore, elimination of this residue leads to probable loss of three of the helix bridging interactions seen in the structure. It is interesting to note that metal-free native con-T has considerable helical content as opposed to both con-G and con-T[K7γ][7]. This is probably due to two factors. First, electrostatic repulsions between the four Gla residues on one side of a helix almost certainly lead to destabilization. Gla7 contributes significantly to this destabilization because it resides in the center of the helix. Replacement of Gla7 with a positively charged side-chain (e.g., Lys), would significantly reduce this electrostatic repulsion and may possibly add electrostatic stabilization by forming salt bridging interactions with both Gla3 and Gla10. Consistent with this is the fact that the con-G[γ7K] mutant is significantly more helical in the absence of metal ions than is native con-G [8]. The con-T peptide is also four residues longer than con-G, and all of these residues are helical in the crystal structure. It is well acknowledged that longer helices are more stable than shorter helices.

Less clear is what causes the different dimerization interface of con-T[K7γ] relative to con-G. Although eight amino acids differ between the two peptides, all of the residues that are involved in Ca^{2+} binding and dimerization are conserved. The overlay of the con-G and con-T[K7γ] helices shows that most of the residues in the interface have identical conformations in the two structures. The most significant difference is in the conformation of Gln6, which is flipped into the interface and chelates Ca2B (or Ca2A for the B helix) in con-G, but is flipped out of the interface

in con-T[K7γ] and interacts with Tyr5, a residue that is not conserved in the two peptides. Hence, the conformation of Gln6 appears to be critical to both dimer interfaces. As one of the cross-helical Ca^{2+} ligands in the con-G interface, the flipped-out Gln6 conformation would compromise this mode of Ca^{2+} coordination. On the other hand, in con-T[K7γ], the flipped-in conformation of Gln6 would sterically clash with Arg13 from the other helix and prevent the formation of the salt bridge between Arg13 and Gla3 in con-T[K7γ]. Overlaying the two helices indicates that the Gln6 flipped-out conformation is opposed in con-G both by the conformation of Gln9, which is also not conserved in con-T, and the resulting water structure around Gln9 and Glu2. The flipped-out conformation would bring the Gln6 side-chain oxygen atom within 3 Å of the Gln9 oxygen atom and within 2 Å of two very well-ordered water molecules that form a bridge between Glu2 and Gln9, all of which represent electrostratic and steric clashes. The side-chain oxygens of both Gln6 and Gln9 are well-defined, since the Gln6 side-chain nitrogen must interact with Glu2, and the Gln9 nitrogen must interact with the main-chain carbonyl of Leu5. Therefore the dimer interface is defined primarily by the metal chelation of the four critical Gla residues, Gla 3, 7, 10 and 14, but is also strongly affected by subtle, indirect effects on the conformation of other side-chains as well.

4.5 Mutational Studies of the con-G and con-T Peptides

The systematic mutational studies of both con-G and con-T [8,9] have shed much light on our structural results. For example, deletion of residues from the C-terminus of either the con-T or con-G helices leads to reduced helicity. This is not surprising given that longer helices are usually found to be more stable than shorter

ones and our structure demonstrates that an end-to-end helix exists for both peptides. Further, individual replacement of Gla residues 3, 7, 10 or 14 leads to significantly reduced helicity for con-G [8], consistent with our observation that chelation of Ca^{2+} to these residues defines and stabilizes the helical peptide structure. Though replacement of many of the other residues also leads to some loss of helicity, it is particularly interesting to note that in con-G, the R13A variant manifests a nearly complete loss of helicity in the presence of Ca^{2+}. In the case of con-T, the R13A replacement likewise diminishes the helix content of both Ca^{2+}-bound and metal-free peptide forms [8,9]. In our structure, Arg13 makes a tight salt bridge with Gla10 of the same helix in both con-G and con-T[K7γ], and in the latter makes an additional salt bridge with Gla3 of the adjoining helix. Thus, it appears that Arg13 is important for stabilization of the dimer form, through neutralization of the considerable negative charge within the helix-helix interface and also participates in intrapeptide electrostatic contacts that stabilize the monomeric form of apo con-T.

4.6 The con-G and con-T Structures and NMDA Receptor Binding

Those residues found by alanine replacement to be important for NMDA receptor binding or antagonism define one face of both the con-G and con-T helices (Figure 2-13 and 2-15) [9]. The N-termini of the peptides are especially critical for biological activity, consistent with the fact that the N-termini of con-T and con-G are identical through the first four residues. Gla4 is particularly important, since mutation of this residue, even conservatively to Asp or Glu, results in no detectable NMDA receptor activity [9]. The fact that these residues are localized to one face of the helix strongly indicates that both peptides bind the receptor in a helical

conformation, even though replacements that significantly destabilize the con-G or con-T helix do not always have a significant effect on NMDA receptor binding (e.g., γ10A). However, for both con-G and con-T, the γ3A replacement results in the reduction of IC_{50} values by at least an order of magnitude. Our hypothesis is that NMDA receptor binding is mostly dependent on the maintenance of the helical conformation in the N-terminal portion of the helix. Alanine replacement of γ3 results in helix fraying at the N-terminus, possibly through loss of the stabilizing effect that the negatively charged Gla residue imparts to the helix macro-dipole at the N-terminus thereby reducing peptide affinity for the receptor. In further support of the helical hypothesis is con-G[γ7P], which is incapable of helix formation and is completely inactive at the NMDA receptor [8]. On the other hand, the C-terminal part of the helix is less critical for the interaction and therefore loss of helicity in this region is less deleterious to NMDA receptor binding.

4.7 Comparison of the Crystal Structures to the NMR Structures

Several NMR-based structures of both metal-bound and apo forms of con-T and con-G have been determined [2-5,10]. Most relevant to this discussion are the NMR-derived structures of Ca^{2+}-complexed con-G [25-27]. While the helical backbones of these structures are very similar to the X-ray structure determined herein (Figure 4-8), neither of the NMR-derived structures predicts the dimerization interface. This is consistent with our data, which indicate that there are no interhelical proton distances closer than 5 Å. This would obviate detection of the dimeric species by 2D-NMR, since NOEs do not typically evolve from proton-proton distances with the mixing time employed in the afore-referenced studies.

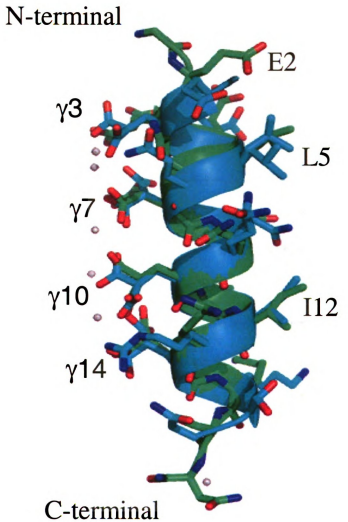

Figure 4-8 Overlay of the NMR and X-ray crystal structures of Ca^{2+}/con-G. Cyan, the X-ray structure of Ca^{2+}/con-G; green, a low energy NMR structure[27]; white, Ca^{2+}. In all cases, peptide side-chain hetero-atoms are colored by atom type.

While side-chain orientations on the Gla-rich face of the helix are similar between the NMR and X-ray structures, profound differences are observed at the opposite face and at the termini. These latter disparities likely reflect some of the uncertainties involved in garnering NMR-based structural information in the more flexible regions of the peptide, wherein an ensemble of solution conformers must be averaged. Furthermore, computational docking of Ca^{2+} into the structure of con-G [27] yields stoichiometry and placement that is different from that associated with the crystal structure. A total of four bound Ca^{2+} are predicted from the NMR data whereas three exist per monomer, in the crystal structure. While three of these Ca^{2+} ions form a similar coordination network in both structures, the NMR-derived complex deviates from our structure in that it invokes Gla3 and Gla4 as ligands for a fourth Ca^{2+}. Also, the Ca^{2+}-docked NMR structure does not reveal Ca^{2+} coordination to the side-chain carbonyl of Gln6 of con-G.

It is difficult to accurately predict metal positions in peptides and proteins solely by using conventional NMR in tandem with modeling techniques. While single crystal X-ray diffraction is a powerful approach for elucidating the subtle details of metal ion binding, it is not without uncertainties caused by the structural ramifications of crystal packing interactions, whereas dynamic solution structures are without these constraints. Thus, both NMR and X-ray crystallography continue to provide valuable information even though there are some differences using such disparate approaches.

4.8 Conclusions From the Calcium-Bound Structures

Limited success has been obtained in the past in the design of short peptides that adopt stable, structurally homogeneous α–helical conformations. Almost all helices are stabilized by the hydrophobic effect, promoting the formation of a hydrophobic helical face that is critical to stabilize the structure. It is also the primary force that drives helix self-assembly as seen in coiled-coil multimers, e.g., the leucine zipper motif. A possible alternative to the hydrophobic effect is the use of metal ion coordination to electrostatically stabilize the helical conformation. However, to date, there are no examples of a peptide that is solely dependent on metal ion chelation for both helix formation and self-assembly. Further, there were no high resolution crystal structures of any of the metal-promoted helices, leaving open the full nature of the metal coordination. Employing the con-G and con-T[K7γ] peptides, we provide the first example of helix stabilization and self-assembly that is completely dependent on electrostatics and metal coordination, with no component of the stabilization due to the hydrophobic effect. The unique dispensation of Gla residues on one face of the helix, resulting in a regular coordination of metal ions along this face, give rise to this effect, resulting in what we call a "metallo-zipper motif". In addition, the natural peptides form specific anti-parallel helix dimers in the presence of Ca^{2+}, and fold into monomeric helices in the presence of Mg^{2+}. Our structures of these peptides show very clearly the nature of the metal chelation that leads to helix formation and also elaborate the dimerization interface in both peptides. The structures also demonstrate that while the metal chelation is very similar in the two peptides, they surprisingly result in completely

different dimer interfaces. The dimer interface appears to be exquisitely sensitive to amino acid changes that are not directly involved in the metal chelation that results in dimerization. This is the first example of a short peptide helix that is completely stabilized by metal ion chelation. Helix formation depends only on the disposition of the metal chelating Gla residues. The dimer interface consists exclusively of metal coordination interactions in the case of the con-G dimer, while a salt bridge interaction plays an important role in defining the con-T[K7γ] interface. These structures provide a fertile platform for the design of short peptide helices that are comparatively insensitive to amino acid sequence in all but a few positions. The structures also provide a much better understanding of the mechanism of metal-dependent oligomerization.

4.9 The Overall Structure of Cd^{2+}/Mg^{2+}/con-T[K7γ]

The X-ray crystal structure of the Cd^{2+}/Mg^{2+}/con-T[K7γ] has been determined at high resolution. The overall structure is shown in Figures 2-18 and 2-19. The con-T[K7γ] α-helix is approximately 30 Å in length with a total of 4 Cd^{2+}, 1 Mg^{2+}, and 61 water molecules. Alternative side-chain conformations exist for residues Met8 and Val17. Electron density was also seen for the C-terminal NH_2 atom.

Much like other conantokin structures such as Ca^{2+}/con-G and Ca^{2+}/con-T[K7γ], Gla residues 3, 7, 10, and 14 reside on the same face of the helix. However, three cadmiums are seen at the Gla interface. In Figure 4-9a, Cd2 is six-coordinate chelating residues Gla3 and Gla7 of the helix as well as Gla3 of a crystallographic-related symmetry molecule. Cd2 also coordinates with water W20. However, Cd2

is only 50% occupied. It is unclear as to whether this partially occupied Cd^{2+} is also partially occupied by Mg^{2+} since magnesium does not have an anomalous signal (Figure 2-18). Figure 4-9b shows that Cd3 is tetrahedrally coordinated involving γ4 and γ7 of the helix as well as γ4 and γ7 of a symmetry-related molecule. Also, Cd3 sits directly on a crystallographic two-fold axis. Unlike Cd2, Cd3 has 100% occupancy. The Cd3 metal ion site is unique since the previous crystal structures of both con-G and con-T[K7γ] bound to Ca^{2+} do not have a metal ion coordinating γ4 and γ7. Figure 4-9c shows Cd4 coordination. Cd4 is octahedrally coordinated and involves γ10 of the helix and a symmetry-related γ10. Cd4 also coordinates with two waters W22a and W51a as well as their symmetry-related waters W22b and W51b. Much like Cd3, Cd4 sits directly on a crystallographic two-fold. One Mg^{2+} is also seen at this interface (Figure 4-10). Mg1 is octahedrally coordinated involving residues γ10 and γ14. Mg1 makes bidentate contacts with both carboxylates of γ10 and γ14 as well as coordinating with waters W23 and W24. However, W23 exits in two different conformations (W23a and W23b) and Mg1 makes contact with W23a only since W23b is too far (2.53 Å). Unlike the other Cd^{2+}, Mg1 does not coordinate with a crystallographic symmetry-related molecule. The primary role of Mg1 involves Gla residue chelation of a single helix. On the opposite face of the helix, Cd1 coordinates with the helix and a symmetry-related molecule (Figure 4-11). Cd1 coordinates with both the nitrogen and oxygen atoms of G1 of the helix. Cd1 also coordinates with E16 and K19 of a crystallographic related molecule. Two waters are also involved at this site. W18 and W19 both coordinate with Cd1. The Cd^{2+}

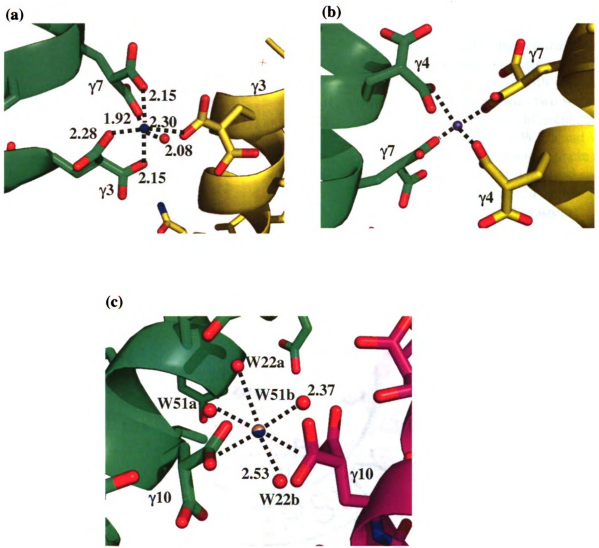

Figure 4-9 Cadmium coordination at the calcium interface. (a) *Cadmium Cd2 is six-coordinate.* The overall structure of Cd^{2+}/Mg^{2+}-con-T[K7γ] is shown in green. Cd2 is a blue sphere and the water W20 is a red sphere. A crystallographic-related molecule is colored yellow. (b) *Cadmium coordination at Cd3 is tetrahedrally coordinated.* Cd3 sits directly on a crystallographic two-fold axis and coordinates both γ4 and γ7 of the structure and its symmetry-related molecule. Distances to each oxygen atom range from about 2.3-2.6Å. The structure is shown in green and the symmetry-related molecule is colored yellow. Cd3 is shown as a blue sphere. (c) *Cadmium coordination at Cd4 is octahedrally coordinated.* Cd4 coordinates with γ10 of the structure (green) and a crystallography-related molecule (magenta). Cd4 also coordinates with waters W22a and W51a and the symmetry-related waters W22b and W51b. Cd4 is partially colored as a blue sphere and all waters are colored as red spheres. All distances are measured in Angstroms. The side-chains are colored by atom.

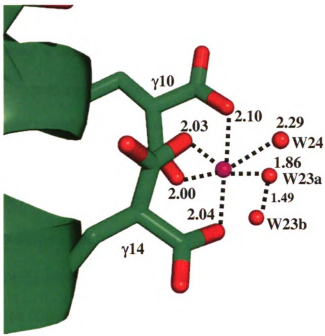

Figure 4-10. *Magnesium coordination of Mg1 is octahedrally coordinated.* Mg1 coordinates bindentately with γ10 and γ14. Mg1 also coordinates with two waters W24 and W23. W23 has two different conformations shown as W23a and W23b. Mg1 is colored as a magenta sphere and the water molecules are shown as red spheres. The side-chains are colored by atom. All distances are shown in Angstroms.

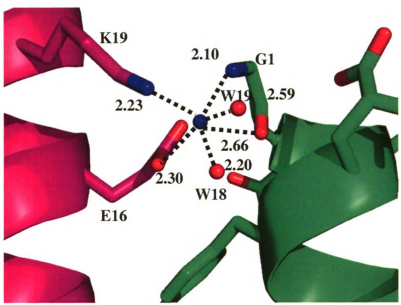

Figure 4-11. O*n the opposite side of the calcium interface cadmium coordination at Cd1 is six coordinate.* Cd1 coordinates with E16 and K19 of the crystallographic symmetry-related molecule. Cd1 also coordinates with G1 and two waters W18 and W19. All distances are shown in Angstroms. The structure is colored green and a symmetry-related molecule is colored magenta. Cd1 is colored as a blue sphere and the waters are shown as red spheres. All side-chains are colored by atom.

distances range from about 1.9 Å – 2.5 Å whereas the Mg^{2+} distance ranges from about 1.8 Å – 2.3 Å. Cadmium is essential for crystallization since no crystals grew in the absence of cadmium or decreased concentrations of cadmium. The structure confirms the need for cadmium during crystallization since three out of the four cadmium sites reside on crystallographic axises proving that cadmium is crucial to form a crystalline lattice.

4.10 Comparison between Ca^{2+}-con-T[K7γ] and Cd^{2+}/Mg^{2+}-con-T[K7γ]

An overlay of the Cd^{2+}/Mg^{2+}-con-T[K7γ] onto the Ca^{2+}-con-T[K7γ] dimer is shown in Figure 4-12. Gla residues 3, 4, and 7 are move significantly relative to the metal ion binding sites whereas γ10 and γ14 are virtually in the same position. γ4 moves the most because of the γ4-γ7 Cd^{2+} chelation that exists only in the Cd^{2+}/Mg^{2+}-con-T[K7γ] structure. Interestingly, K18 is also able to move significantly. However, this is due to crystal packing interactions. Also, Mg1 is near one of the Ca^{2+} (Ca1B) and Cd2 is near Ca1A. However Cd4 is not close to Ca2A and there is no Ca^{2+} near Cd3.

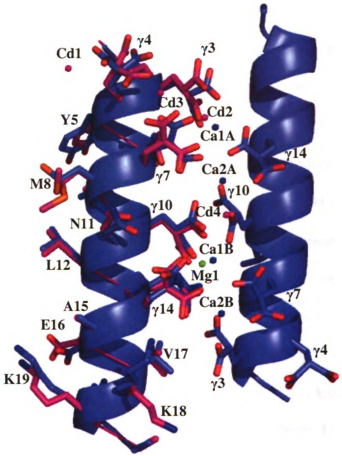

Figure 4-12 Gla residues 3, 4, and 7 are able to move relative to the specific metal cation. Overlay of the Cd^{2+}/Mg^{2+}-con-T[K7γ] structure (magenta) onto the Ca^{2+}-con-T[K7γ] dimer (blue) structure. Cd^{2+} are shown as magenta spheres, Mg^{2+} is shown as a green sphere, and Ca^{2+} are blue spheres. The side-chains are colored by atom.

In order to understand why Mg^{2+} does not cause dimerization in either con-T[K7γ] and con-G, we must take a closer look at the Mg^{2+} metal cation site. Figure 4-12 shows that the Mg^{2+} metal cation is relatively close to one of the Ca^{2+} metal sites (Ca1B). Figure 4-13 shows an overlay of the Cd^{2+}/Mg^{2+}-con-T[K7γ] structure onto the Ca^{2+}-con-T[K7γ] dimer. The distance between Mg1 and Ca1 is about 0.84Å. Ca1 coordinates with γ10 and γ14 of one helix and γ7 and γ10 of the dimer helix whereas Mg1 only coordinates with γ10 and γ14. The distances for Mg1 coordination with the Ca^{2+} dimer helix residues γ10 and γ7 are 3.08Å and 2.91Å, respectively. This distance is much too far for Mg^{2+}. If you were to replace Ca1 with Mg1, the chelation distances would still be too far for Mg^{2+}. Mg^{2+} generally chelates with a distances between 1.7-2.3 Å whereas Ca^{2+} coordinates with distances ranging from 2.2-2.6 Å. This is expected since Mg^{2+} is a much smaller metal cation than Ca^{2+}. Another difference occurs in the coordination requirements. Ca^{2+} is not as rigid in its requirements for coordination. For example, in the structures of Ca^{2+}-con-G and Ca^{2+}-con-T[K7γ], Ca^{2+} chelation is six-coordinate and occasionally even seven-coordinate. However, Mg^{2+} metal cations are most often octahedrally coordinated. We postulate that the reason why Mg^{2+} does not form the dimer is because Mg^{2+} is too small of a cation to make the larger distance contacts seen at the calcium induced dimerization interface. Since Mg^{2+} coordinates with distances closer than that of Ca^{2+}, Mg^{2+} will not be able to reach critical Gla residues to interact with the dimer. This explains why con-T[K7γ] and con-G do not dimerize in the presence of Mg^{2+}.

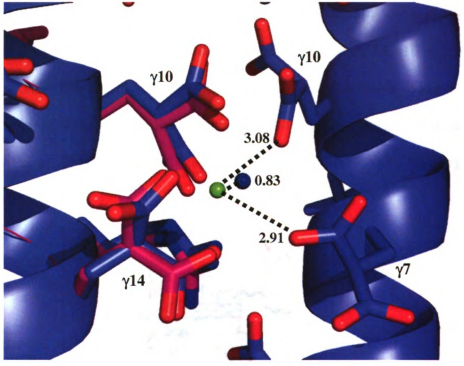

Figure 4-13 An overlay of the Cd^{2+}/Mg^{2+}-con-T[K7γ] structure onto the Ca^{2+}-con-T[K7γ] dimer. The Ca^{2+}-con-T[K7γ] dimeric structure is colored blue and the Cd^{2+}/Mg^{2+}-con-T[K7γ] structure is colored magenta. Mg1 is shown as a green sphere whereas Ca1 is shown as a blue sphere. The distances shown are all in Angstroms. The side-chains are colored by atom.

Another issue that must be understood is why Mg^{2+} only chelates γ10 and γ14 and not any of the other Gla residues specifically γ3 and γ7. Figure 4-14 shows a closer view of the Mg^{2+} site and the γ3-γ7 chelation pair. The structure shows that R13 stabilizes the negative charge of the Gla residues (γ10 and γ14) allowing the Mg^{2+} to bind better. The γ3-γ7 chelation pair is not as negative as the γ10-γ14 pair because of R13. The Mg^{2+} has also forced γ10 and γ14 to move closer to one another (Figure 4-13) than from the Ca^{2+} bound con-T variant structure. This cannot happen at the γ3-γ7 pair since γ4 has flipped away to stabilize the negative charge and chelate Cd^{2+}.

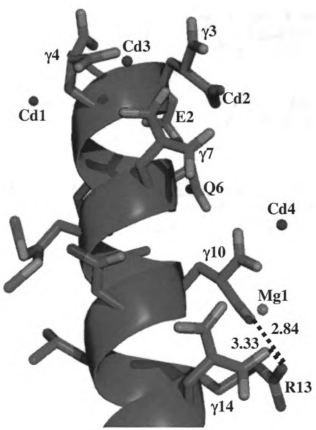

Figure 4-14 A closer view of the Mg^{2+} site and the γ3-γ7 chelation pair. All distances are shown in Angstroms. All side-chains are colored by atom. The magenta atom is Mg1 and the blue atoms are Cd^{2+}.

4.11 Literature Cited

1. Prorok, M., Warder, S.E., Blandl, T. & Castellino, F.J. Calcium binding properties of synthetic gamma-carboxyglutamic acid-containing marine cone snail "sleeper" peptides, conantokin-G and conantokin-T. *Biochemistry* **35**, 16528-34 (1996).

2. Skjaerbaek, N., Nielsen, K.J., Lewis, R.J., Alewood, P. & Craik, D.J. Determination of the solution structures of conantokin-G and conantokin-T by CD and NMR spectroscopy. *J Biol Chem* **272**, 2291-9 (1997).

3. Rigby, A.C., Baleja, J.D., Furie, B.C. & Furie, B. Three-dimensional structure of a gamma-carboxyglutamic acid-containing conotoxin, conantokin G, from the marine snail Conus geographus: the metal-free conformer. *Biochemistry* **36**, 6906-14 (1997).

4. Rigby, A.C. et al. Role of gamma-carboxyglutamic acid in the calcium-induced structural transition of conantokin G, a conotoxin from the marine snail Conus geographus. *Biochemistry* **36**, 15677-84 (1997).

5. Chen, Z. et al. Conformational changes in conantokin-G induced upon binding of calcium and magnesium as revealed by NMR structural analysis. *J Biol Chem* **273**, 16248-58 (1998).

6. Dai, Q., Prorok, M. & Castellino, F.J. A new mechanism for metal ion-assisted interchain helix assembly in a naturally occurring peptide mediated by optimally spaced gamma-carboxyglutamic acid residues. *J Mol Biol* **336**, 731-44 (2004).

7. Dai, Q., Castellino, F.J. & Prorok, M. A single amino acid replacement results in the Ca2+-induced self-assembly of a helical conantokin-based peptide. *Biochemistry* **43**, 13225-32 (2004).

8. Blandl, T., Prorok, M. & Castellino, F.J. NMDA-receptor antagonist requirements in conantokin-G. *FEBS Lett* **435**, 257-62 (1998).

9. Warder, S.E., Blandl, T., Klein, R.C., Castellino, F.J. & Prorok, M. Amino acid determinants for NMDA receptor inhibition by conantokin-T. *J Neurochem* **77**, 812-22 (2001).

10. Warder, S.E. et al. The NMR solution structure of the NMDA receptor antagonist, conantokin-T, in the absence of divalent metal ions. *FEBS Lett* **411**, 19-26 (1997).

APPENDIX

Appendix 5.1 Scalepack output file of Angiostatin/VEK-30 $P6_122$

Shell Lower Limit	Upper limit	I/Sigma in resolution shells: % of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	4.95	0.9	1.4	1.6	1.9	2.1	4.7	19.4	66.4	85.8
4.95	3.93	0.6	1.0	1.2	1.4	1.6	3.0	14.0	80.4	94.4
3.93	3.44	0.9	1.0	1.3	1.6	2.4	4.8	18.2	79.4	97.6
3.44	3.12	0.5	1.0	1.8	2.2	3.5	8.4	23.1	76.9	99.9
3.12	2.90	0.7	1.2	2.2	3.1	6.8	17.4	38.0	62.0	99.9
2.90	2.73	1.1	2.8	4.6	6.6	11.8	26.1	50.6	49.4	100.0
2.73	2.59	2.0	5.1	8.8	13.2	23.6	45.5	73.2	26.7	99.9
2.59	2.48	1.6	9.2	19.7	30.2	44.8	66.8	84.7	13.4	98.1
2.48	2.38	3.2	13.2	26.7	38.3	53.5	70.4	82.1	5.6	87.7
2.38	2.30	5.2	23.1	40.9	51.7	64.4	73.6	79.4	2.5	81.9
All hkl		1.6	5.7	10.5	14.5	20.7	31.1	47.2	47.2	94.4

Summary of reflections intensities and R-factors by shells

R linear = SUM (ABS(I - <I>)) / SUM (I)

R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)

Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))

In all sums single measurements are excluded

Shell Lower limit	Upper Angstrom	Average I	error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
50.00	4.95	6673.7	308.1	224.6	1.169	0.068	0.082
4.95	3.93	7495.9	274.7	228.8	1.908	0.079	0.095
3.93	3.44	5756.9	203.1	167.8	2.287	0.095	0.124
3.44	3.12	2377.0	66.7	57.2	2.294	0.108	0.144
3.12	2.90	1197.1	39.3	35.3	1.921	0.117	0.137
2.90	2.73	745.7	28.8	28.7	2.217	0.141	0.201
2.73	2.59	526.6	32.9	32.8	2.100	0.203	0.548
2.59	2.48	339.8	36.9	36.9	1.220	0.161	0.145
2.48	2.38	239.3	39.9	39.9	0.925	0.198	0.175
2.38	2.30	161.9	45.3	45.3	0.693	0.247	0.201
All reflections		2648.7	110.3	91.7	1.860	0.094	0.112

Appendix 5.2 Scalepack output file of Angiostatin/VEK-30 $P6_1$

Shell I/Sigma in resolution shells:

Lower limit	Upper limit	% of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	5.81	0.2	0.4	0.7	1.1	2.4	7.1	43.8	51.0	94.8
5.81	4.62	0.8	1.4	1.9	2.5	4.1	11.4	40.7	57.4	98.1
4.62	4.03	1.1	1.7	3.1	4.6	9.4	25.0	64.3	34.3	98.6
4.03	3.66	1.1	2.7	5.5	9.8	21.6	48.0	83.3	15.0	98.3
3.66	3.40	1.4	5.5	13.5	24.1	41.8	72.6	95.7	3.2	98.9
3.40	3.20	3.2	16.4	37.6	57.0	77.2	93.7	98.0	0.5	98.6
3.20	3.04	7.2	37.9	64.3	79.6	91.6	96.2	97.1	0.0	97.1
3.04	2.91	9.5	46.3	63.8	69.3	72.3	73.5	73.6	0.0	73.6
2.91	2.80	4.3	19.6	24.7	25.9	27.0	27.3	27.4	0.0	27.4
2.80	2.70	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
All hkl		2.9	13.1	21.4	27.2	34.6	45.3	62.3	16.3	78.5

Summary of reflections intensities and R-factors by shells

R linear = SUM (ABS(I - <I>)) / SUM (I)

R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)

Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))

In all sums single measurements are excluded

Shell limit	Lower Angstrom	Upper	Average I	Average error	Norm. stat.	Chi**2	Linear R-fac	Square R-fac
50.00	5.81		10401.5	514.7	347.8	1.087	0.065	0.077
5.81	4.62		5723.4	259.5	218.4	1.462	0.069	0.083
4.62	4.03		3866.7	220.2	195.8	1.220	0.081	0.093
4.03	3.66		2660.2	212.2	201.4	0.968	0.102	0.111
3.66	3.40		1532.3	206.0	201.8	0.577	0.129	0.139
3.40	3.20		671.6	198.0	198.0	0.329	0.211	0.210
3.20	3.04		346.9	208.4	208.4	0.254	0.320	0.323
3.04	2.91		197.3	224.6	224.6	0.205	0.408	0.388
2.91	2.80		167.5	227.9	227.9	0.208	0.445	0.372
All reflections			3159.9	255.1	224.4	0.837	0.084	0.083

Appendix 5.3 Scalepack output file of Ca^{2+}/con-G

Shell I/Sigma in resolution shells:

Lower limit	Upper limit	% of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
50.00	2.80	0.0	0.1	0.1	0.2	0.5	1.9	12.4	77.1	89.5
2.80	2.22	0.0	0.1	0.3	0.4	1.5	4.0	12.2	87.7	99.9
2.22	1.94	0.3	0.4	1.0	1.5	3.0	8.2	18.5	81.3	99.8
1.94	1.76	0.4	1.7	4.1	5.6	9.8	18.2	37.6	62.2	99.8
1.76	1.64	1.9	3.9	7.2	10.7	19.0	31.3	57.1	42.6	99.7
1.64	1.54	3.0	7.0	12.5	18.1	27.4	47.6	75.5	24.1	99.6
1.54	1.46	3.5	10.1	18.9	26.4	41.6	62.7	86.2	13.3	99.5
1.46	1.40	4.4	15.3	28.3	0.2	57.0	77.8	92.2	7.1	99.3
1.40	1.35	8.9	27.8	43.6	55.3	71.5	87.2	97.7	1.6	99.3
1.35	1.30	8.7	32.1	53.2	66.9	79.2	92.7	98.5	0.5	99.0
All hkl		3.1	9.8	16.8	22.4	30.9	42.9	58.5	40.0	98.5

Summary of reflections intensities and R-factors by shells

R linear = SUM (ABS(I - <I>)) / SUM (I)

R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)

Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))

In all sums single measurements are excluded

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
50.00	2.80		84068.8	3227.9	1681.8	1.916	0.062	0.071
2.80	2.22		41083.1	1074.5	435.1	1.548	0.057	0.064
2.22	1.94		20749.8	504.1	285.4	1.764	0.063	0.069
1.94	1.76		7616.3	225.4	178.1	1.700	0.085	0.085
1.76	1.64		3596.7	149.8	140.0	1.494	0.111	0.110
1.64	1.54		2105.6	131.4	126.4	1.284	0.153	0.139
1.54	1.46		1387.0	125.3	122.5	1.129	0.206	0.180
1.46	1.40		866.5	119.3	118.0	0.779	0.274	0.207
1.40	1.35		493.6	117.1	116.7	0.756	0.444	0.383
1.35	1.30		360.9	117.5	117.2	0.654	0.535	0.423
All reflections			15851.8	563.1	323.6	1.285	0.071	0.069

Appendix 5.4 Scalepack output file of Ca^{2+}/con-T[K7γ]

Shell Lower limit	Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	3.45	0.7	1.5	2.3	3.4	5.1	8.4	18.2	77.3	95.5
3.45	2.74	0.5	2.0	3.7	5.3	7.7	15.0	31.0	69.0	100.0
2.74	2.39	2.1	5.5	8.6	10.5	15.7	25.7	48.8	51.1	99.9
2.39	2.17	2.6	5.1	8.1	11.4	16.0	30.7	58.3	41.7	100.0
2.17	2.02	2.6	7.0	13.2	18.5	28.3	47.5	79.7	20.3	100.0
2.02	1.90	3.6	11.4	19.6	27.1	42.2	67.2	94.0	6.0	100.0
1.90	1.80	6.9	18.0	29.9	40.7	55.8	81.2	98.3	1.6	99.9
1.80	1.72	11.7	28.0	44.6	57.9	75.7	94.2	99.7	0.3	100.0
1.72	1.66	12.7	33.9	52.9	66.6	82.4	97.2	99.8	0.1	99.9
1.66	1.60	14.8	42.2	63.3	77.1	90.0	99.0	100.0	0.0	100.0
All hkl		5.7	15.1	24.1	31.2	41.0	55.6	71.7	27.7	99.5

Summary of reflections intensities and R-factors by shells

R linear = SUM (ABS(I - <I>)) / SUM (I)

R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)

Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))

In all sums single measurements are excluded

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
40.00	3.45		21439.2	642.3	327.1	1.427	0.060	0.070
3.45	2.74		12805.1	386.4	234.9	1.152	0.069	0.074
2.74	2.39		5409.5	218.8	169.6	0.899	0.090	0.093
2.39	2.17		4343.7	209.4	173.0	0.782	0.103	0.106
2.17	2.02		2563.9	182.8	165.9	0.682	0.148	0.144
2.02	1.90		1430.5	165.5	158.4	0.599	0.245	0.249
1.90	1.80		918.1	152.8	148.9	0.567	0.361	0.292
1.80	1.72		477.9	141.8	140.4	0.523	0.638	0.615
1.72	1.66		368.1	143.1	142.2	0.547	0.895	0.856
1.66	1.60		275.6	146.2	145.6	0.565	0.000	0.000
All reflections			5180.1	242.9	182.1	0.775	0.099	0.079

Appendix 5.5 Scalepack output file of Cd^{2+}/Mg^{2+}/con-T[K7γ]

Shell Lower limit	Upper limit	I/Sigma in resolution shells: % of of reflections with I / Sigma less than 0	1	2	3	5	10	20	>20	total
40.00	3.26	0.5	1.0	1.3	1.8	3.5	6.1	13.7	83.9	97.5
3.26	2.59	0.4	0.4	1.1	1.5	3.3	6.6	14.7	85.1	99.8
2.59	2.26	0.6	1.5	2.8	3.4	5.2	8.9	16.4	83.4	99.8
2.26	2.05	0.4	1.3	2.7	4.0	6.0	10.9	22.5	77.5	100.0
2.05	1.90	0.6	1.1	2.7	4.4	7.0	13.3	30.0	70.0	100.0
1.90	1.79	2.3	4.8	6.5	8.8	11.3	18.5	36.9	63.1	100.0
1.79	1.70	1.8	5.3	7.6	9.2	13.5	22.1	46.3	53.7	100.0
1.70	1.63	0.8	2.6	5.9	8.5	14.0	24.0	52.8	47.2	100.0
1.63	1.57	1.9	4.3	7.6	10.7	14.8	31.3	58.8	41.2	100.0
1.57	1.51	0.4	4.5	8.6	11.5	20.0	37.9	67.5	32.5	100.0
1.51	1.46	1.2	5.1	9.9	16.4	29.1	53.0	82.2	17.8	100.0
1.46	1.42	2.4	10.9	19.6	25.5	39.0	63.0	91.7	8.3	100.0
1.42	1.39	3.9	14.1	26.8	35.5	48.2	75.2	94.7	5.3	100.0
1.39	1.35	5.6	20.6	37.5	47.6	61.7	84.1	95.6	2.8	98.4
1.35	1.32	5.4	23.4	39.9	48.4	65.7	83.3	92.1	0.6	92.7
1.32	1.29	3.5	24.0	37.6	48.8	63.6	77.6	81.3	0.4	81.7
1.29	1.27	4.9	22.2	34.6	43.6	53.4	60.1	62.5	0.2	62.7
1.27	1.24	2.8	12.4	19.8	24.4	30.5	34.6	34.8	0.0	34.8
1.24	1.22	0.2	1.4	2.4	4.8	6.2	8.2	8.5	0.0	8.5
1.22	1.20	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
All hkl		1.9	7.9	13.4	17.5	24.3	35.3	49.4	34.7	84.2

Summary of reflections intensities and R-factors by shells

R linear = SUM (ABS(I - <I>)) / SUM (I)

R square = SUM ((I - <I>) ** 2) / SUM (I ** 2)

Chi**2 = SUM ((I - <I>) ** 2) / (Error ** 2 * N / (N-1)))

In all sums single measurements are excluded

Shell limit	Lower Angstrom	Upper	Average I	Average error	Average stat.	Norm. Chi**2	Linear R-fac	Square R-fac
40.00	3.26		30150.7	716.3	525.2	1.596	0.056	0.069
3.26	2.59		17688.5	395.7	300.3	1.282	0.059	0.064
2.59	2.26		12812.1	278.2	220.4	1.040	0.059	0.067
2.26	2.05		9338.9	249.6	177.7	0.731	0.062	0.070
2.05	1.90		7049.7	204.4	150.4	0.699	0.064	0.066
1.90	1.79		4685.8	153.1	122.7	0.617	0.072	0.078
1.79	1.70		3458.2	126.2	105.0	0.590	0.081	0.084
1.70	1.63		2799.6	111.0	94.1	0.610	0.091	0.088
1.63	1.57		2148.1	96.9	84.9	0.589	0.105	0.108
1.57	1.51		1888.8	91.2	80.9	0.568	0.110	0.095
1.51	1.46		1563.5	95.0	86.2	0.542	0.108	0.069
1.46	1.42		1099.2	105.2	99.2	0.450	0.131	0.115
1.42	1.39		966.0	123.4	117.9	0.429	0.134	0.109
1.39	1.35		768.7	146.8	142.6	0.418	0.156	0.125
1.35	1.32		620.2	166.4	163.5	0.394	0.163	0.169
1.32	1.29		660.7	189.4	187.1	0.395	0.138	0.127
1.29	1.27		550.4	203.9	202.3	0.358	0.151	0.166
1.27	1.24		591.3	231.5	230.9	0.543	0.213	0.259
1.24	1.22		1072.6	283.5	282.0	0.666	0.206	0.224
All reflections			6170.2	210.3	173.1	0.752	0.067	0.068