The Mechanism of Action of the Scorpion toxin, Chlorotoxin, a Bioinformatics Approach

Sheikh Adnan

A thesis submitted in partial fulfilment of the requirements of
Sheffield Hallam University
for the degree of Masters of Philosophy

March 2017

ABSTRACT

Glioblastoma is a very invasive and metastatic cancer with low patient survival-rate and high resistance to current treatment. Matrix metalloproteinase 2 (MMP2) is upregulated in glioblastoma multiforme (GBM) and its high expression is directly correlated with invasion and metastasis of GBM cells. Annexin A2 is involved in angiogenesis, and it is overexpressed in GBM. MMP2 and annexin A2 are potential targets for effective glioma therapy. Chlorotoxin is a promising toxin peptide for glioma targeted therapy. Chlorotoxin decreased migration and metastasis of GBM cells, although mechanism of action of the toxin is unclear due to the complicated nature of GBM, slow and costly experiments and challenging techniques. The relationship of chlorotoxin binding to *both* MMP2 and annexin A2 is controversial and further work is needed to clarify a complex situation. Chlorotoxin-like short toxins are homologous peptides that may also have potential therapeutic effects for glioma.

In this study homology modelling was used to predict 3D models of chlorotoxin-like short toxins. Stereochemical quality estimation of 3D models was performed with QMEAN, Z-Score and Procheck server. ClusPro docking programme was used to dock 3D structures of chlorotoxin and chlorotoxin homologues with their putative binding partners MMP2 and annexin A2 to provide a first step to understand the mechanism of action of the toxin. Chlorotoxin shows binding interaction with both MMP2 and annexin A2. Chlorotoxin-like short toxins also showed stronger binding interaction with activated MMP2 than annexin A2. The positively charged surface of chlorotoxin is mainly involved in interaction with MMP2. All three chlorotoxin-like short toxins bind to a similar site on MMP2 but on separate sites on annexin A2 hence suggesting the significance of observed interaction between chlorotoxin and MMP2.

Table of Contents **page**

Word Count 20150

1 INTRODUCTION

1.1 Glioblastoma.

Glioma is a lethal malignant primary brain cancer (Bayat *et al.,* 2016). Glioma consists of several subtypes including astrocytoma, oligodendroglioma, ependymoma and glioblastoma multiforme (Stupp *et al*., 2014). Glioblastoma multiforme (GBM) is the highest grade and most malignant type of glioma (Blume *et al*., 2013; Lee *et al*., 2015). Current treatment includes surgical resection of cancerous tissues, where possible, as well as radiotherapy followed by chemotherapy with the DNA alkylating agent temozolomide (Yin *et al*., 2007; Ganau *et al.,* 2015). However, the median survival time is approximately 16 months with current treatment and there is a 5-year survival time of 5% (Ganau *et al.,* 2015). Surgery and chemotherapy are challenging due to the diffusive and metastatic nature of glioma (Kesari *et al*., 2011, Sun *et al.,* 2012). Moreover, many chemotherapeutic agents cannot cross the blood brain barrier (Blume *et al*., 2013; Ramirez *et al*., 2013).

Tumour cells secrete proteolytic enzymes such as matrix metalloproteinases (MMPs) which degrade components of the extracellular matrix (ECM) and then invade surrounding healthy brain tissues (Madsen *et al*., 2015). It is crucial to understand underlying factors of GBM metastasis for effective anticancer therapy and to increase patient survival-time (Yan *et al*., 2011). GBM metastasis is a very complicated process as GBM cells are very diffusive in nature and exhibit tumour heterogeneity (Soeda *et al*., 2015). Several proteases play important role in facilitating cancer cell adhesion, migration and metastasis (Wang *et al*., 2003). Adhesion of the GBM cells to the ECM occurs which provides traction for cancer cell mobility in a liquid filled 3D environment (Wang *et al*., 2003). Glioma cells and endothelial cells secrete MMPs including MMP2, MMP9 and MT1-MMP which play

an important role in the degradation of the ECM of brain tissue (Goldbrunner *et al*., 1999). Ion fluxes generated by K^+ and Cl^- ion channels remarkably alters morphology of GBM cells by forming membrane protrusions and changes the cell volume via loss of water and osmotic pressure, helping GBM cells to shrink and squeeze through the ECM and metastasize surrounding tissues (Soroceanu *et al*., 1999; Turner *et al*., 2014). Cell surface receptors such as integrins and the hyaluronate receptor CD44 are involved in GBM cell adhesion (Goldbrunner *et al*., 1999). GBM adhesion to the ECM occurs via integrin $\alpha_V\beta_3$ as well as $\alpha_V\beta_5$ and the process may have rearranged cytoskeleton via cytoplasmic mediators such as focal adhesion kinase (FAK) and Pyk2. These integrin facilitate cancer cell mobility and invasion (Rupp *et al*., 2008). Integrin $\alpha_V\beta_3$ can directly bind to MMP2 and increases enzyme concentration on GBM cell surface thus enabling tumour cells to regulate MMP2 function and metastasise surrounding brain tissues (Munaut *et al*., 2003; McFerrin *et al*., 2006). Abnormal angiogenesis is also associated with tumour metastasis (Das and Marsden, 2013). GBM is one of the most vascular-rich tumour containing blood vessels that are highly permeable, with abnormal endothelial wall, twisted with an irregular pattern (Yi *et al*., 2016). Annexin A2 is involved in angiogenesis and metastasis of GBM.

1.1.1 Matrix metalloproteinases

Matrix-metalloproteinases (MMPs) are zinc-dependent endopeptidases involved in the degradation of the extracellular matrix (Nagase and Woessner, 1999). The activity of MMPs is highly regulated under normal conditions because they are involved in many physiological processes such as tissue remodelling, angiogenesis, wound healing, embryogenesis, nerve growth and apoptosis (Rahme *et al*., 2015). In contrast, MMPs are overexpressed and unregulated in pathological conditions such as cancer and cardiovascular diseases (Nagase and Woessner, 1999). Tissue-

inhibitor of metalloproteinases (TIMPs) are natural inhibitors of MMPs that control their activity (Nakopoulou *et al*., 2003).

MMPs are secreted as latent zymogens (Strongin *et al*., 1995), In order to activate MMPs, the activation peptide is cleaved which exposes the catalytic site (Morgonuva *et al*., 1999; Hur *et al*., 2000). In the activation of MMP2, MT1-MMP (MMP14) forms a complex with TIMP2 on the cell surface (Gingras *et al.,* 2000). The C-terminal domain of proMMP2 binds to TIMP2, forming a MT1-MMP/TIMP2/proMMP2 complex. Activation is mediated by another free MT1-MMP which interacts with complexed MT1-MMP via its C-terminal domain. MT1-MMP can cleave its substrate but cannot activate proMMP2 without TIMP2 (Bernardo and Fridman, 2003).

1.1.2 High expression of MMP2, MMP9 and Annexin A2 in GBM

MMP2 and MMP9 are involved in the degradation of ECM components including collagen and other matrix substrates (Liotta *et al*., 1980; Yong *et al.,* 2001). Enzyme-linked immunosorbent assay and gelatin zymography showed overexpression of MMP2 in GBM compared to low-grade glioma (Sawaya *et al*., 1996; Choe *et al*., 2002). Furthermore, immunohistochemical technique showed the localisation of MMP2 and MMP9 in only glioma cells and its blood vessels (Sawaya *et al*., 1996; Ramachandran *et al*., 2017). MMP2 was highly expressed in GBM tissue samples as well as all transformed cell lines but not in healthy brain tissues (Deshane *et al*., 2003; Kargiotis *et al*., 2008; Hagemann *et al*., 2010; Rahme *et al*., 2015).

Molecular mechanism of GBM invasion and metastasis is very complex (Kleber *et al*. 2008). Overexpression of MMP2 regulated by various molecular cascades leads to the invasion of GBM cells in healthy brain tissues (Hagemann *et al*., 2010). Interleukin-6 (IL-6), IL-10 and Th2 cytokine activates signal transducer and activator of transcription 3 (STAT3), a transcription factor which facilitates glioma invasion and metastasis (Brantley *et al*., 2008; Vasquez-Dunddel, 2013). STAT3 regulates and -

promotes overexpression of MMP2 including proteins such as cyclin D1, a cell cycle progression regulator (Peng2016).

GBM cells use MMP2 to initiate cancer cell invasion and metastasis regulated by CD95/CD95L signalling pathway. CD95 ligand (CD95L) is a natural ligand for CD95 cell surface receptor of tumour necrosis receptor family (Peter *et al*., 2015). CD95L induces apoptosis by initiating death-inducing signalling pathways by recruiting pro-apoptotic factors including caspase-8 to CD95 (Kleber *et al*., 2008). However, in apoptotic resistant GBM cells, CD95L induce tumorigenesis mediated by MMP2 via activation of Src-PI3K signalling pathway.

Transcription of MMP2 and MMP9 is initiated by β-catenin translocation into the nucleus and the process is initially induced by CD95 ligand (CD95L). CD95L allows binding of the Src family kinase (SFK; nonreceptor tyrosine kinase) member Yes and p85 subunit of phosphatidylinositol 3-kinase (PI3K) to CD95 resulting in the activation of AKT, a serine-threonine protein kinase also known as protein kinase B (PKB). Active AKT results in phosphorylation and inhibition of glycogen synthase kinase 3-β (GSK3-β), initiating translocation of β-catenin into the nucleus hence subsequent expression of MMPs particularly MMP2 (Kleber *et al*., 2008). Furthermore, MMP2 upregulation and invasion of GBM was mediated by forkhead transcription factor (FoxM1B) as well as insulin-like growth factor binding protein-2 (IGFBP2) (Wang *et al*., 2003; Levitt *et al*., 2005). In human GBM specimen, FoxM1B was upregulated as well as elevated MMP2 expression was observed (Dai *et al*., 2007). In addition, FoxM1B transformed normal astrocytes to GBM via process that degraded phosphatase and tensin homologue (PTEN) and activated AKT (Dai *et al*., 2010). The low-density lipoprotein receptor-related protein 1 also elevated MMP2 and MMP9 in an extracellular signal-regulated kinase (ERK) dependent pathway (Song *et al*., 2009). This suggest role of MMP2 in signal transduction and glioma proliferation and invasion, hence targeting MMP2 is a potential therapeutic strategy.

A recent study showed that glioma influence microglia activation and initiates microglial infiltration which induces activation of the oncogenes and production of cytokines. MT1-MMP activated by this mechanism degrades the ECM and facilitates tumour progression (Resende *et al*., 2016). Interleukin-1 (IL-1) production also activates STAT3 as well as MMP2 upregulation in glioblastoma which leads to invasion and metastasis (Tarassishin *et al*.,2014). Angiogenesis is promoted by IL-1 production in glioma cells as IL-1 induced glioma secretosome that contains factors which leads to vascular angiogenesis (Tarassishin *et al*., 2014). Tumour associated endothelial cells mediate secretion of MMP2 thus contributing to aggressive proliferation of GBM cells and establishment of a hostile microenvironment (Soda *et al*., 2011). Annexin A2 (AnxA2), a Ca^{2+} dependent, phospholipid-binding protein, is localised at the surface of endothelial cells as well as glioma cells (Tatenhorst *et al*., 2006). AnxA2 is overexpressed in many cancers including glioma, pancreatic cancer, colorectal and breast cancer (Kesavan *et al*., 2010). AnxA2 is involved in angiogenesis and more highly expressed in high-grade glioma than in low-grade glioma (Onishi *et al*., 2015). AnxA2 is involved in migration of the vascular endothelial cells and formation of the abnormal tumour vasculature in high-grade glioma (Kesavan *et al*., 2010). Recently, molecular targeted therapy is being considered as an effective anti-cancer therapy (Wang *et al*., 2015).

1.1.2 Biochemical studies and structure prediction, based on amino acid sequencing and cDNA cloning.

Full-length human MMP2 (gelatinase A) is a 72kDa enzyme involved in the degradation of collagens and gelatins (Morgunova *et al*., 1999). An active intermediate of the MMP2 is a 64 kDa enzyme which is subsequently cleaved to form the stable 62kDa active peptide (Lafleur *et al*., 2001). The full-length MMP2

comprises 660 amino acids (AAs) including a signal peptide (residues 1-29). The activation peptide of MMP2 (residues 30-109) contains the cysteine-switch motif *PRCGVPD* (Fig 1) which is important for keeping MMP2 in its inactive form. The catalytic region (residues 110-465) of MMP2 encompasses a catalytic domain and three repeats of a fibronectin type-II (Fib-II) domain, inserted into the catalytic domain. The catalytic domain of MMP2 has one catalytic Zn^{2+} ion and one structural Zn^{2+} ion as well as 3 Ca^{2+} ions. The Fib-II domains are important invcollagen, elastin and laminin binding as well as gelatin (Ghermann *et al*., 2002). The Fib-II domain 1 consists of 63 AAs, the Fib-II domain 2 consists of 60 AAs and the Fib-II domain 3 consists of 72 AAs. The hemopexin-like domain (residues 466-660) is linked to the catalytic domain via a hinge peptide (Fig1). The hemopexin-like domain consists of four subunits and its subunits 1 and 4 are connected via a disulphide bond (Fig 1).

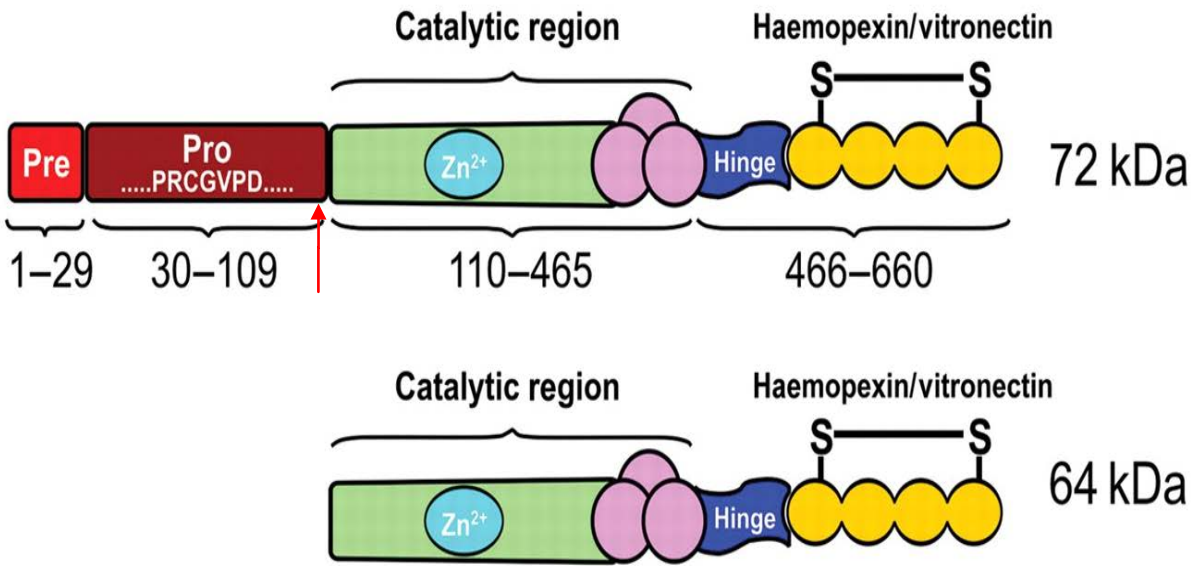

Figure 1. Domain structure of the MMP2. The signal peptide is indicated as 'Pre', activation peptide as 'Pro' which contains the cysteine-switch motif *PRCGVPD*. The catalytic domain contains Zn^{2+} ions (coloured turquoise). The Fib-II domains are represented by pink circles included in the catalytic region. The hemopexin-like domain (yellow) linked to the catalytic domain via a hinge peptide (dark blue). The 72kDa enzyme represents full-length inactivated MMP2 whereas 64kDa represents intermediate active MMP2. Red arrow indicates cleavage site to form 64kD MMP2 (Kandasamy *et al*., 2010).

1.1.4. 3D structures of MMP2 and its various domains.

A number of 3-dimensional (3D) structures of MMP2 have been determined experimentally by X-ray diffraction and nuclear magnetic resonance (NMR). These structures include full-length MMP2 (PDB code: 1CK7), the catalytic domain of MMP2 (PDB code 1QIB), first Fib-II domain (PDB code: 1KS0), second Fib-II domain (PDB code: 1CXW), third Fib-II domain (PDB code: 1J7M) and the hemopexin-like domain (PDB code: 1RTG).

1.1.4.1 Full-length human MMP2 (Gelatinase A) chain A.

The full-length human MMP2 3D structure was experimentally determined by X-ray diffraction to 2.8Å resolution by Morgunova *et al*. (1999) (PDB code: 1CK7) and consists of 631 residues corresponding the full-length sequence of the protein excluding the signal peptide (residues 1-29).

The protein from which the structure of proMMP2 was derived has a Glu404Ala mutation (Morgunova *et al*., 1999). Glu404 is involved in the catalytic activity of MMP2 (Hangauer *et al.,* 1984) and was replaced with alanine to allow crystallisation of MMP2 and provide stability against autoproteolysis. This mutation did not affect the architecture of the catalytic site of protein (Morgunova *et al.,* 1999).

The activation peptide of 1CK7 consists of 3 α-helices as well as short loops. A specific cleavage site for proteases, called the ‘bait-region’, is present at the connecting loop between helix 1 and helix 2 of the activation peptide (Fig 1.1). The activation peptide has an extended loop after helix-3 which contains the conserved cysteine-switch motif, interacting with the active zinc site of the catalytic domain (Fig 1.1).

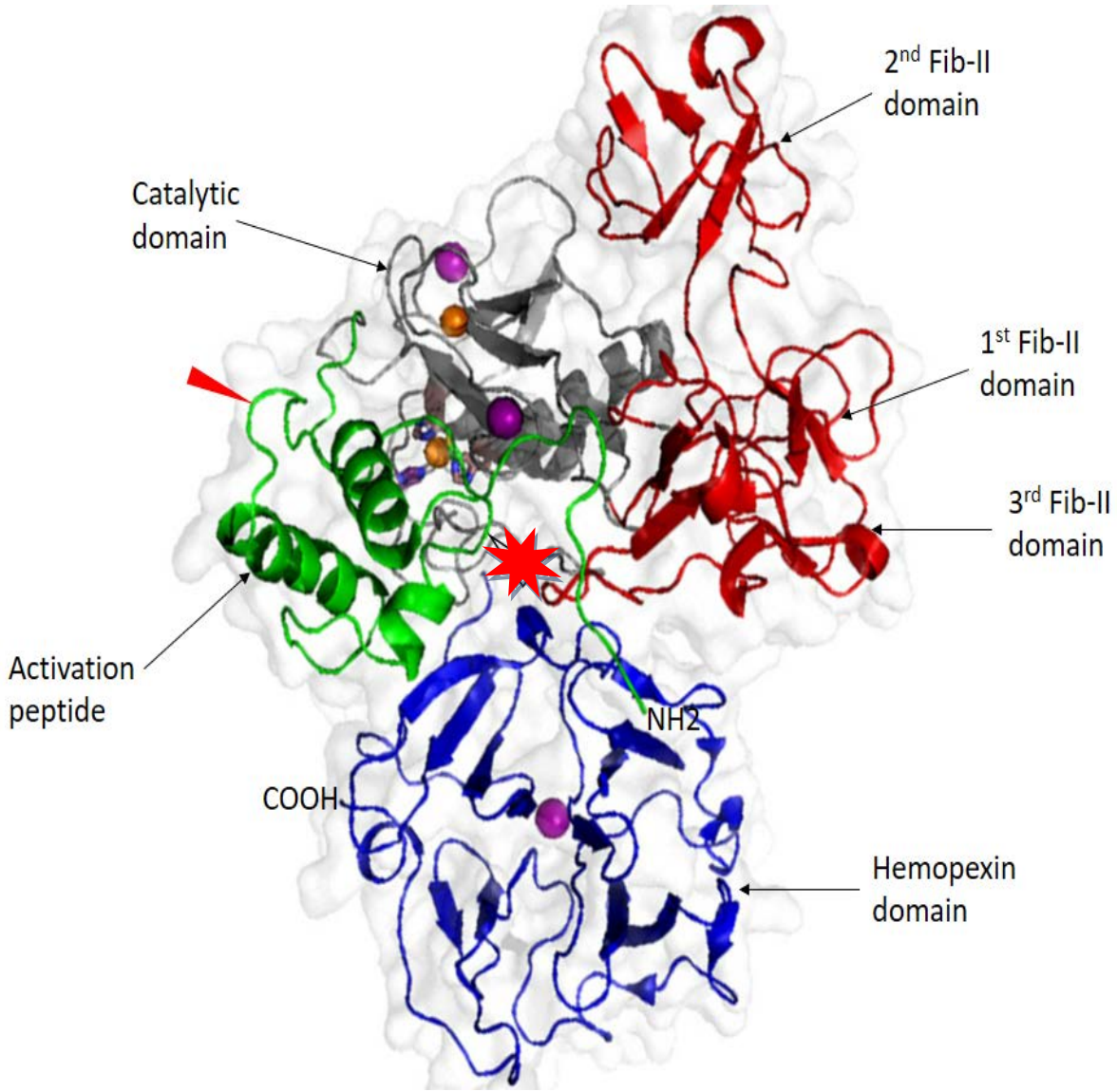

Figure 1.1 3D structure of proMMP2 (72kDa). The MMP2 (PDB code: 1CK7) activation peptide (green) contains 3 α-helices and connecting loops as shown in Fig 1.1. The protease-susceptible bait-region and cysteine-switch is represented by red asterisk and red triangle respectively. The catalytic domain (grey), three Fib-II domains (red) as well as hemopexin domain (blue) are shown. The Zn^{2+} ions are orange spheres and Ca^{2+} ions are purple (Sela-Passwell *et al*., 2009).

1.1.4.2 3D structures of various regions of MMP2.

The 3D structures of various regions of MMP2 have already been determined by X-ray crystallography and NMR studies (Visse and Nagase, 2003). The crystal structure of the catalytic domain of MMP2 was experimentally determined by X-ray diffraction at 2.8 Å resolution by Dhanaraj *et al.,* (1999) (PDB code: 1QIB). The catalytic domain contains 3 Ca^{2+} ions as well as one structural Zn^{2+} ion and one catalytic Zn^{2+} ion where three histidine residues (H403, H407 and H413) and a water molecule which is also found in the crystal structure, all interact with the catalytic Zn^{2+} ion (Dhanaraj *et al*., 1999).

The 1QIB protein was expressed in *E. coli* and the amino acid sequence of the fibronectin domains was deleted during expression and crystallisation of 1QIB (Fig 1.2a). The 3D structure of 1QIB consists of 3 α-helices, 5 stranded β-sheets and connecting loops. The catalytic domain has a specificity pocket S1' pocket located at the right side of the CAT-Zn represented by stick residues (Fig 1.2a) as well as a substrate binding-cleft formed by the β-sheet IV, helix-B followed by an extended peptide region (Fig 1.2a).

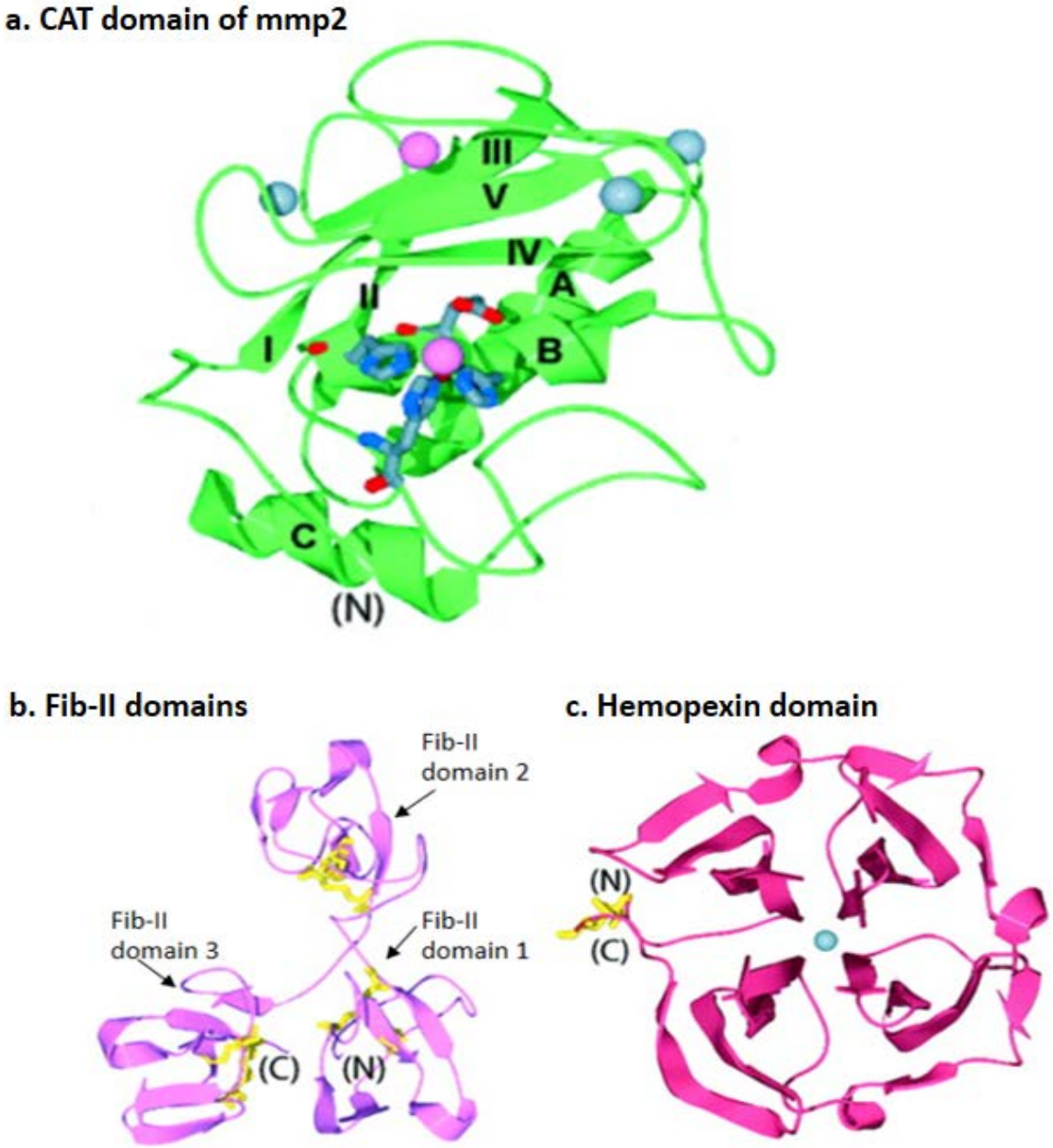

Fig 1.2. 3D structures of various regions of MMP2. (a). The 3D structure of the catalytic domain contains 3 α-helices (indicated as A to C) and 5 β-sheets (indicated as I-V). β-sheets and α -helices of 1QIB are arranged as 'NH2-I-A-II-III-IV-V-B-C-COOH'. The Zn^{2+} ions are shown as pink spheres and Ca^{2+}ions are shown as blue spheres (b) three Fib-II domains of the MMP2, (c) the hemopexin-like domain of MMP2 consists of 4 propeller-like blades of β-sheets with a disulphide present between blade I and IV indicated with yellow lines (Visse and Nagase, 2003).

1.1.4.3 The Fibronectin Domains of the MMP2

MMP2 consists of an array of three Fib-II domains which bind to gelatin, elastin, collagen type I and IV and laminin (Ghermann *et al.,* 2002). The 3D structures of the Fib-II domains 1, 2 and 3 were determined by NMR spectroscopy (Briknarova *et al.,* 1999 and Briknarova *et al*., 2001; Ghermann *et al.,* 2002). The 3D structure of third Fib-II domain (PDB code: 1J7M) has a E11G mutation verified by DNA sequencing. NMR spectroscopy showed that the codon for glutamic acid at position 11 (GAA) in the 1J7M expression plasmid mutated to GGA (Gly amino acid) during plasmid propagation. This mutation did not affect overall 3D structure (Briknarova *et al.,* 2001).

The Fib-II domains are inserted between the 5^{th} β-sheet and 2^{nd} α-helix of the CAT domain. Each Fib-II domain consists of a pair of β-sheets and a small helix stabilised by two disulphide bonds that is indicated by yellow sticks (Fig 1.2b). The hydrophobic pockets formed by the β-sheets of the fibronectin domains are involved in substrate (gelatin) binding hence determining substrate specificity (Murphy and Knauper, 1997; Ghermann *et al*., 2002).

1.1.4.4 The C-terminal domain of MMP2.

The 3D crystal structure of the C-terminal domain (hemopexin-like domain) of MMP2 was determined by X-ray diffraction at 2.6 Å resolution by Gohlke *et al.,* (1996) (PDB code: 1RTG). Amino acid residues start from 451 to 660 with 33% coverage of gelatinase A. There are 2 Ca^{+} co-factors and 1 Cl^{-} ion bound to the structure. The structure (1RTG) consists of 4 bladed propeller-like discs as shown in Fig 1.2c, where blade I and IV are connected via a disulphide bond (Fig 1.2c).

1.1.5 Activation of proMMP2, the Cysteine Switch.

MMP2 is secreted as an inactive zymogen (Nagase and Woessner*.,* 1999). An extended loop after helix-3 of the activation peptide contains a conserved cysteine. The sulfhydryl group of this cysteine co-ordinates with the zinc ion of the CAT domain keeping MMP2 in its zymogen form (Bescond *et al.,* 1999).

Inactive proMMP2 (72kDa) is activated when the proteinase-sensitive 'bait region' between helix-1 and helix-2 of the activation peptide is cleaved by proteases such as plasmin, which disrupts the Cys-Zn coordination (Fig 1.3). Though only a fragment of the activation peptide is cleaved, the rest of the activation peptide is cleaved by the MMP intermediate or by other active MMPs which activates MMP2. Finally, a fully activated MMP2 (62kDa) is generated (Fig 1.3).

ProMMP2 can be activated by thiol modifying agents such as 4-aminophenylmercuric acetate (APMA), N-ethylmaleimide (NEM) as well as sodium dodecyl sulphate (SDS), oxidised glutathione (GSSG) and hypochlorous acid (HOCl) (Chakraborti *et al*., 2003; Nagase *et al*., 2006). The thiol-modifying agents disrupt the Cys-Zn interaction by modifying the sulfhdryl (SX) group exposing the catalytic site (Fig 1.3). The initial cleavage occurs within the activation peptide region and this reaction is intramolecular; this is followed by removal of the rest of the activation peptide, due to the intermolecular reaction of intermediate compounds (Nagase *et al*., 2006).

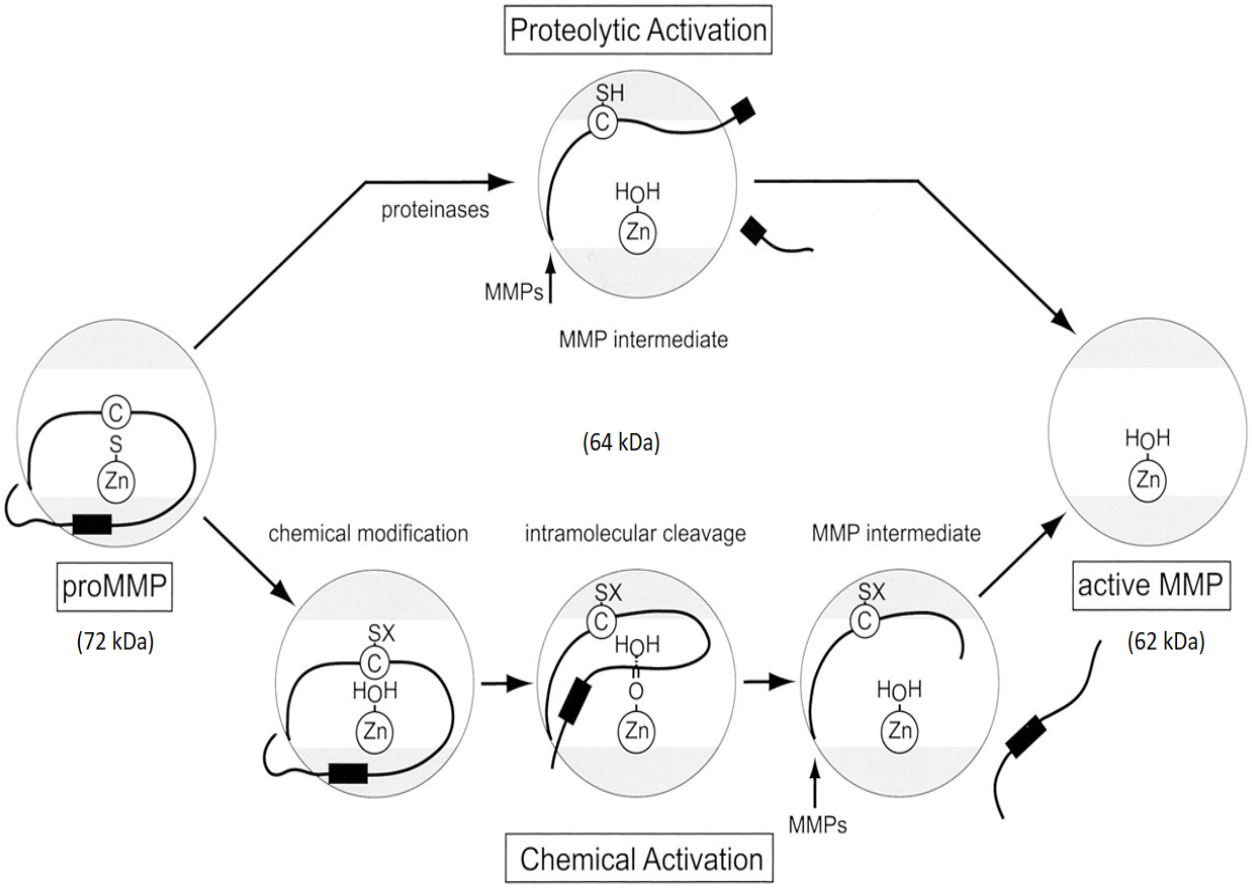

Figure 1.3. Cysteine switch mechanism. ProMMP2 can be activated chemically (bottom pathway) as well as by proteinases (top pathway). Black rectangle represents 'bait region' and 'C' represents the cysteine switch in propeptide. Sulfhydryl of cysteine is represented by 'SH' in top pathway and 'SX' in bottom pathway. Thiol modifying agents such as AMPA and NEM as well as proteases such as plasmin, disrupt the Cys-Zn coordination exposing the catalytic site. MMP2 is fully activated after removal of the activation peptide by intermediate compounds (Visse and Nagase, 2003).

MMP2 can be activated biologically as well as chemically. MT1-MMP (MMP14) complexed with TIMP2 leads to activation of proMMP2 (Cao *et al.,* 1995). The complex acts as a receptor for proMMP2; the N-terminal domain of TIMP2 binds to MT1-MMP on the cell surface, allowing the C-terminal domain to interact with proMMP2, then this is activated by another free MT1-MMP. Another free MT1-MMP interacts with the MT1-MMP, TIMP2, ProMMP2 complex and activates proMMP2 (Fig 1.4) (Bernardo and Fridman, 2003).

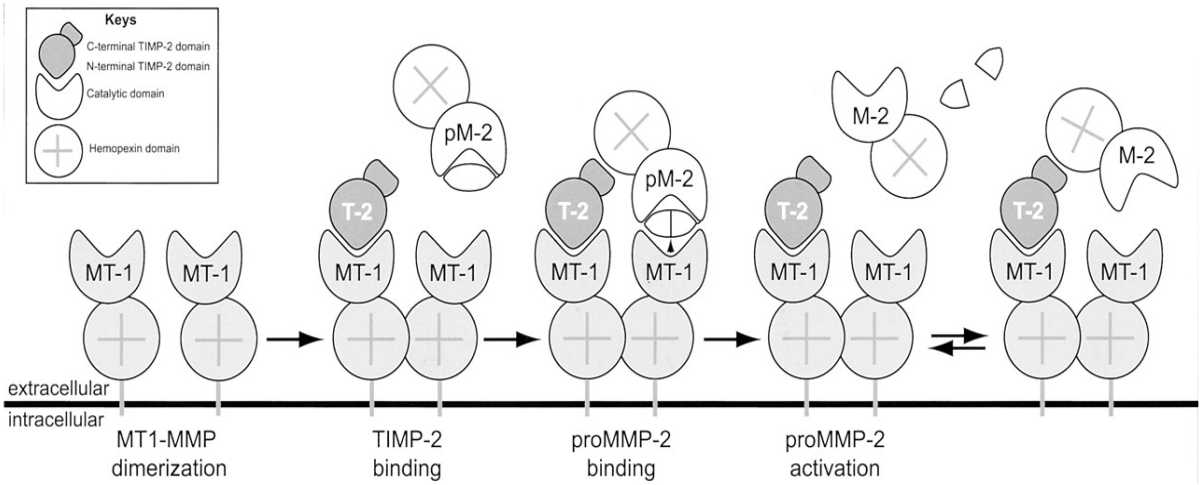

Figure 1.4. Activation of proMMP2 by MT1-MMP/TIMP2 complex. An active MT1-MMP (MT-1) complexed with TIMP2 on cell surface serves as a receptor for proMMP2 (pM-2). The hemopexin domains of two MT-1 (MMP14) on cell surface can interact to form a dimer. The C-terminal domain of pM-2 binds to MT-1/T-2 complex and the activation peptide binds to free MT-1 activating pM-2. MT-1 can cleave its substrate but cannot activate pM-2 without T2 (Visse and Nagase, 2003).

1.2 ANNEXIN A2

1.2.1 Annexin A2 in GBM

Annexin A2 (AnxA2) is a calcium-binding cytoskeletal protein which is localised at the surface of endothelial cells as well as glioma cells (Tatenhorst *et al*., 2006). It is more highly expressed in high-grade GBM than low-grade (Beckner *et al*., 2005; Zhai *et al*., 2011). GBM over-expression of AnxA2 is correlated with an increase in metastasis and abnormal angiogenesis of GBM cells by promoting degradation of the ECM, vascular endothelial cell migration and vascular remodelling (Zhao *et al*., 2010). AnxA2 mediates glioma cell adhesion to endothelial cells and promotes production of angiogenic factors such as vascular endothelial growth factor (VEGF) and platelet-derived growth factor (PDGF) thus promotes neovascularization and invasion (Maule *et al*., 2016). In glioblastoma specimen, multiple satellites of tumour cells containing dilated blood vessels were found in brain parenchyma near brain tumour centre (Onishi *et al*., 2015). Immunohistochemical and qRT-PCR studies showed high expression of AnxA2, VEGF and PDGF in glioblastoma cells. Moreover, staining intensity of VEGF and PDGF positively correlated with AnxA2 overexpression hence suggesting AnxA2 role in regulating angiogenesis (Onishi *et al*., 2015).

Moreover, surface biotinylation assay showed elevated AnxA2 expression on GBM cell surfaces and tumour associated endothelial cells (Kesavan *et al*., 2010). Reduced migration of GBM cells has been shown in an AnxA2 knockout mouse model (Tatenhorst *et al*., 2006). AnxA2 silencing by small-interfering RNA (siRNA) in human glioblastoma cell lines (U87MG and U373MG) also resulted in reduced migration (Zhao *et al*., 2011). AnxA2 regulates expression of VEGF through protein kinase C (PKC) signalling pathway and a positive feedback mechanism (Zhao *et al*.,

2011). In AnxA2 depleted glioblastoma cells, expression of proangiogenic factors was inhibited (Bao *et al*., 2009).

AnxA2 mediates the activation of a serine protease enzyme called plasmin (Sharma *et al*., 2007). Plasminogen activators such as tissue plasminogen activator (tPA) and urokinase plasminogen activator (uPA) convert the inactive plasminogen to active plasmin (Diaz *et al*., 2004). AnxA2 complexed with p11 serves as a co-receptor for tissue plasminogen activator (tPA) and plasminogen on the cell surface which mediates the conversion of the inactive plasminogen to plasmin (Valapala *et al*., 2011). This process also induces endothelial cell and tumour cell mediated release of growth factors such as basic fibroblast growth factor (bFGF) and PDGF (Maule *et al*., 2016). Plasmin is involved in angiogenesis by fibrinolysis, degradation of the ECM components such as laminin, proteoglycans and fibronectin as well as activation of the proMMP2 (Tatenhorst *et al*., 2006; Lokman *et al*., 2011). Plasmin promotes angiogenesis by binding to integrin αVβ3 and facilitating endothelial cell migration (Tarui *et al*., 2002). This process is facilitated by high expression of (bFGF) as well as attracting human peripheral monocytes and degrading platelet aggregation (Tarui *et al*., 2002). All these pathological processes are important for AnxA2 mediated angiogenesis and subsequent invasion and metastasis of GBM cells. This suggests that targeting AnxA2 would be an effective anti-cancer therapy and may improve patient survival (Ichikawa *et al.,* 2014).

1.2.1 Structure of Annexin A2.

The 3D structure of AnxA2 was determined from X-ray diffraction studies by Shao *et al*., (2006). AnxA2 consists of a N-terminal ‘head-domain’ of ~33 amino acids and a highly conserved core motif of 31-338 amino acids, making a curve-shaped protein. The core motif of the carboxyl domain consists of four homologous repeats also called annexin repeats (I-IV); each repeat encompasses 70-80 residues, arranged in

a near-parallel shape (Fig 1.5b). All homologous repeats have a similar structure, each repeat consists of 5 α-helices and connective loops (Fig 1.5b). The N-terminal domain is involved in the tPA binding site whereas, the C-terminal domain is involved in calcium binding as well as phospholipid binding and F-actin binding (Jones *et al*., 1992; Rety *et al*., 1999).

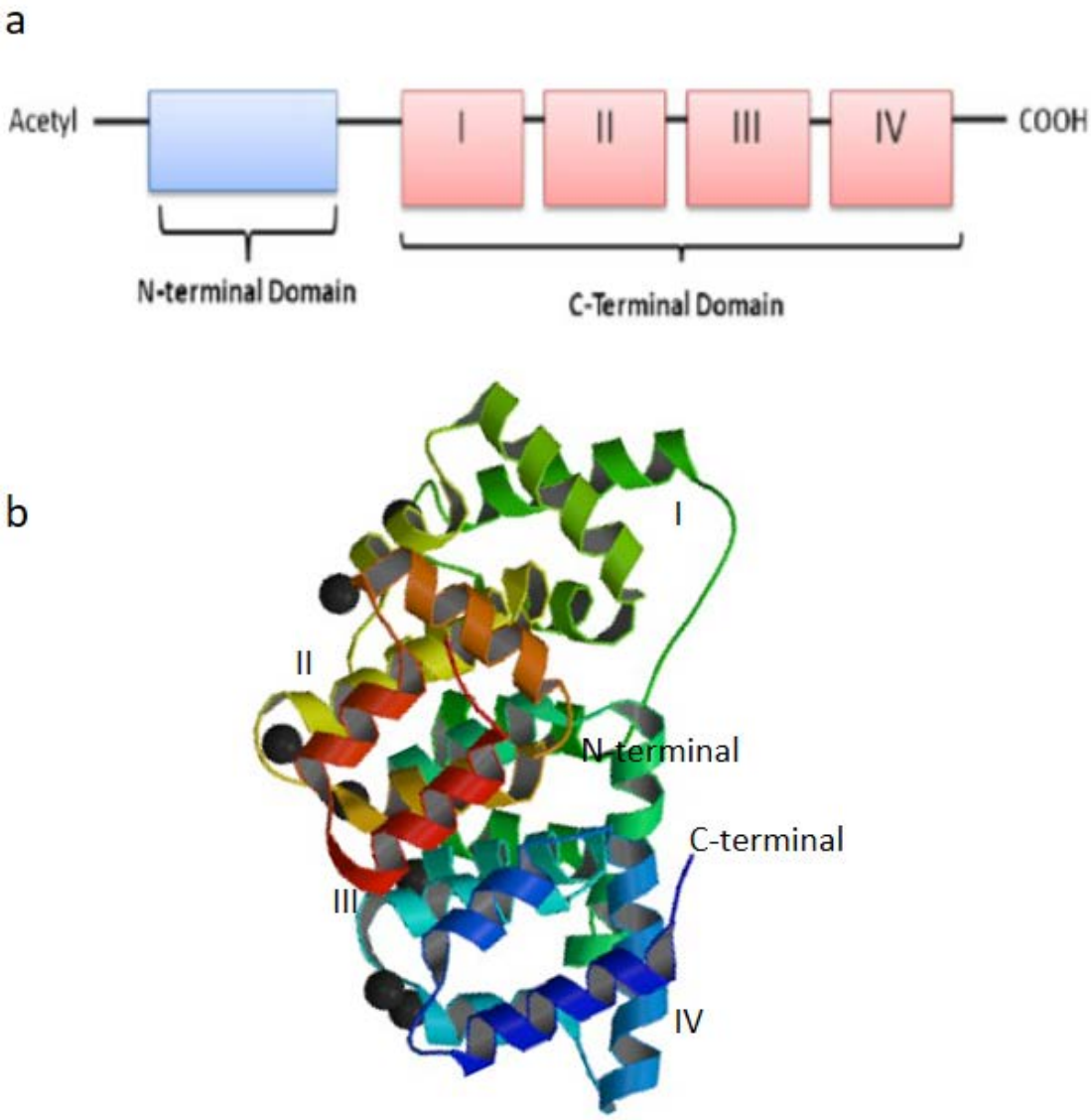

Fig 1.5. Structure of AnxA2 protein. (a) Domain structure of AnxA2 where C-terminal core region is represented by pink cubes and N-terminal 'head domain' by blue rectangle. The NH_2 terminus of AnxA2 is acetylated (b) 3D structure of AnxA2 (PDB code: 2HYW), four homologous annexin repeats are labelled and each repeat consists of 5 α-helices linked to each other via connective loops. Ca^{2+} ions are represented by black spheres (Lokman *et al*., 2011; RCSB PDB databank).

MMP2 and AnxA2 are potential candidates for molecular targeted therapy for glioma because they are involved in GBM invasion, adhesion, angiogenesis and metastasis (Kling *et al*., 2016). GBM is resistant to surgery and chemotherapy due to its diffuse nature (Ramirez *et al*., 2013). Studies have shown that even low-grade glioma invade healthy brain tissues at a fast rate; it is therefore crucial to target these proteins to inhibit glioma metastasis and improve effectiveness of glioma treatment. A number of studies are under investigation to target MMP2 and AnxA2 thus inhibit GBM metastasis and increase survival-rate (Kesavan *et al*., 2010; Alam *et al*., 2010; Badiga *et al*., 2011; Ali *et al.*, 2016).

1.3 Chlorotoxin

Chlorotoxin (Cltx) is a 36 amino acid peptide derived from venom of the scorpion (*L. quinquestriatus*) (DeBin *et al*., 1993). Cltx specifically binds to GBM and inhibits its invasion and metastasis without having an effect on healthy brain tissues (Soroceanu *et al*., 1998; Kovar *et al*., 2013; Butte *et al*., 2014). Cltx also binds to tumours of neuroectodermal origin such as small-cell lung carcinoma, melanoma, suggesting that it might be used to treat a range of cancers (Lyon *et al.,* 2002). Cltx is in phase-III clinical trials for glioma treatment and is being used for diagnostic and imaging of glioma (Veiseh *et al*., 2007; Zhang *et al*., 2016).

The 3D structure of Cltx was experimentally determined by NMR spectroscopy (Lippens *et al*., 1995). It has a compact structure comprised of 3 short anti-parallel β-sheets tightly packed around an α-helix (Fig 1.6). The compact structure of Cltx is stabilised by four disulphide bonds linking the cysteine residues.

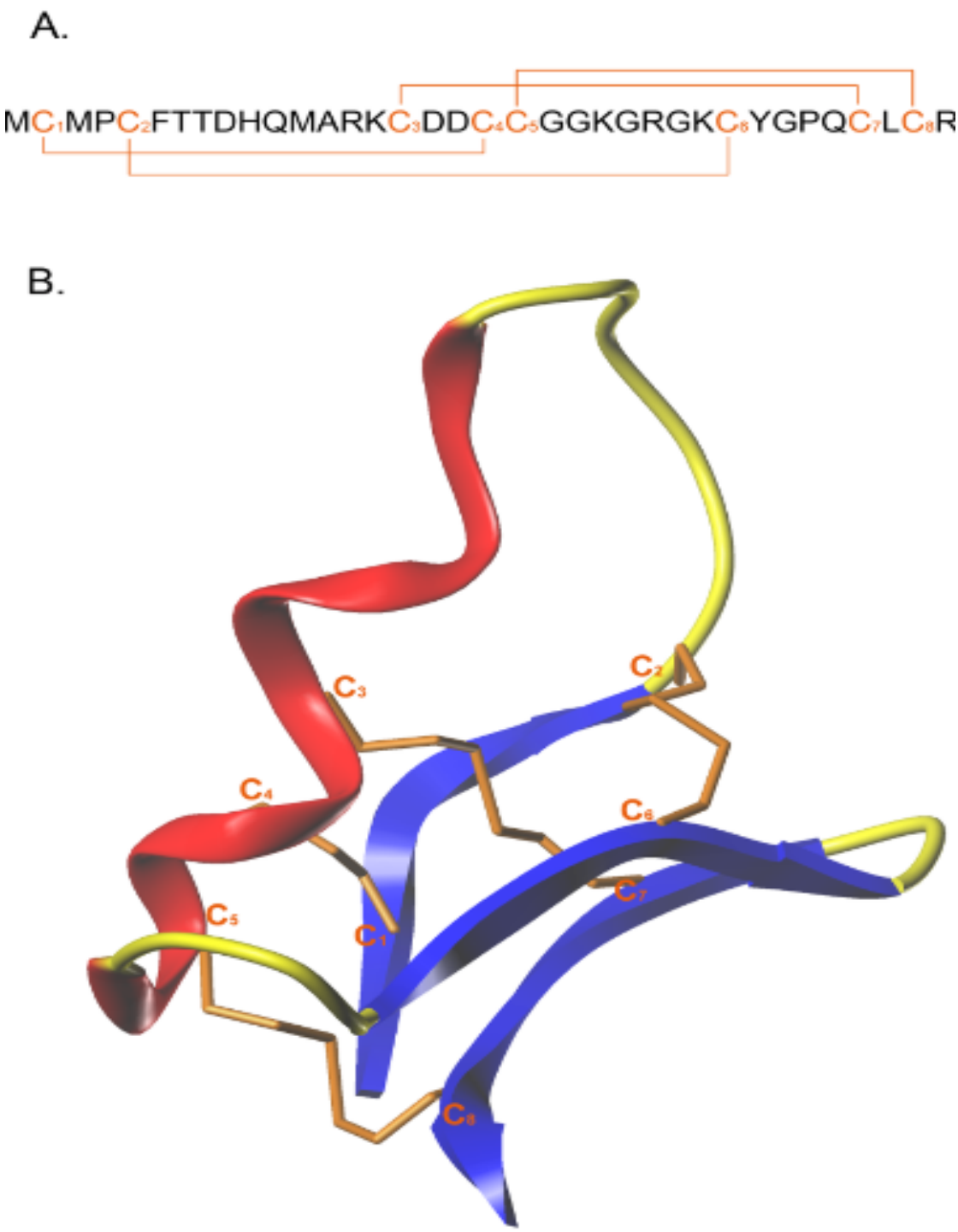

Figure 1.6. Primary structure and 3D structure of Cltx. (A) Cltx consists of 36 amino acids, the sequence contains 8 highly conserved cysteine residues (B) Ribbon structure of Cltx (PDB code: 1CHL), peptide consists of an alpha helix and 3 anti-parallel β-sheets connected with loops. Four disulphide bridges are present between highly conserved cysteines. β-sheets are coloured in blue and α-helix is coloured in red; disulphide bridges of the toxin are coloured in orange (Dardevet *et al*., 2015).

At first, chlorotoxin was believed to inhibit chloride fluxes by blocking small-conductance chloride channels in epithelial cells (DeBin *et al*., 1993). Cltx was shown to bind selectively to glioma specific Cl^- channels in acute slices of human gliomas (Ullrich *et al*., 1998). Later, studies showed that chlorotoxin specifically interacts with a lipid raft-anchored complex that contains MMP2, MT1-MMP, TIMP2 and chloride channels (ClC3) facilitating internalisation of the cell surface complex thus reducing glioma metastasis (Kesavan *et al*., 2010).

Fortunately, Cltx is accessible to glioma cells as it can cross well defined blood brain barrier (BBB) or at least leaky blood brain tumour barrier (BBTB). Cltx was shown to diffuse deeply in glioma whereas specific antibodies were not able to cross BBB (Mamelak *et al*., 2007; Veiseh *et al*., 2009). Also, Cltx conjugated fluorescent dye Cy5.5 (Cy5.5-Cltx) labelled glioma cells without affecting toxin transport across the BBB in a transgenic mice model (Veiseh *et al*., 2007).

Cltx specifically inhibited MMP2 activity but did not interact with either MMP1 and MMP9 which are expressed on the surface of glioma cells (Deshane *et al*., 2003). Cltx inhibited MMP2 activity via fluorescence resonance energy transfer (FRET) substrate assay and gelatin zymography (Alam *et al*., 2012). Inhibition of MMP2 activity resulted in the prevention of glioma cell shrinkage and release from the ECM. Thus, reduced migration of GBM cells occurred through the extracellular space of brain tissue (McFerrin and Sontheimer 2006). Moreover, Cltx demonstrated a high affinity binding site and a low binding affinity site for radio-iodinated Cltx (^{125}I-Cltx) suggesting more than one receptor for Cltx on the glioma cell surface (Soroceanu *et al*., 1998).

A synthetic version of Cltx has been synthesized under the name of TM601 (Kesavan *et al*., 2010). TM601 showed both anti-tumour activity against glioma and anti-angiogenic activity against human umbilical vein endothelial cell (HUVECs). TM601 is being used in phase-III clinical trials in the US. Radio-labelled TM601

(radioactive iodine ^{131}I conjugated at Tyr^{29}) is being used for diagnostic imaging and radiotherapy of glioma. Cltx-Cy5.5 is an infrared dye which is being used for imaging and surgical resection of glioma (Vieseh *et al*., 2007; Stroud *et al*., 2011).

Cltx-conjugated nanoparticles are under investigation for glioma imaging, therapy and drug delivery (Hsiao *et al*., 2015; Han *et al*., 2016; Chen *et al*., 2016; Tamborini *et al*., 2016).

To complicate the situation further, another target for Cltx was subsequently identified (Kesavan *et al*., 2010). Cltx was shown to have anti-angiogenic effects by binding to human umbilical vein endothelial cells (HUVECs) (Zhai *et al*., 2011). Cltx also binds to AnxA2 on glioma cells thus inhibiting migration and metastasis (Kesavan *et al*., 2010; Onishi *et al*., 2015). Small interfering RNA (SiRNA) knockout AnxA2 cells showed reduced Cltx binding as well as reduced migration (Kesavan *et al*., 2010). Cltx has a promising therapeutic target to glioma and related solid tumours. Although, the mechanism of action of Cltx is very complicated and far from clear at both cellular and molecular levels. The mechanism of chlorotoxin action needs to be investigated if clinical studies are to progress satisfactorily. The relationship of Cltx binding to *both* MMP2 and AnxA2 is controversial and further work is needed to clarify a complex situation. A number of chlorotoxin-like short-toxins (Cltx-like short toxins) have been isolated from venom of different scorpion species (Wudayagiri *et al*., 2005; Newton *et al*., 2007; Ali *et al*., 2016). Chlorotoxin homologues might have therapeutic value in glioma and range of cancers.

In-vitro experiments are costly and time consuming hence *in-silico* computational studies are being considered as a reliable alternative (Coelho *et al*., 2016; Kumar *et al*., 2015) used to study protein-protein interaction, enzyme inhibitor complex and drug design. *In-silico* computational studies are fast, reliable approach in selecting broad range of inhibitor of an enzyme, drug development, drug selectivity (Coelho *et al*., 2016).

1.4 Homology modelling

Determination of 3D structures of protein by X-ray crystallography and NMR spectroscopy are slow and time-consuming methods without guaranteed success especially for membrane proteins (Floudas *et al*., 2006). There are many more available sequences of proteins in databases than solved crystal structures of proteins and there is a constant increase in the generation of the sequences of proteins than solved 3D structures (Moult *et al*., 2014). Homology modelling is a reliable and accurate method used to generate 3D models of a protein from its amino acid sequence, using an existing crystal structure (template) for an evolutionarily related protein (target). The critical assessment of structure assessment (CASP) is world-wide study to test protein-prediction methods (Moult *et al*., 2014). Mainly, CASP experiments test and provide help to advance the methods of 3D model prediction of proteins via blind prediction (Moult *et al*., 2014). Homology modelling is a reliable procedure for protein prediction assessed by CASP. The Swiss-Model is a fully automated homology modelling web-server used to predict 3D structure of a protein from its amino acid sequence as PDB format output (Bordoli *et al*., 2009). Recently, Swiss-Model has been used extensively world-wide, predicting 0.9 million requests for protein models each year (Biasini *et al*., 2014).

1.5 Protein-protein docking

Protein-protein docking is performed to investigate the interaction between two proteins or between a drug and a protein (Coelho *et al*., 2016). Docking is performed to analyse binding affinities of a protein on its target molecule which is helpful in enzyme-inhibitor interaction, drug design etc. (Kozakov *et al*., 2013; Coelho *et al*., 2016).

It is possible to study the interaction of chlorotoxin and related Cltx-like short toxins with putative binding targets such as MMP2 and AnxA2 using docking studies. Study of their interaction, evaluate binding sites on targets, binding affinities and to select toxins that have higher binding affinities. X-ray crystallography is a slow and time-consuming method. To date only the 3D structures of Cltx and Insectotoxin I5A have been determined so far. Homology modelling has provided a fast, robust and reliable method to predict 3D models of Cltx-like short toxins from their sequences and then to evaluate quality assessment in order to find their usefulness in a subsequent docking study (Moult *et al*., 2014).

2 AIMS and OBJECTIVES

2.1 AIMS

The aims of the project are to study interaction of chlorotoxin and related toxins with its proposed putative binding partners MMP2 and annexin A2, using an *in-silico* approach in order to understand the mechanism of action of chlorotoxin.

2.2 OBJECTIVES

- To study existing 3D models of MMP2, perform quality assessments and then to dock chlorotoxin to MMP2 models using the ClusPro docking programme.

- To study which domains of MMP2 (the catalytic domain, fibronectin module type-II domains and the hemopexin domain) are important in binding.

- To study annexin A2 3D models, and to assess stereochemistry of models. To dock chlorotoxin to annexin A2 models and study their interaction.

- To use molecular modelling to predict 3D models of the chlorotoxin-like short-toxins and docking techniques to study the mechanism of action of chlorotoxin-like short-toxins with a view to the development of new clinical agents.

3 CHLOROTOXIN DOCKING with MMP2 and ANNEXIN A2

3.1 Introduction

Glioblastoma (GBM) is the most common metastatic primary brain cancer (Lee *et al*., 2015). MMP2 proteolytic enzyme is overexpressed in GBM and involved in its invasion and metastasis (Deshane *et al*., 2003). AnxA2 is a calcium-binding cytoskeletal protein highly expressed in GBM and involved in angiogenesis (Tatenhorst *et al*., 2006). Both MMP2 and AnxA2 are potential candidates for targeted therapy for GBM as discussed in previous chapter. A number of studies are under investigation to target MMP2 and AnxA2 thus inhibit GBM metastasis and increase survival-rate.

Chlorotoxin (Cltx) is a 4kD peptide derived from venom of *Leiurus quinquestriatus* scorpion (DeBin *et al*., 1993). It consists of 36 amino acids and its sequence is MCMPCFTTDHQMARKCDDCCGGKGRGKCYGPQCLCR (DeBin *et al*., 1993). It has a compact structure which is supported by four disulphide bridges which link all cysteine residues present in its sequence. It contains 3 small β-sheets packed around an α-helix (Lippens *et al*., 1995).

Cltx was originally shown to block small-conductance epithelial chloride channels when applied internally (DeBin *et al*., 1993). Later, studies suggested that Cltx binds to MMP2 and facilitates its internalisation, thus reducing MMP2 enzymatic activity on the surface (Deshane *et al*., 2003). To complicate matters further, annexin A2 (AnxA2) was proposed a potential target for Cltx (Kesavan *et al*., 2010). However, the mechanism of action of Cltx interaction with its proposed targets is far from clear.

In-silico homology modelling and docking is a reliable method to study molecular interactions (Vyas *et al*., 2012). Protein modelling and docking is a reliable and cost-effective method which is used to provide information on predicted interactions (Moult *et al.,* 2014). This type of *in-silico* approach is used to find interaction of Cltx with its putative targets MMP2 and AnxA2 and shed light on its mechanism of action.

This approach has been used before in the study of scorpion toxins. Charybdotoxin (ChTx) is a short scorpion venom peptide that blocks voltage-gated potassium channels (Froy *et al*., 1999). The 3D structure of ChTX was experimentally determined by NMR spectroscopy (Bontems *et al*., 1992). Khabiri *et al*., (2011) used *in-silico* modelling and docking in order to study the interaction of ChTX with the potassium channel Kv1.3. Kumar *et al*., (2015) investigated potential inhibitors of ChTX toxin via *in-silico* study. Recently, Farsani *et al*., 2015 modelled a Cltx homologue (Cltx-like peptide MeICT) and studied its interaction with the catalytic domain of MMP2 (1QIB structure) using an *in-silico* approach. The 3D structure of full-length MMP2 is present in mutant form (PDB code: 1CK7). Morgunova *et al*., (1999) introduced a mutation (E404A) in order to prevent the protein from auto-proteolysis during crystallisation.

In order to mimic the *in-vivo* scenario, 1CK7 structure was un-mutated to form a native MMP2 model. To represent Cltx interaction with an activated MMP2, the activation peptide of MMP2 was cleaved. Then, interaction of Cltx was studied with its putative binding partners, MMP2 and AnxA2.

3.2 Methodology.

3.2.1 3D protein structures and visualisation by RASMOL and Swiss Deep-View.

The amino acid sequences of Cltx, MMP2 and AnxA2 were extracted from the UniProt database (http://www.uniprot.org/). The 3D structures of Cltx (1CHL) and annexin A2 (2HYW) were obtained from RCSB PDB-databank (www.rcsb.org/) in PDB format. Various 3D structures of MMP2 and its individual domains were also downloaded from RCSB-PDB databank (www.rcsb.org/), full-length MMP2 (1CK7, Morgunova *et al*., 1999), the catalytic domain of MMP2 (1QIB, Dhanaraj *et al*., 1999), first Fib-II domain (1KS0, Gehrmann *et al*., 2002), second Fib-II domain (1CXW, Briknarova *et al*., 1999), third Fib-II domain (1J7M, Briknarova *et al*., 2001) and the C-terminal domain of MMP2 (1RTG, Gohlke *et al*., 1996). The 3D structure of charybdotoxin (2CRD) was downloaded in a PDB format from PDB-databank.

RASMOL and Swiss Deep-View software, developed by Swiss Institute of Bioinformatics (SIB), were used to visualise and analyse one or several protein models or structures at same time (Ramadan *et al*., 2014; SIB ExPASY, 2016). Swiss Deep-View can analyse, manipulate as well as superimpose proteins to visualise different aspects of a protein. Swiss Deep-View superimposes proteins by using root mean square deviation (RMSD) algorithm by measuring the average distance between the atoms of superimposed molecules (ExPASY, 2016). Swiss Deep-View was used to superimpose 3D structures of various regions of MMP2 on full-length MMP2 using the ‘magic fit’ tool. Then, all models were highlighted in different colours using the ‘colour layer’ tool. Charged surfaces of a 3D model or structure can also be determined by Swiss Deep-View (Ramadan *et al*., 2014). The electrostatic surfaces of Cltx and MMP2 were calculated using Swiss Deep-View

programme using the Poisson-Boltzmann method. The default software parameters were used to calculate charged surfaces. Calculated positively charged surfaces are coloured in blue (1.8 kT/e), negatively charged are in red (-1.8 kT/e), and neutral regions (no charge) are coloured in white, where ‘k’ is the Boltzmann constant, ‘T’ is the temperature in Kelvin, and ‘e’ is the charge of the electron.

3.2.2 Homology Modelling

Swiss Institute of Bioinformatics (SIB) Swiss-Model (https://swissmodel.expasy.org/) uses automated comparative modelling to predict 3D models of a protein from its amino acid sequence using an existing crystal structure ‘template’ for an evolutionarily related protein ‘target’. Three isoforms of MMP2 (1CK7) structure were predicted and downloaded in PDB format. In order to predict the native 3D structure of MMP2, the Ala404 mutation was deleted and Glu amino acid was inserted in its place. Then, the sequence was uploaded to Swiss-Model to predict the native 3D model using automated comparative modelling. Native MMP2 structure ($1CK7^{-}$) was predicted based on MMP2 template (PDB code: 1CK7) and MMP2 in conjunction with TIMP2 (PDB code: 1GXD) template in PDB format. In order to predict the 3D model of activated MMP2, the activation peptide (30-109 amino acids) of MMP2 (1CK7) was cleaved by Signal P server. Sequence was then uploaded to Swiss-Model to predict 3D model of activated MMP2 ($1CK7\ A^{-}$). To predict native activated MMP2, the activation peptide was cleaved, the Ala404 mutation was deleted and Glu amino acid was inserted in its place and then ($1CK7^{-}A^{-}$) was predicted.

3.2.3 Quality assessment

It is very important to estimate the quality of the 3D models as they may contain geometrical errors such as disorientated side chains. Qualitative Model Energy

ANalysis (QMEAN) is an estimate of the quality of a 3D model using several geometrical aspects and a comparison to experimental structures of similar sizes (Benkert *et al*., 2011). The Z-Score compares the QMEAN6 score to scores of a set of non-redundant high-resolution 3D structures of similar size experimentally determined by X-ray crystallography (Korneta *et al*., 2012).

Models were subjected for quality checks using Swiss-Model workspace (http://swissmodel.expasy.org/workspace/). QMEAN6 and Z-Score measures absolute quality and studies structural geometrical features of a 3D model compared to experimentally determined 3D structure of similar size (Korneta *et al*., 2012). QMEAN uses a reliability score ranging between 0 to 1; a value of a model closer to 1 represents a better quality model and indicates 'the degree of near-nativeness' of a predicted model (Benkert *et al*., 2011).

The QMEAN6 consists of a combination of 6 structural descriptors using statistical potentials used for quality estimation of a 3D model (Benkert *et al*., 2011). A torsion angle potential over three consecutive residues examines the back-bone local geometry of a 3D model. Long-range interactions of the 3D model are evaluated by two distant-dependent interactions potentials called Cβ interaction and all-atom interaction. The residue level interactions are assessed based on C-β atoms only, while all-atom interactions are used to assess long range interactions. The solvation potential identifies burial status of amino acids. In addition, QMEAN6 uses 2 extra statistical terms which describe the agreement between predicted (from sequence) and calculated secondary structure as well as solvent accessibility. DFIRE (distance-scaled, finite ideal-gas reference state) is an all atom potential energy function of a protein 3D structure (Zhou *et al*., 2002). Lower the DFIRE energy values means best quality protein models (Mishra and Hoque, 2015).

Procheck Ramachandran plot evaluates the stereochemistry of a predicted 3D model (Laskowski *et al*., 1993). Ramachadran plot investigates the backbone dihedral

angles ψ and φ of the residues of the 3D model and compare it to the stereo-chemical values of the experimentally derived 3D structures (Krishnamoorthy *et al*., 2016).

3.2.4 Protein-protein docking

ClusPro is a rigid-body docking programme used to evaluate protein-protein interactions (Kozakov *et al*., 2013). The ClusPro docking programme (www.cluspro.bu.edu) was used to dock Cltx with its putative binders MMP2 and AnxA2. ClusPro covers translational and rotational space of relative positions between peptide and protein and generates thousands of docked complexes. It represents the receptor molecule on a fixed grid and the ligand on a movable grid and rotates the ligand in a rotational and translation confirmation relative to the receptor. It generates 1000 complexes, clusters them and then minimises them via CHARMM (Chemistry at HARvard Macromolecular Mechanics) molecular simulation programme to find near-native complexes (Kozakov *et al*., 2013). Cluspro identifies near-native complexes by assuming that a near-native structure will have the largest cluster together with the lowest energy (stable clusters) (Xia *et al*., 2016). The number of docked structures within the top cluster above 122 members indicate a well-defined set of encounter complexes and provides assurance that the clustering of docked structures occurs in the neighbourhood of the native state. The lowest energy of the top cluster represents stability of the cluster and further increases confidence in its near-nativeness (Kozakov *et al*., 2017).

MMP2 isoforms and AnxA2 were selected as 'receptor' and Cltx was selected as a 'ligand' and then submitted to ClusPro server under default software parameters. As a control, charybdotoxin (ChTX) was selected as ‘ligand’ and then docked to the native activated MMP2.

The top cluster (with largest number of members) was downloaded in PDB format and then visualised in PyMOL. PyMOL software (www.pymol.org/) is a molecular visualisation system developed by Schrödinger Inc. Ligand and receptor were highlighted in separate colours to facilitate visualisation of protein complexes. Residues of ligand and receptor were visualised at 20 angstroms and residues of interest were coloured. Docked pockets were predicted by selecting ‘surface’ for receptor and ‘stick’ display for ligand.

3.3 RESULTS

3.3.1 Quality assessment of the MMP2 models.

The 3D structures of MMP2 (1CK7), Cltx (1CHL) and AnxA2 (2HYW) were downloaded from the RCSB-PDB databank; the 3D models of MMP2 isoforms were predicted utilising Swiss-Model.

The quality of 3D models of MMP2 was assessed using Swiss-Model workspace to find their usefulness in docking studies. Firstly, the quality of the 3D structure of MMP2 (1CK7) was assessed by QMEAN6, Z-Score and Ramachandran plot statistics (Fig 3.1). The QMEAN6 score for MMP2 structure (0.68) was within the acceptable range (The QMEAN6 score range from 0 to 1 with higher score for more reliable models). The Z-Score of MMP2 represented by a red star (Fig 3.1b) is within the Z-Scores of the reference 3D structures located in PDB-databank. The Z-Score for MMP2 was -0.92 which means the MMP2 3D structure is of comparable quality to high-resolution experimental structures of similar sizes (structure size +/- 10%) located in the PDB-databank. The Ramachandran plot (Fig 3.1c) shows that 88% residues of 1CK7 exhibit dihedral (φ and ψ) values in core region, 11.7 % allowed, 0% generously allowed and 0% in the disallowed region.

The quality of the other 3 isoforms of MMP2 was then assessed and compared to parent 1CK7 (see appendix for QMEAN6, Z-Score, DFIRE and Ramachandran plot). Overall, all models have good quality; QMEAN6 and Z-Score for all models are within an acceptable range (Table 4.1.0). All models had low energy values hence low steric hindrance and no clashes in atoms which means that the models have good stability (table 3.2). None of residues were in the disallowed region for all other isoforms of MMP2. Both 1CK7 and 1CK7^{-}A^{-} showed a similarity in Z-Score values that of -0.923 and -0.974 respectively. QMEAN6 values for all three models are in a

similar range (Table 3.2). In general, the quality data of all models falls within an acceptable range, showing a high similarity to parent 1CK7 (Table 4.1.0).

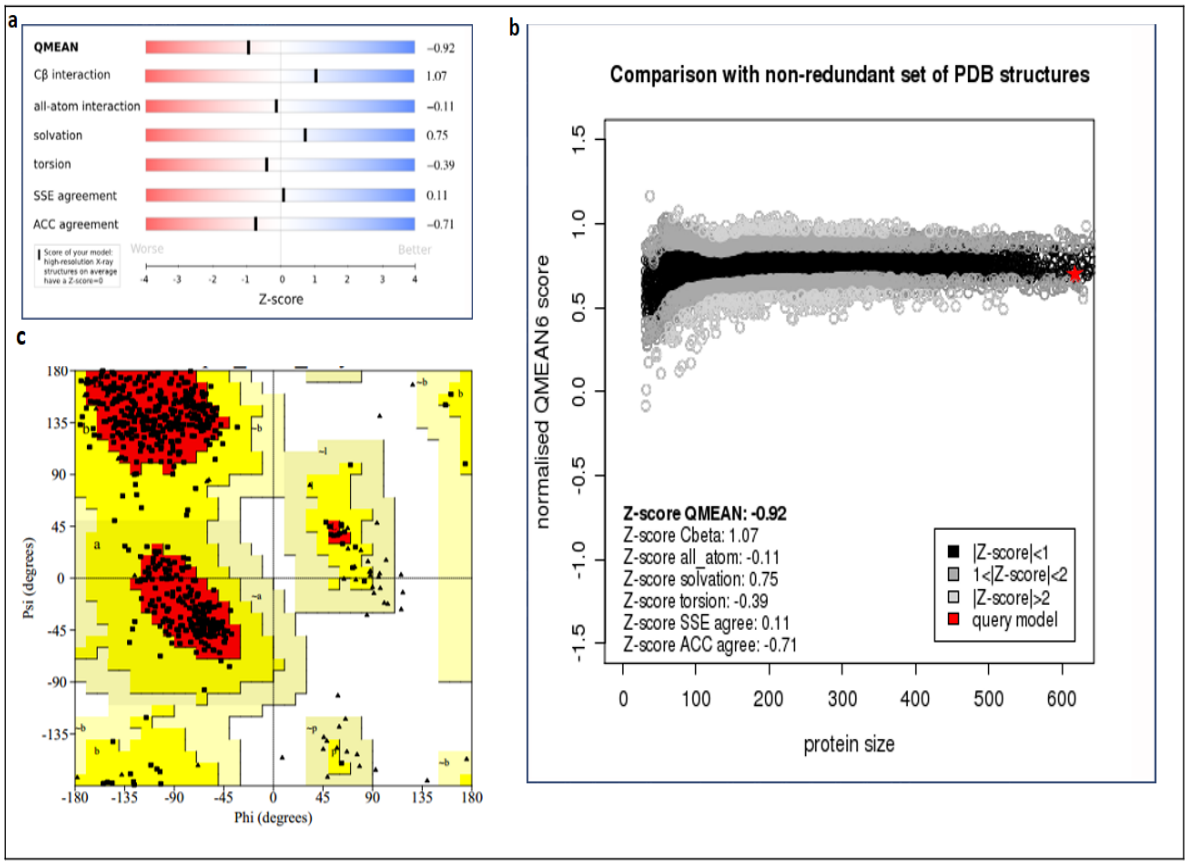

Fig 3.1. Stereochemistry and quality assessment of MMP2 models. (a) QMEAN6 score of full-length MMP2 (1CK7, E404A mutant with activation peptide), and identification of the score of individual geometrical features (b) QMEAN Z-Score of 1CK7, different shades of grey and black colour represent Z-Scores of the reference structures located in PDB-databank and red star represents MMP2 (c) shows Ramachandran plot statistics of 1CK7 3D structure. Red coloured region of Ramachandran plot shows residues in the most favoured region and yellow region show residues in the allowed region, grey region shows residues in generously allowed region and white region shows residues in the disallowed region. Glycine can be located in white regions as it does not contain a side chain.

Models	**DFIRE energy**	**QMEAN6 (Z-Score)**	**Residues in disallowed region % (Ramachandran)**
1CK7	-886.77	0.68 (-0.923)	0
1CK7^{-}	-868.60	0.717 (-0.512)	0
1CK7 A^{-}	-749.73	0.69 (-0.815)	0
1CK7^{-} A^{-}	-768.63	0.676 (-0.974)	0

Table 3.2. The quality assessment of 3D models of MMP2 isoforms. QMEAN6 and Z-Score values of all 4 models are within an acceptable QMEAN range. All models have a low energy hence less steric clashes. All models are good structures and comparable to a set of experimentally determined 3D structures of similar sizes from PDB databank.

The quality data suggested that the MMP2 3D structure as well as 3D models of MMP2 isoforms had an overall good quality and could be used in docking studies.

3.3.2 Charged surfaces of MMP2 models and effect of the point mutation.

The electrostatic surface potential of different MMP2 models was predicted by Swiss Deep-View and the effects of a point mutation on charged surfaces were studied. Charged surfaces of full length recombinant MMP2 (1CK7), native MMP2 with an activation peptide and without mutation (1CK7^{-}), recombinant MMP2 with mutation of E404A and without activation peptide (1CK7 A^{-}) and MMP2 without activation peptide and point mutation (1CK7^{-} A^{-}) were calculated (Fig 3.3).

The electrostatic charged surfaces reveal that all four models of MMP2 have an overall negative charge (Fig 3.3). The full-length MMP2 (1CK7) and native MMP2 contain a neutral charged region as well as a positively charged region at the activation peptide (Fig 3.3a and b). MMP2 models without the activation peptide have a highly negatively charged region around the catalytic domain. Both MMP2 models 1CK7 and 1CK7^{-} show a similar charge distribution around their surfaces. There are no differences in overall charged surfaces between these two models (Fig 3.3a and b). The addition of Glu at position 404 in the catalytic domain did not affect charged surfaces of native MMP2 model (1CK7^{-}). Both models have an overall negatively charged region around the catalytic and fibronectin domains. The C-terminal domain has neutral and positive regions particularly around blade 3 and 4. However, the activation peptide has neutral (no charge) as well as positive surfaces near the helix regions. Figures 3.3c show a highly negatively charged surface is exposed by removal of activation peptide near catalytic domain, this region is highly negative in both 1CK7 A^{-} and 1CK7 $^{-}$ A^{-} isoforms. The point mutation does not affect charged surfaces of MMP2, however removal of the activation peptide shows a highly negatively charged region (Fig 3.3), although there is a slight difference in the orientation between these two models.

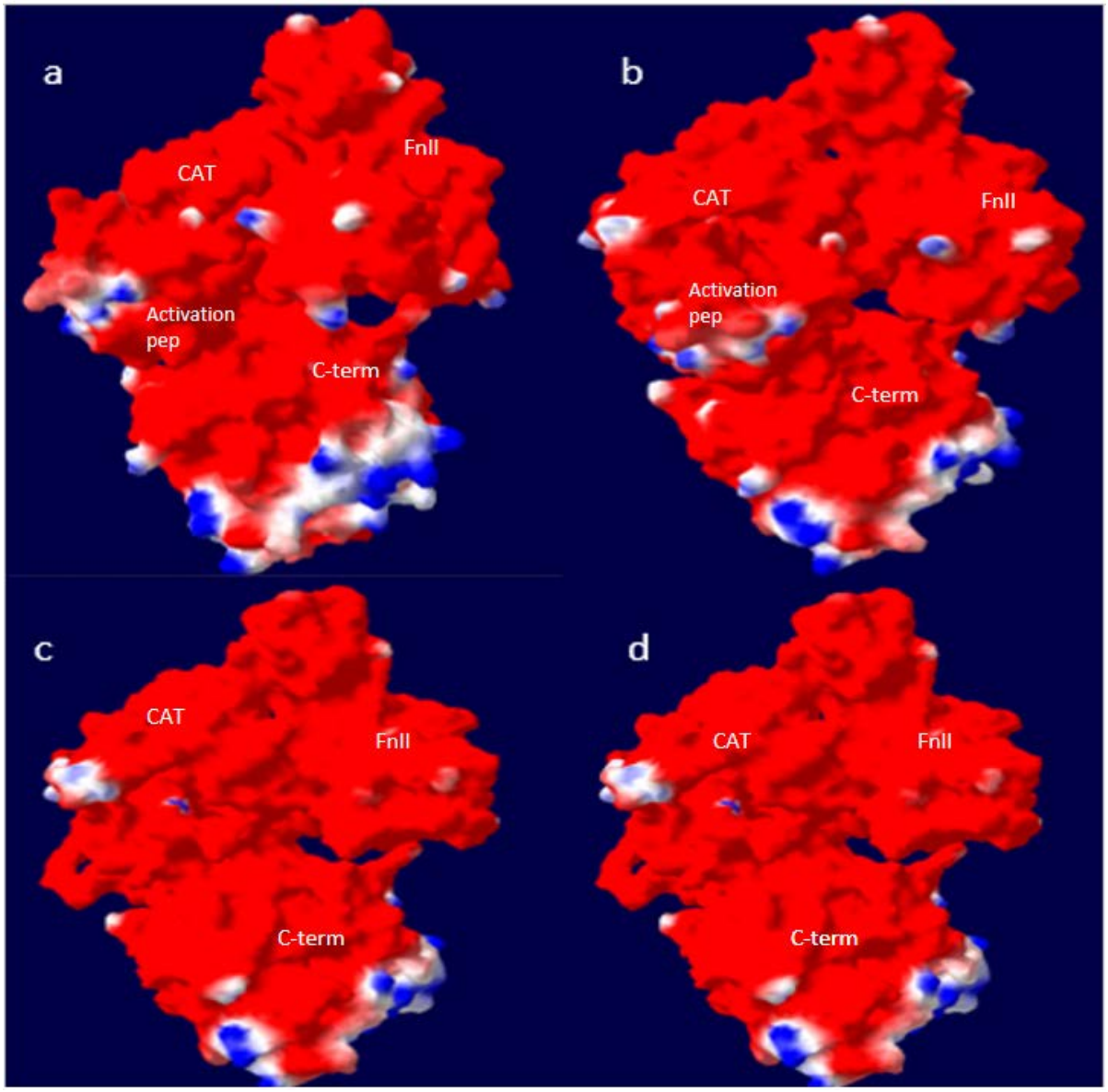

Figure 3.3. Charged surfaces of MMP2 isoforms. (a) Charged surfaces of 1CK7, (b) Native MMP2 1CK7^{-}, (c) 1CK7 A^{-}, (d) native activated MMP2 1CK7^{-} A^{-}. Different colour codes are red (-1.8 kT/e) and blue (1.8 kT/e) where 'k' is the Boltzmann constant, 'T' is the temperature in Kelvin, and 'e' is the charge of the electron. Default charge parameters of the programme were selected for calculation of charged surfaces of different isoforms of protein models.

3.3.3 Superimposition of 3D structures of MMP2, structure similarities and differences.

The MMP2 3D model was superimposed on the 3D structures of various regions of MMP2 to identify similarities and differences between different structures. The MMP2 3D structure was superimposed on the catalytic domain of MMP2 (1QIB), the first Fib-II domain of MMP2 (1KS0), the second Fib-II domain (1CXW), the third Fib-II domain (1J7M) and the C-terminal domain of MMP2 (1RTG). The catalytic domain of MMP2 (1QIB) show high structural similarity to the standard domains of 1CK7 A^{-}. Both 1CK7 A^{-} and 1QIB structures have β-sheets and 3-α helices which superimpose perfectly (Fig 3.4). However, there is small difference between the connective loops at N-terminal and C-terminal region which is highlighted with red arrows. The Fib-II domains are superimposed on Fib-II domains of 1CK7 A^{-}. Figure 3.4 shows that β-sheets of all three Fib-II domains superimpose well on 1CK7 A^{-} with a small difference in loop alignment at the N-terminal and the C-terminal domains of Fib-II domains, which is highlighted with green arrows. Differences in loops are present due to the fact that each individual Fib-II domain is superimposed on full-length 1CK7 A^{-}.

The C-terminal domain of MMP2 (1RTG) is superimposed on the MMP2 model as shown in Fig 3.4. Both structures contain four disc-shaped propellers and equal number of superimposed β-sheets. However, disc III of 1RTG has a β-sheet that is slightly longer than full-length MMP2 C-terminal domain, which is highlighted by a brown arrow (Fig 3.4). In general, all different models of MMP2 have similar structures to 1CK7 and there is only a small difference, particularly in loop regions between the structure of full-length MMP2 and its various published domains. Either full-length MMP2 or MMP2 model without its activation peptide can be used in docking studies rather than using each individual model, as it may represent a reliable interaction and an *in-vivo* scenario. Tao *et al*., (2009) cleaved the activation

peptide and used activated MMP2 in a protein docking study, in the present study it was decided to use either full-length MMP2 or an activated MMP2 model.

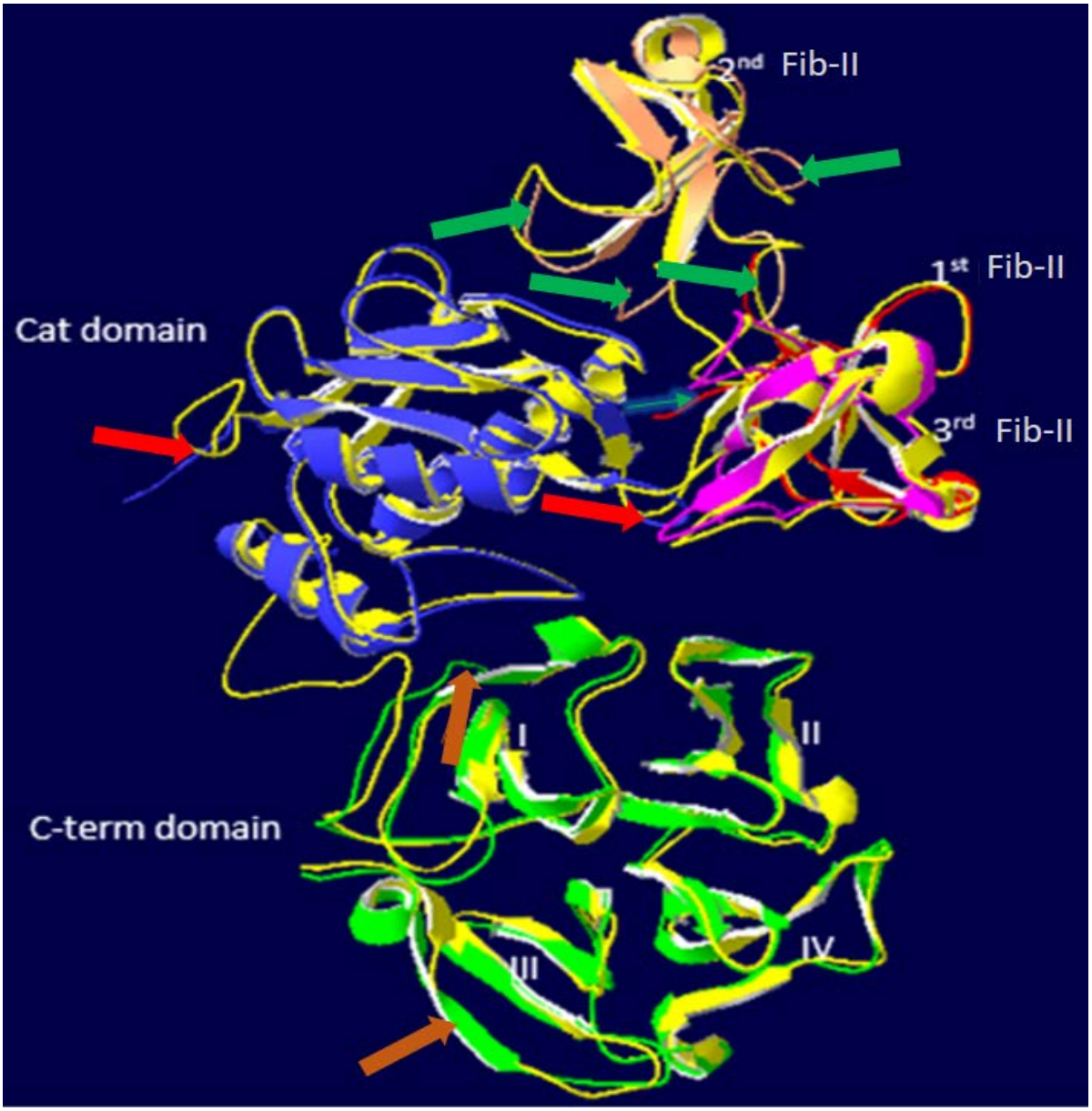

Fig 3.4. Superimposition of MMP2 1CK7 A$^-$ on different structures of MMP2 by Swiss Deep-View. The full-length MMP2 without activation peptide (1CK7 A$^-$) is coloured yellow, the catalytic domain only (1QIB) is coloured in blue, 1st Fib-II (1KS0) is coloured in red, 2nd Fib-II domain (1CXW) is coloured in grey, 3rd Fib-II (1J7M) domain is coloured in pink and C-terminal domain (1RTG) is coloured in green. Red arrows highlight differences in structures of 1CK7 A$^-$ and 1QIB, green arrows highlight differences in structures of the Fib-II domains while brown arrows represent differences in 1CK7 A$^-$ and C-terminal domain (1RTG). A brown arrow in the show different sizes of a β-sheet of the C-terminal domain.

3.3.4 Chlorotoxin docking to MMP2

MMP2 is a proposed target for Cltx however the mechanism of action is unknown hence Cltx was docked to full-length MMP2 to assess its feasibility as a binding partner. ClusPro generated a top cluster of Cltx-MMP2 complexes and this top cluster contained the highest number of docked structures. Top clusters and their energy scoring function were obtained in order to analyse accurate predictions of binding affinity of the peptide-protein complexes. The top cluster showing Cltx interaction with full-length MMP2 (1CK7) was further examined. Cltx is bound to the catalytic domain and second Fib-II domain in the top cluster structure (Fig 3.5.1a). The N-terminal region of Cltx is docked to the 2^{nd} Fib-II domain and β-sheets are docked near the catalytic domain of MMP2 (Fig 3.5.1a). Similarly, Cltx is also docked to native MMP2 ($1CK7^{-}$) near the catalytic and the 2^{nd} Fib-II domain (Fig 3.5.1b) and a similar face of Cltx is interacting with both recombinant and native MMP2 targets. This interaction provides evidence that E404A does not affect binding locus of Cltx on full-length MMP2. However, both MMP2 models contain an activation peptide at N-terminal domain, which is cleaved during *in-vivo* activation of MMP2 (Springman *et al*., 1990).

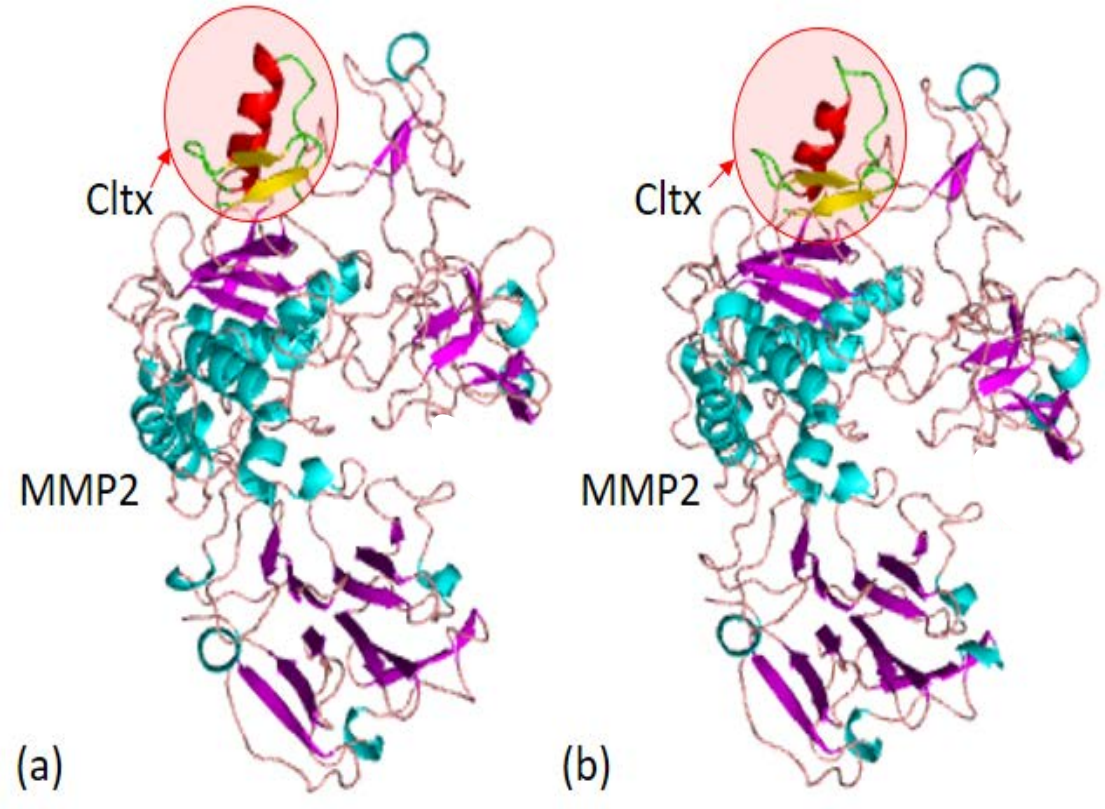

Figure 3.5.1. The top clusters of Cltx docked to human MMP2 via ClusPro. Cltx is highlighted in red circle, α-helix of Cltx is shown in red colour while β-sheets are shown in yellow colour. The α-helices of MMP2 are coloured in cyan and β-sheets are coloured in magenta. (a) Cltx is docked to full-length MMP2 (1CK7) at the catalytic and Fib-II domain of MMP2. (b) Cltx is docked to full-length MMP2 without point mutation (1CK7^{-}) in a similar position.

The binding pockets of the docked complex was determined by PyMOL. The top docked clusters of Cltx-1CK7 and Cltx-1CK7^{-} were analysed in order to find the binding locus of Cltx on MMP2. The top cluster shows that Cltx is docked in a pocket of MMP2, especially between the catalytic and the 2nd Fib-II domain. The N-terminal region of Cltx is present in close contact with the 2nd Fib-II domain whereas the C-terminal domain is in close contact with the catalytic domain of MMP2 (Fig 3.5.2).

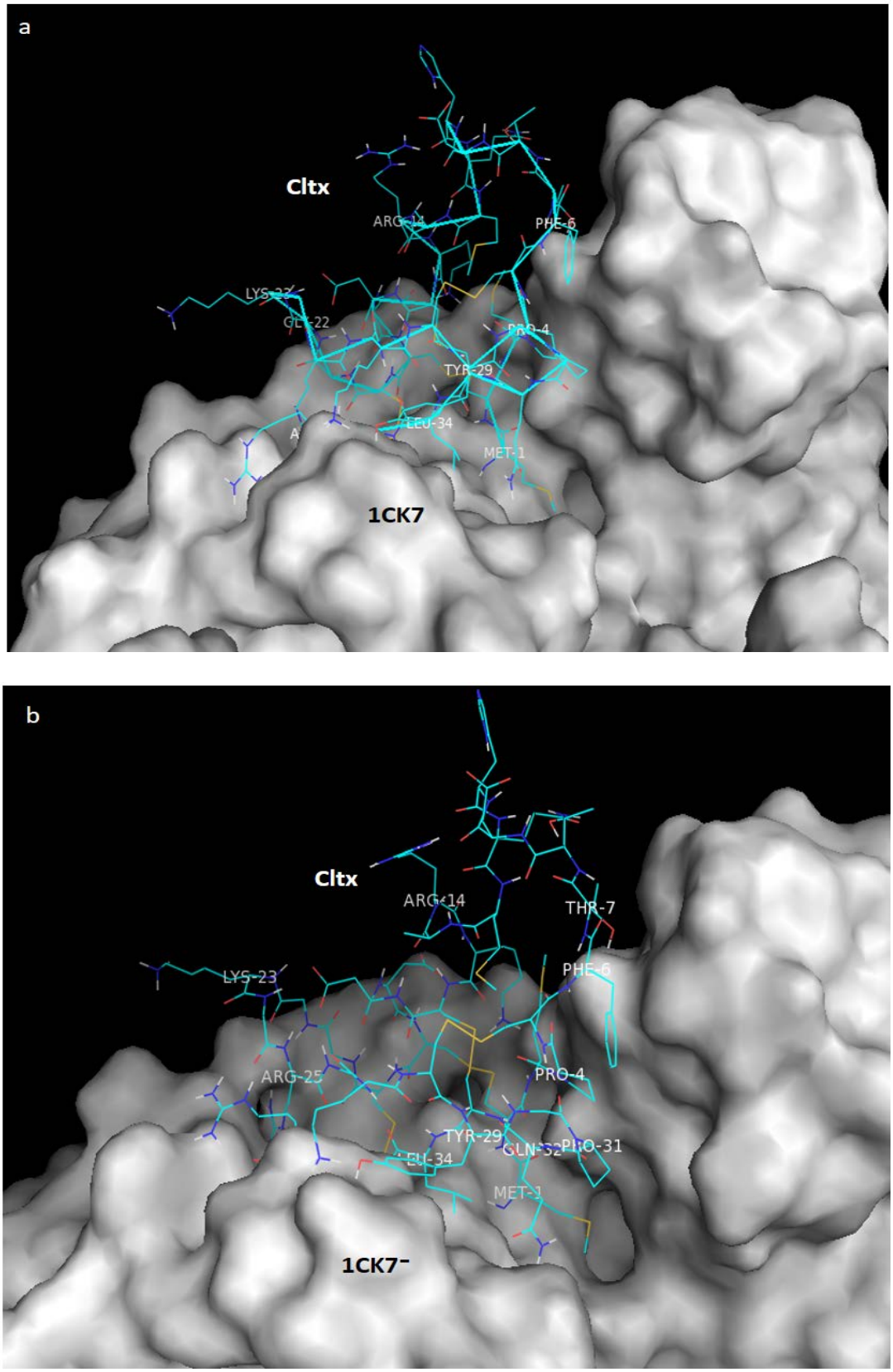

Figure 3.5.2. Cltx docked to MMP2. MMP2 is shown as grey surface and Cltx is shown in cyan. (a) Cltx is bound in a pocket between the catalytic domain and the Fib-II domain of MMP2 (1CK7). (b) Cltx is docked to 1CK7^{-} particularly bound in a pocket between the catalytic domain and second Fib-II domain.

3.3.5 Chlorotoxin docking with MMP2 without activation peptide.

Cltx was docked to recombinant MMP2 without the activation peptide (1CK7 A^-). The top cluster shows that Cltx interacts with 1CK7 A^- and is docked at the catalytic domain of MMP2 (Fig 3.5.3a). Cltx was then docked to native MMP2 without the activation peptide (activated MMP2, 1CK7$^-$ A^-) as this would represent an *in-vivo* interaction. A similar docking pose is observed in the top cluster of Cltx and 1CK7$^-$ A^- complex (Fig 3.5.3b). The top cluster shows that Cltx is bound near α-helices of the catalytic domain of 1CK7$^-$ A^-. In both clusters, Cltx is docked to MMP2 without activation peptide differently than MMP2 with an activation peptide (Fig 3.5.3).

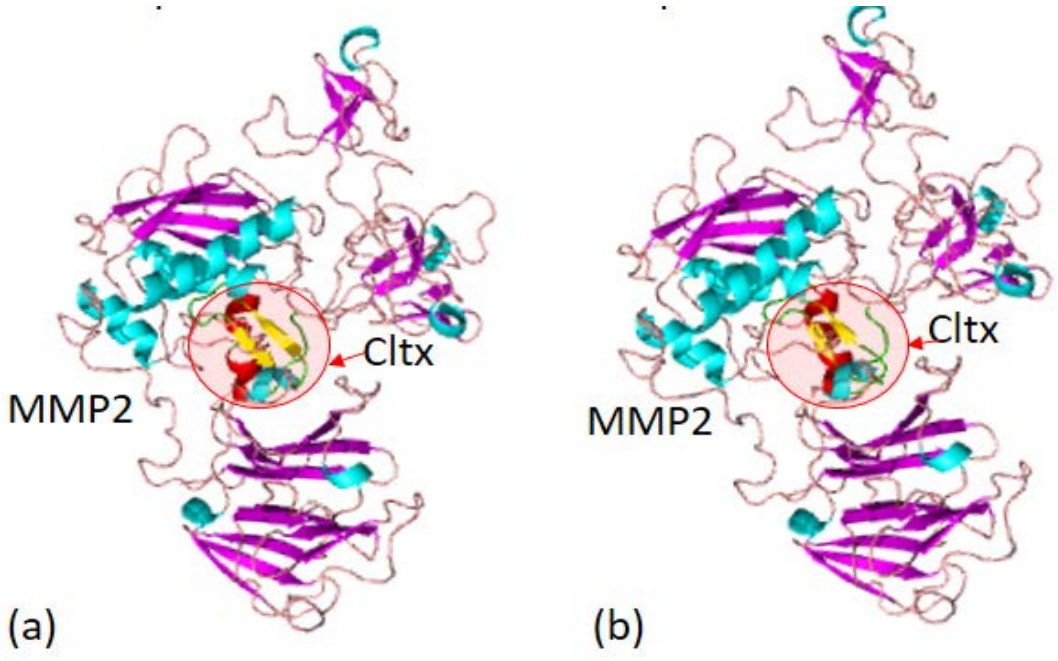

Figure 3.5.3. Cltx is docked to MMP2 without activation peptide. Cltx is highlighted in red circle, α-helix of Cltx is coloured in red and β-sheets are coloured in yellow. The α-helices of MMP2 are coloured in cyan and β-sheets are coloured in magenta. (a) Cltx is docked at the catalytic domain of recombinant MMP2 (E404A) without an activation peptide (1CK7 A^-), (b) shows Cltx docked to native MMP2 without activation peptide (that is Glu404, 1CK7$^-$ A^-).

Subsequently, the top clusters of Cltx and MMP2 without activation peptide were analysed by PyMOL programme. Cltx is docked in deep pockets of MMP2 target at the catalytic domain which represents interaction between Cltx and its target protein (Fig3.5.4). Cltx is interacting with both the mutated and native variant as shown in figure below.

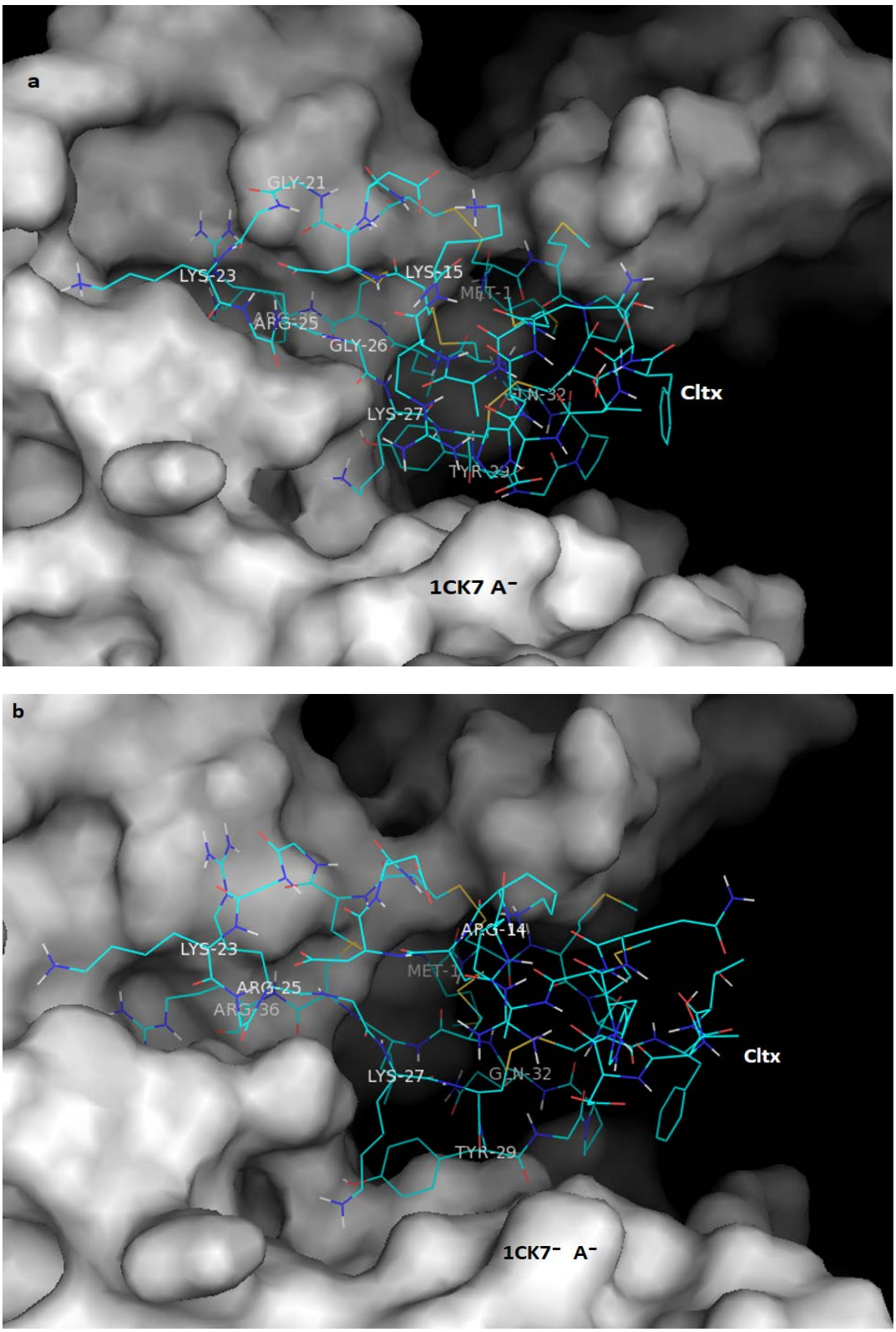

Figure 3.5.4 Cltx docked to MMP2 without activation peptide. Grey surface represents MMP2 target and Cltx is shown in cyan colour (a) Cltx is docked at the catalytic domain of MMP2. Amino acids of Cltx that are in close contact with target are labelled. (b) Cltx is docked at the catalytic domain of MMP2 without activation peptide and without mutation.

The MMP2 model without activation peptide and point mutation represents *in-vivo* activated enzyme, the top docked cluster of Cltx and $1CK7^{-}A^{-}$ may provide an *in-silico* model for the *in-vivo* interaction. The interaction of Cltx/$1CK7^{-}A^{-}$ was therefore analysed at the residue level by using PyMOL programme. Cltx is bound in deep pockets of the MMP2 catalytic domain as shown in Fig 3.5.5. Analysis of the top cluster shows that amino acids located at the C-terminal domain of Cltx are particularly positive residues such as Arg25 and Arg36 present in close contact to MMP2 (Fig 3.5.5). Amino acid Lys23 of Cltx is present in close contact with Asp77 and Asp80 of MMP2 and Arg25 and Arg26 positive residues of Cltx are present in close contact of Glu296 and Glu304 respectively which represent electrostatic interactions between two proteins. Another positive residue Lys27 of Cltx is also penetrated deeply in the catalytic domain of MMP2. An aromatic residue Tyr29 of Cltx is present in proximity of Thr318 of MMP2. However, Arg14 and Lys15 are present in free space and are not involved in interaction with target MMP2 (Fig 3.5.5).

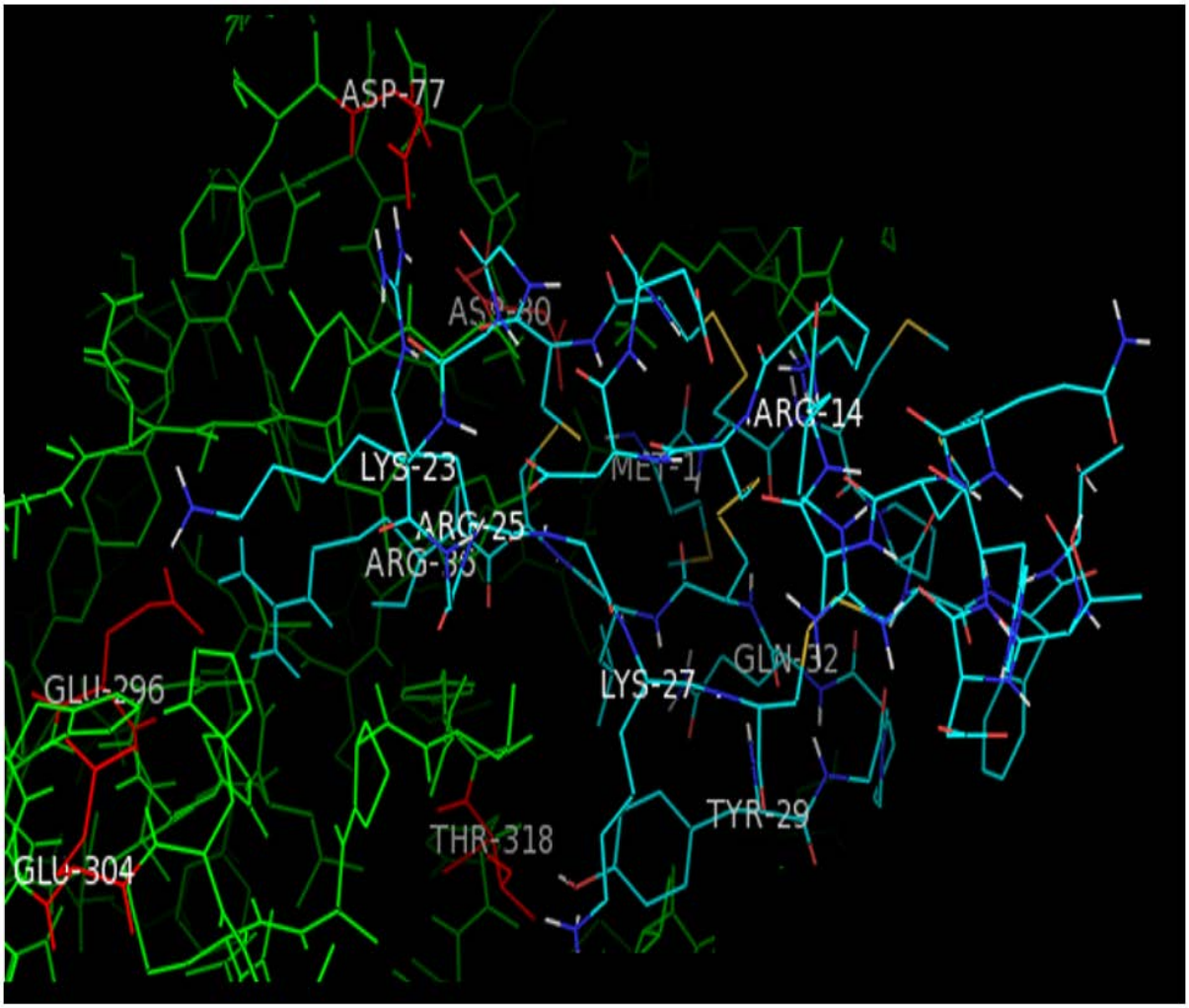

Figure 3.5.5. Interaction of Cltx and activated MMP2 (1CK7^{-}A^{-}) in the top docked cluster. Cltx ligand is coloured in cyan and target MMP2 is coloured in green. Amino acids of MMP2 that are contacting with Cltx are coloured red. Amino acids of interests on both Cltx and MMP2 are labelled. Mostly, positive residues of Cltx at C-terminal domain of the peptide are present in close contact to negative amino acids of MMP2.

3.3.6 Energy Scoring of docked complexes of Cltx-MMP2.

Energy scoring provides information on stability and binding affinity of docked complexes of peptide and target proteins. The most likely, near-native confirmation is determined by the largest cluster size (that is the highest number of docked neighbour structures in a cluster) and lowest energy (Kozakov *et al*., 2013). Lowest energy regions tend to predict bigger clusters of docked complexes. Differences in binding energy values of docked complexes represent differences in binding affinity.

According to table 3.6, the top cluster of Cltx and native activated MMP2 (Cltx/1CK7^{-}A^{-}) is a large cluster with strong binding affinity. This complex has the highest number of docked structures within the cluster and include 183 members. This complex also has the lowest energy of -956.5 which represent strong binding between two proteins. The centre docked structure of Cltx/1CK7^{-}A^{-} in the top cluster has energy of -709.5 which represents the stability of the cluster and a stronger binding affinity than other docked clusters (table 3.6). The top cluster of Cltx docked with mutant MMP2 without the activation peptide (Cltx/1CK7 A^{-}) ranked second on the list. The Cltx/1CK7 A^{-} cluster contains 179 members and a docked structure within cluster has -890 lowest energy. This suggests a strong binding affinity of Cltx with 1CK7 A^{-}, however target MMP2 has a mutation. Hence, Cltx/1CK7^{-}A^{-} large cluster together with lowest-energy indicates that the proposed docking complex is likely to reflect the position of near-native binding.

Peptide/Protein Complex	Cluster	Members	Representative	Weighted Score
Cltx/1CK7	**1**	**135**	**Centre**	**-729.5**
			Lowest energy	**-829**
Cltx/1CK7^{-}	**1**	**153**	**Centre**	**-766**
			Lowest energy	**-778**
Cltx/1CK7 A^{-}	**1**	**179**	**Centre**	**-890**
			Lowest energy	**-890**
Cltx/1CK7^{-}A^{-}	**1**	**183**	**Centre**	**-709.5**
			Lowest energy	**-956.5**

Table 3.6 Energy score profile of Cltx-MMP2 docked clusters. Members represents size of the docked complexes, that is number of docked structures within a cluster. Energy of cluster centre represents energy of docked complex that has the highest number of member or neighbours in the cluster and energy of the docked complex in cluster with lowest energy.

3.3.7 Charybdotoxin docking with MMP2.

Charybdotoxin (ChTX) is a calcium-activated potassium channel toxin isolated from *Leiurus quinquestriatus* scorpion venom (Laurent *et al.,* 1993). The 3D structure of ChTX was downloaded in PDB format from PDB-databank (PDB code: 2CRD) and then docked to the native activated MMP2 for negative control. The top cluster show that ChTX is docked near the first Fib-II domain and the second Fib-II domain (Fig 3.7). The ChTX is bound very differently to MMP2 compared to that of Cltx. The energy score shows a very binding affinity between ChTX and MMP2. The top cluster contains a structure with lowest energy of -512.5 hence very weak binding compared to that of Cltx binding to MMP2. This indicates that ChTX does not bind to MMP2 and therefore data of the present study with Cltx is validated, as one would expect based on experimental studies that Cltx interacts with MMP2 (Deshane *et al*., 2003; Alam *et al.,* 2012; Xu *et al*., 2016).

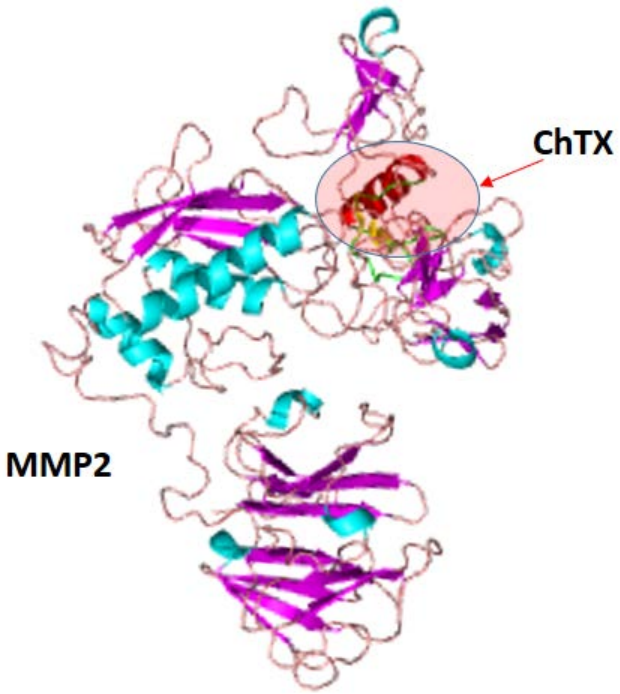

Figure 3.7. Top docked cluster of ChTX and MMP2. The ChTX is highlighted in red circle which is docked to the Fib-II domains of MMP2. The β-sheets of MMP2 are coloured in magenta and alpha-helices are coloured in cyan.

3.3.8 Cltx docking with Annexin A2

3.3.8.1. Cltx docking with human Annexin A2 with calcium bound

AnxA2 has been proposed a putative target for Cltx as the activity of AnxA2 was inhibited by Cltx in glioma studies (Kesavan *et al*., 2010). Cltx was docked to human AnxA2 in order to study their binding interaction. The 3D structure of AnxA2 (PDB code: 2HYW, chain A) was determined by X-ray diffraction studies at 2.10Å by Shao *et al.,* (2006). 2HYW contains 4 structurally similar homologous repeats of 5 α-helices (I-IV) arranged in a near-parallel shape, starting from Asn^{31} to Asp^{338} making a curved disk shape structure (Shao *et al.,* 2006). 2HYW shares structural similarity to human AnxA2 (PDB code: 1XJL), both 2HYW and 1XJL contain 7 Ca^{2+} ligands. Analysis of the top cluster show Cltx docked to repeat-I and IV of AnxA2 (Fig 3.8). The N-terminal domain of Cltx as well as β-sheets in the C-terminal domain are interacting with AnxA2 whereas α-helix of Cltx does not interact with the protein.

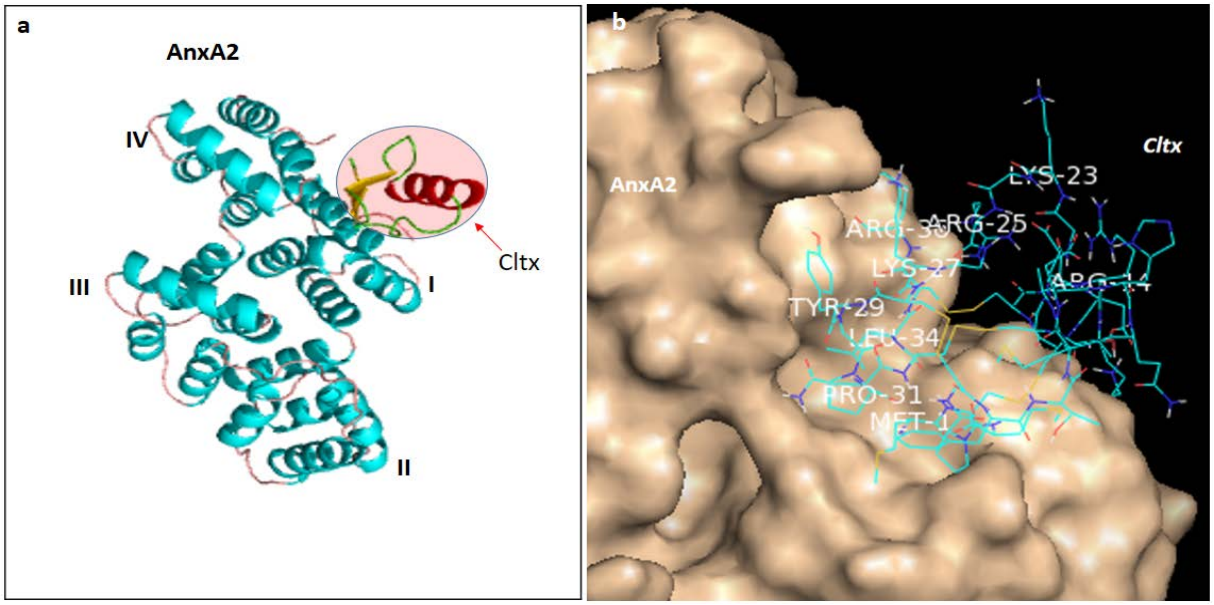

Figure 3.8. The top docked cluster of Cltx and AnxA2. (a) α-helices of AnxA2 are coloured in cyan and 4 homologous repeats that consists of α-helices are labelled. Cltx is highlighted in red circle, α-helix of Cltx is shown in red colour while β-sheets are coloured in yellow. (b) Cltx (cyan) docked in AnxA2 pockets (wheat).

The energy profile show that Cltx has binding affinity to AnxA2. The top cluster has large size and there are 288 docked members present in cluster 1 (table 3.9). The structure with strong binding affinity has lowest energy of -659.5 and a structure within cluster has centre energy of -561 (table 3.8). Although, Cltx-AnxA2 has highest number of members docked at similar locus on AnxA2 and there is stable binding between Cltx and AnxA2. However, binding affinity is weaker as compared to that of Cltx-MMP2 cluster (table 3.9).

Peptide/Protein Complex	**Cluster**	**Members**	**Representative**	**Weighted Score**
Cltx-AnxA2	**1**	**288**	**Centre energy**	**-561**
			Lowest energy	**-659.5**

Table 3.9. Energy score profile of Cltx-AnxA2 cluster. The top cluster contains 288 docked members i.e. docked structure at similar locus in the top cluster. Strong binding affinity of a structure within top cluster is represented by lowest energy. In addition, a docked structure at centre of the cluster also has good binding affinity with energy of -561.

3.4 Discussion.

MMP2 and AnxA2 are over expressed in GBM hence targeting these proteins is crucial for GBM treatment (Lokman *et al*., 2011). Cltx was docked to its proposed putative targets MMP2 and AnxA2 in order to study their interaction. All 3D structures of MMP2 superimpose and there is no big difference between the models, hence Cltx docking to individual MMP2 regions is unnecessary. Thus, Cltx was docked to native MMP2 as well as activated MMP2 in the present study.

The results in the present study suggest that Cltx interacts with MMP2. This correlates with Deshane *et al*., (2003) who showed that Cltx interacted with MMP2 and reduced its enzymatic activity via matrigel invasion assay and an enzyme-linked immunosorbent binding assay as well as gelatin zymography. Alam *et al*., (2012) and Ali *et al*., (2016) showed that Cltx inhibited the activity of MMP2 via FRET substrate assay and gelatin zymography.

The results show that Cltx is docked to full-length MMP2 between the catalytic domain and the Fib-II in the top cluster. Docking Cltx to the catalytic domain only (1QIB) or any other individual 3D structure of a region of MMP2 is not justified as it may not represent an *in-vivo* scenario.

Tao *et al*., (2009) cleaved the activation peptide of MMP2 in order to access the catalytic site to mimic the *in-vivo* scenario and then docked it to SB-3CT inhibitor. In our study, the activation peptide was cleaved and then MMP2 1CK7 protein was un-mutated in order to study Cltx interaction with native MMP2. Removal of the activation peptide exposes a site that was covered by the activation peptide in full-length MMP2. Results indicated that Cltx has a stronger binding affinity with native activated MMP2 than the other forms of MMP2 tested. Farsani *et al*., (2015) conducted an *in-silico* experiment and docked Cltx-like peptide MeICT (a Cltx homologue) with the catalytic domain of MMP2 (1QIB). Although, docking with the

catalytic domain only is not justified as 1QIB does not represent a full *in-vivo* activated MMP2. They showed that the α-helix of Cltx-like peptide MeICT docked to MMP2 catalytic domain near the α-helices. Our results show that the β-sheets and loops of Cltx are docked to MMP2 and α-helix does not seem to be interacting with MMP2.

The present study shows that Cltx has strong binding affinity for activated MMP2 as shown in Fig3.5.3. Electrostatic surface analysis showed that MMP2 has a strong negatively charged surface as shown in Fig 3.3. Positive residues that are located mainly at the C-terminal domain of toxins are present in close contact with MMP2 (Fig3.5.5). The present study shows that the positive residues such as Arg14 and Lys15 do not interact with MMP2 in top cluster and both residues are present in free space. This suggests that both Arg14 and Lys15 may be involved in binding to the chloride channel and may induce a conformational change to the channel. It is possible that Cltx may bind to two separate targets at same time such as MMP2 and AnxA2 or chloride channel. Results show that positive residues such as Arg25, Lys27 and Arg36 are present in close contact with MMP2 but Arg14 and Lys15 amino acids are free, Arg14 and Lys15 may interact with other targets of Cltx that are AnxA2 or chloride channel (Fig4.2.5).

AnxA2 is involved in glioma metastasis and angiogenesis and proposed a target for Cltx (Kesavan *et al*., 2010; Zhai *et al*., 2011). Tatenhorst *et al*., (2006) and Kesavan *et al*., (2010) suggested that Cltx interacts with AnxA2. The present study indicates that Cltx binds to AnxA2 and show that Cltx interacts with AnxA2 near annexin repeat-I and IV. However, binding affinity between Cltx-AnxA2 is lower than binding affinity of Cltx-MMP2 (table 3.6 and 3.9). Sorocenau *et al*., (1998) suggested that Cltx has a lower binding affinity site and a higher affinity site in GBM. Based on our results, it can be assumed that Cltx has stronger binding affinity for MMP2 but lower binding affinity for AnxA2 (table 3.6 and 3.9).

Molecular dynamics would be useful to confirm results although, MD simulations are computationally expensive and requires multicore operating system and performed at industry level. This *in-silico* study provided reliable models and docking conformations with various binding affinities. Hence, laboratory screens on recombinant homologues can be performed based on these results to focus on MMP2 and confirm residues of interests. This rigid-body docking programme did not consider Cltx and its targets MMP2 and AnxA2 as flexible molecules and there was no conformational change upon binding while docking. Although the conformational change for enzyme-inhibitor complex is very small and only occurs in the side chains of ligands (Kozakov *et al*., 2007). Cltx is a small peptide so conformational change upon peptide-target association is moderate which does not affect the complex backbone geometry and binding affinity. In addition to rigid-body docking, flexible docking programme are available such as Rosetta Dock and High Ambiguity Driven Biomolecular Docking (HADDOCK) which consider two molecules as flexible proteins, however, a *priori* biological and structural information and binding locus is required for Rosetta dock. Moreover, HADDOCK requires a *priori* knowledge of interacting residues prior to docking, since there is a limited knowledge about Cltx interaction with its putative partners, Rosetta and HADDOCK are not practical to analyse interaction studies. However, rigid-body programme uses a global search of two interacting molecules, which is useful if there is a limited data of interacting molecule is present hence ClusPro is fast and produced good complexes in our study (Kozakov *et al*., 2013).

Critical assessment of prediction of interactions (CAPRI) is a communitywide experiment that monitors progress of docking programmes (Lensink *et al*., 2016). Cluspro web-server has been used in more than 400 publications (Xia *et al*., 2016). Cluspro ranked top one in CAPRI 2016 blind protein docking, placing it the first best predictor web-server (Lensink *et al*., 2016).

4 CHLOROTOXIN-LIKE SHORT TOXINS DOCKING with MMP2 and ANNEXIN A2

4.1 Introduction

Although a few studies have attempted to understand the mechanism of action of Cltx, there is very limited data available about the structure-functional relationship. Attempts have been made to study new Cltx homologues to investigate their therapeutic potential (Rjeibi *et al*., 2010; Farsani *et al*., 2015; Xu *et al*., 2016). A number of chlorotoxin-like short toxins have been isolated from the venom of different scorpion species. Chlorotoxin-like short toxins (Cltx-like short toxins) share sequence similarity to Cltx and they may have same therapeutic potential as Cltx (Xu *et al*., 2016). Cltx-like short toxins affecting chloride channels are constrained by 4 disulphide bridges and they have a consensus formula of $CX_2CX_{10}CX_2CCX_{5\text{-}7}CX_{3\text{-}4}CXCX_n$. Toxins have been isolated (purified) as mature peptide from venom of scorpions as well as identified from cDNA libraries using venom gland mRNA.

Around 34 Cltx-like short toxins have been described, however, only a few toxins have been used in experimental studies involving glioma cells. The BMK CT isolated from Chinese scorpion *Buthus martensii Karsch* and Chlorotoxin-like peptide (AaCTx) from *Androctonus australis* share 68% and 61% sequence identity to Cltx respectively (Zhao *et al*., 2010; Rjeibi *et al*., 2010). Sequence similarity between Cltx and these toxins suggests that these toxins have similar functions and molecular targets. Both Cltx-like peptide AaCTX and BMK CT inhibited glioma proliferation and metastasis in a dose-dependent manner with no effect on normal glial cells (Zhao *et al*., 2011; Rjeibi *et al*., 2010). Activity of Cltx-like peptide AaCTX was less than 10% compared to Cltx in glioma (Rjeibi *et al*., 2010). It suggests that some toxins may have higher efficacy than others.

Recently, two scorpion toxin derivatives CA4 and CTX23 were designed based on Cltx and BMK CT sequences and tested on glioma (Xu *et al*., 2016). Both CA4 and CTX23 decreased glioma cell growth and proliferation. As expected, CA4 and

CTX23 did not show cytotoxic effects to normal neurons and astrocytes. CA4 and Cltx also inhibited proliferation of endothelial cells demonstrating that both CA4 and CTX23 have anti-glioma properties as well as anti-angiogenic properties (Xu *et al*., 2016).

There are many other peptides known as Lqh8/6, chlorotoxoin-1, Insectotoxin I5A and GaTx1 that share sequence homology to Cltx (Fig 1). To date, only structures of the Cltx and Insectotoxin I5A have been determined (Lippens *et al*., 1995; Arseniev *et al*., 1993). In this chapter, we will use molecular modelling to predict 3D models of the Cltx-like short toxins that have already been used in glioma studies and docking techniques to study the mechanism of action of Cltx-like short toxins with a view to the development of new clinical agents. Biological functions of CTX23, CA4 and Cltx-like peptide AaCTX is known as they have been investigated in anti-glioma studies. Although, mechanism of action of these toxins is unclear. CA4, CTX23 and Cltx-like peptide AaCTX have been used in glioma studies, we decided to study structural-functional relationship of Cltx, CA4, CTX23 and Cltx-like peptide AaCTX.

4.2 Methodology

4.2.1 Primary Structure of Cltx-like short toxins

The amino acid sequences of Cltx-like short toxins based on the consensus formula $CX_2CX_{10}CX_2CCX_{5\text{-}7}CX_{3\text{-}4}CXCX_n$ were extracted from bioinformatics database UniProt server (http://www.uniprot.org/), and primary literature using Pubmed (https://www.ncbi.nlm.nih.gov/pubmed/). Sequences that contained signal peptide were submitted to Signal P 4.1 server (http://www.cbs.dtu.dk/services/SignalP/). Signal P cleaves signal sequence of a protein using hydrophobicity scores as signal sequence are very hydrophobic in nature (Petersen *et al*., 2011) and then sequences were aligned using Clustal Omega server (http://www.ebi.ac.uk/Tools/msa/clustalo/). Clustal Omega uses seeded guide trees and Hidden Markov Model (HMM) profile method to align multiple sequences (Li *et al*., 2015). Amino acids were colour coded based on their properties using CINEMA colour coding scheme. Negative and positive amino acids were coloured in red and blue respectively. Polar amino acids were coloured in green and non-polar amino acids were coloured in grey. Highly conserved cysteine residues were coloured in yellow.

4.2.2 Secondary structure

Secondary structure of all toxins was predicted using PSIPRED server (http://bioinf.cs.ucl.ac.uk/psipred/). PSIPRED uses a secondary structure prediction method to analyse secondary structure of a protein from its primary sequence (Jiang *et al*., 2016). Molecular weight and Isoelectric points of all short-toxins were calculated using ExPASY compute pI/Mw tool (http://web.expasy.org/compute_pi/).

4.2.3 Generation of 3D models of Cltx-like short toxins and their analysis

Swiss-Model was used to generate 3D models of CA4 peptide, CTX23 and Cltx-like peptide AaCTX via automated comparative homology modelling (https://swissmodel.expasy.org/). 3D structures of Cltx (PDB code: 1CHL) and I5A (PDB code: 1SIS) present in RCSB PDB databank (http://www.rcsb.org/PDB/home/home.do) were used as templates for predicted 3D models. Both CA4 and CTX23 were predicted based on 1CHL template whereas, Cltx-like peptide was predicted based on I5A template. Then, 3D models were subjected for quality assessment. The quality of all models was assessed using QMEAN6, Z-Score, DFIRE energy and Procheck Ramachandran plot via Swiss-Model workspace (http://swissmodel.expasy.org/workspace/) as described previously.

RASMOL and Swiss Deep-View v4.1 was used to visualise and analyse 3D models of toxin peptides. Swiss Deep-View was used to superimpose all 3 toxin models and Cltx using ‘magic fit’ tool. The electrostatic surfaces of the 3D models and Cltx was calculated using Swiss Deep-View as described previously in 3.2.

4.2.4 Docking of Cltx-like short toxins with MMP2 and Annexin A2

ClusPro programme (https://cluspro.org) was used to dock activated MMP2 ($1CK7^{-}A^{-}$) and AnxA2 with CA4, CTX23 and Cltx-like peptide AaCTX as described previously in 3.2. The top cluster (with the highest number of members) was downloaded in PDB format and then analysed using Swiss Deep-View and PyMOL. The energy score profile of the top cluster was downloaded in order to study binding affinity of peptide-protein complex. The PDB file of the top cluster was visualised in PyMOL programme at 20 Angstroms. Surface representation was selected for receptor protein and stick representation was selected for ligand toxin. Residues of ligands that were contacting with residues of receptor protein was coloured and labelled.

4.3 Results.

4.3.1 Comparison of primary structure of Cltx-like short toxins

In order to compare the primary structures of all known Cltx-like short toxins, amino acid sequences were extracted from UniProt, GeneBank, TrEMBL, EMBL-EBI and the primary literature. Primary sequence alignment of 34 Cltx-like short toxin peptides is shown. Sequence alignment reveals that chloride channel toxins show high sequence homology to chlorotoxin, 8 conserved cysteine residues form a common disulphide bridge scaffold.

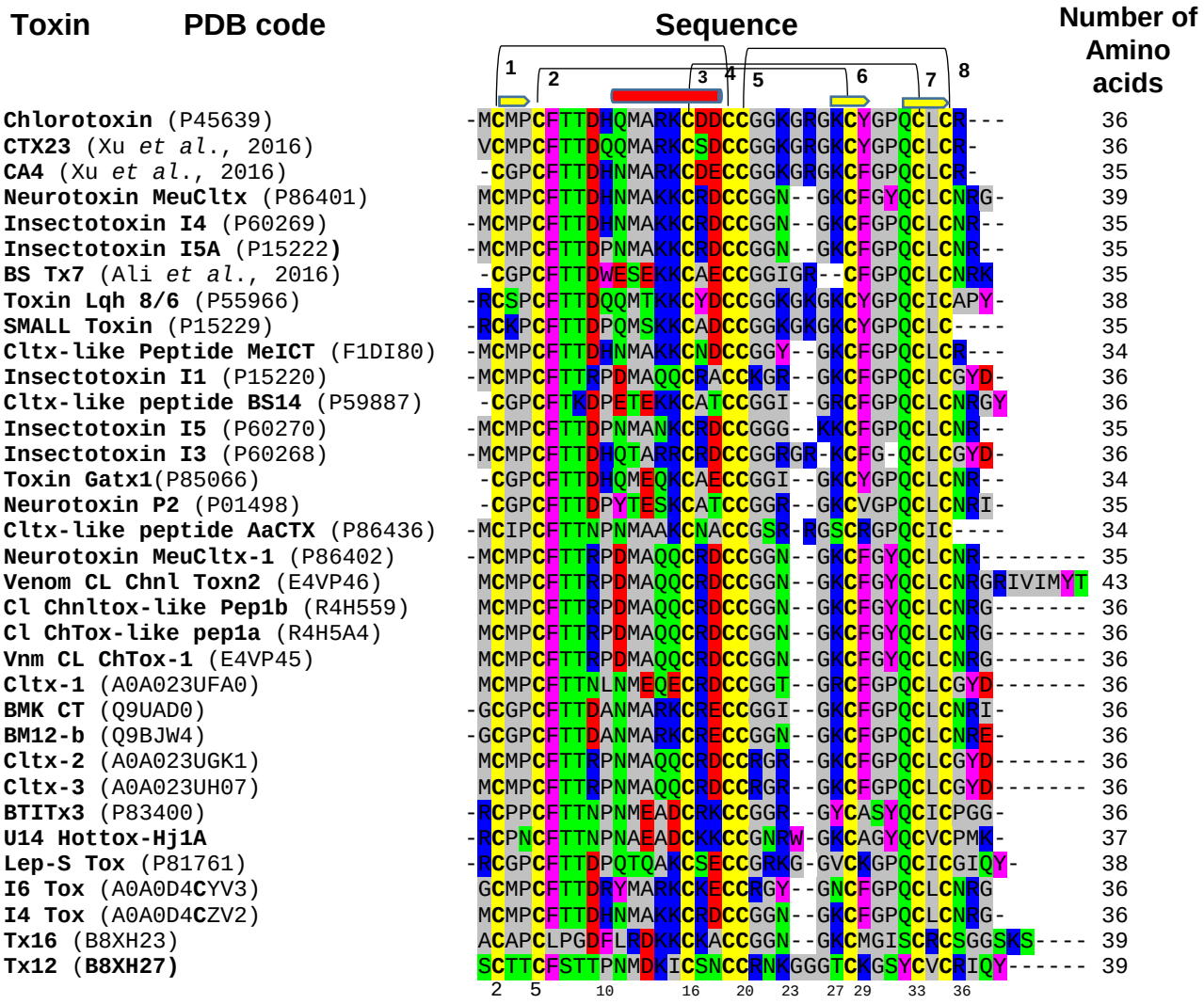

Figure 4.1 Primary Sequence alignment of Cltx-like short toxins. Amino acids are colour coded based on their properties. Positive residues are highlighted blue, negative residues are highlighted red, amino acids with polar side chains are highlighted green and residues with non-polar side chains are highlighted grey. Conserved cysteine amino acids are highlighted yellow and aromatic amino acids are highlighted magenta. Yellow arrows on top of sequences represent three β-sheets and the α-helix is represented by a red cylinder. Disulphide bridge pattern is represented by black lines linked between C1-C4, C2-C6, C3-C7, C5-C8. Where identified signal peptide (from cDNA) have been removed.

All Cltx-like short toxins share highly conserved cysteine residues which produce four disulphide bridges. Number of amino acids is constant between Cys1-Cys2-Cys3-Cy4, though the number of amino acids varies between Cys5-Cys6 and Cys6-Cys7, Csy8 at C-terminal domain is variable. There is a semi-conserved *PCFTx* motif at N-terminal domain, present in all toxins except Tx12, Tx16 and U14 hottentoxin-Hj1a (Fig 4.1). Cltx-like peptide Bs14 has a positive residue Lys at position 7 instead of a polar residue Thr7.

Semi-conserved region at C-terminal domain '*GPQCxC'* is present in all toxins except BtITx3, I6 toxin, U14 Hottentoxin-Hj1a, Tx16 whereas, an aromatic residue Tyr31 is present instead of Pro29 in Neurotoxin MeuCltx, Neurotoxin MeuCltx-1, Venom chloride channel toxin-1 and 2, Chloride channel toxin-like peptide 1a and 1b (Fig 4.1). A positive residue Arg is present at the C-terminal domain of Cltx and other toxins except small toxin, toxin Lqh 8/6, I1, I3, Cltx-like peptide AaCtx, Cltx-1, 2 and 3, BtITx3 as well as Lepidopteran-selective toxin. Lysine is present at C-terminal domain instead of Arg in U14 Hottentoxin-Hj1a and Tx16. Positive residues Lys25 and Lys27 is conserved in Cltx and many other toxins except BtITx3, Lepidopteran selective toxin, TX12 and I6 toxin and Cltx-like peptide. Positive residues Arg14 or Lys14 and Lys15 are conserved in many toxins, giving chloride channel toxins a stronger positive charge. Cltx and many other toxins also has a conserved negative Asp9 amino acid (Fig 4.1). Sequence similarity between toxin peptides indicates the possibility of similar structure and function. Although, sequences of many Cltx-like short toxins have been determined, only a few toxins have been used in glioma. To date, 3D structures of Cltx and I5A has been experimentally determined and available in PDB-databank (DeBin *et al*., 1993; Arseniev *et al*., 1993). There is a difference of 10 amino acids between Cltx and BMK CT whereas, CA4 and CTX23 has a difference of 5 and 3 amino acids from Cltx respectively which may represent difference in loop orientation. Cltx-like peptide AaCTX is a 34 amino acids peptide

and it shares 61% sequence identity to Cltx. Both CA4 and CTX23 share 78% and 85% sequence similarity to Cltx. In present study, secondary structure analysis by PSIPRED showed that all toxins share similar secondary structure to parent Cltx and Insectotoxin I5A containing 2 extended β-sheets and one small fragment of β-sheet as well as an α-helix.

Cltx is a 4002Da and the peptides observed masses of all Cltx-like short toxins range between 3598Da and 4173Da. Cltx is a basic peptide at physiological pH with a PI of 8.5 and most other toxins have similar PI values 8.5 (±0.50) except Cltx-1 and Gatx1 with a PI values of 4.32 and 6.69 respectively. Hence, all Cltx-like short toxins are positively charged with exception of Cltx-1 and Gatx1. Sequence homology ranges between 48-100%. CA4 and CTX23 were selected for further docking studies.

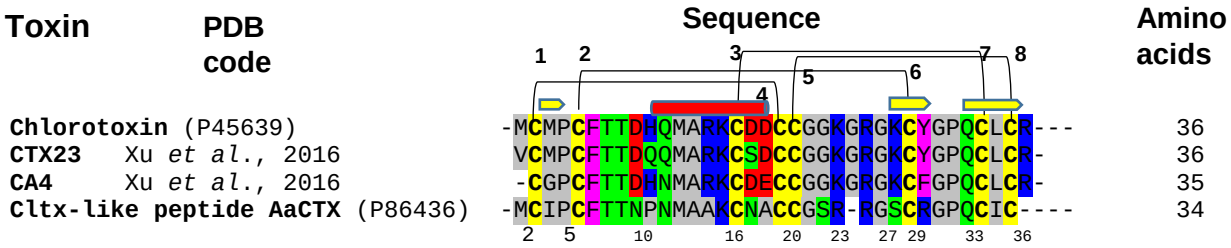

Figure 4.2. Sequence alignment of Cltx, CTX23, CA4 and Cltx-like peptide AaCTX. Amino acids are colour coded in CINEMA colour scheme. Three β-sheets are shown by yellow arrows and the α-helix represented by a red cylinder.

All 4 peptides share semi-conserved region motif '*PCFTT*' at N-terminal domain and '*GPQCxC'* at C-terminal region of toxin sequences. CA4 has a truncated N-terminus beginning with Cys1. Aromatic residue Phe6 is highly conserved in all four Cltx-like toxins (Phe5 in CA4). Both Cltx and CTX23 share conserved positive Arg14, Lys15 amino acids and Lys23, Arg25 as well as Lys27 except Cltx-like peptide AaCTX. CA4 contains conserved positive residues Arg13, Lys14, Lys22, Arg24 and Arg35.

All peptides share Arg35- Arg36 at C-terminal domain except Cltx-like peptide AaCTX. Cltx-like peptide AaCTX has a Cys amino acid at position 34 at C-terminal domain but contains Arg at position 23, 24 and 28. Negative residues Asp17 and Asp18 are located on α-helix of Cltx and Asp16 and Glu17 are present in CA4, Interestingly, CTX23 has a polar residue Ser at position 17 whereas, Cltx-like peptide AaCTX has polar a residue Asn17 and a non-polar residue Ala at position18. There is no single negative residue present in Cltx-like peptide AaCTX indicating a strong positive charged surface of toxin. Both Cltx and CTX23 has an aromatic residue Tyr at position 29 though CA4 has Phe at position 28 indicating a neutral region around this region of the peptide. However, Cltx-like peptide AaCTX contain a positive residue Arg at position 28 indicating a positive charged region around this area. Hydrophobic residues Gly 30, Pro31, Leu34 located of toxins peptide and this motif is conserved in many other toxins (Fig4.2).

4.3.2. Comparison of 3D models of Cltx-like short toxins.

3D models of CA4, CTX23 and Cltx-like peptide AaCTX were generated by homology modelling using Swiss-Model. 3D models of CA4 and CTX23 were generated based on Cltx template (PDB code: 1CHL), while Cltx-like peptide AaCTX was generated based on I5A template (PDB code: 1SIS). Subsequently, the quality of 3D models was assessed using QMEAN6, Z-Score, DFIRE energy and Procheck Ramachandran plot. The quality of 3D models was better than its parent Cltx 3D structure (See appendix). Superimposition of these models is shown in Fig 4.3. A cysteine stabilised α-helix (CSH) fold is linked by 3 disulphide bridges to extended β-sheets and fourth disulphide bridge is linked to loop region. All 4 peptides share 2 small anti-parallel β-sheets (arrows) and a small fragment of β-sheet in extended loop as well as an α-helix. Cltx-like short toxins share CSH structural scaffold stabilised by disulphide bridges which gives toxin a compact structure and stability.

A ribbon representation of superimposed toxins using Swiss Deep-View. All toxins superimpose perfectly as shown in Fig 4.3 The α-helix of Cltx-like peptide AaCTX is slightly longer than Cltx, CTX23 and CA4. Blue arrows show that there is small deviation between loop region of Cltx-like peptide AaCTX as compared to other toxins.

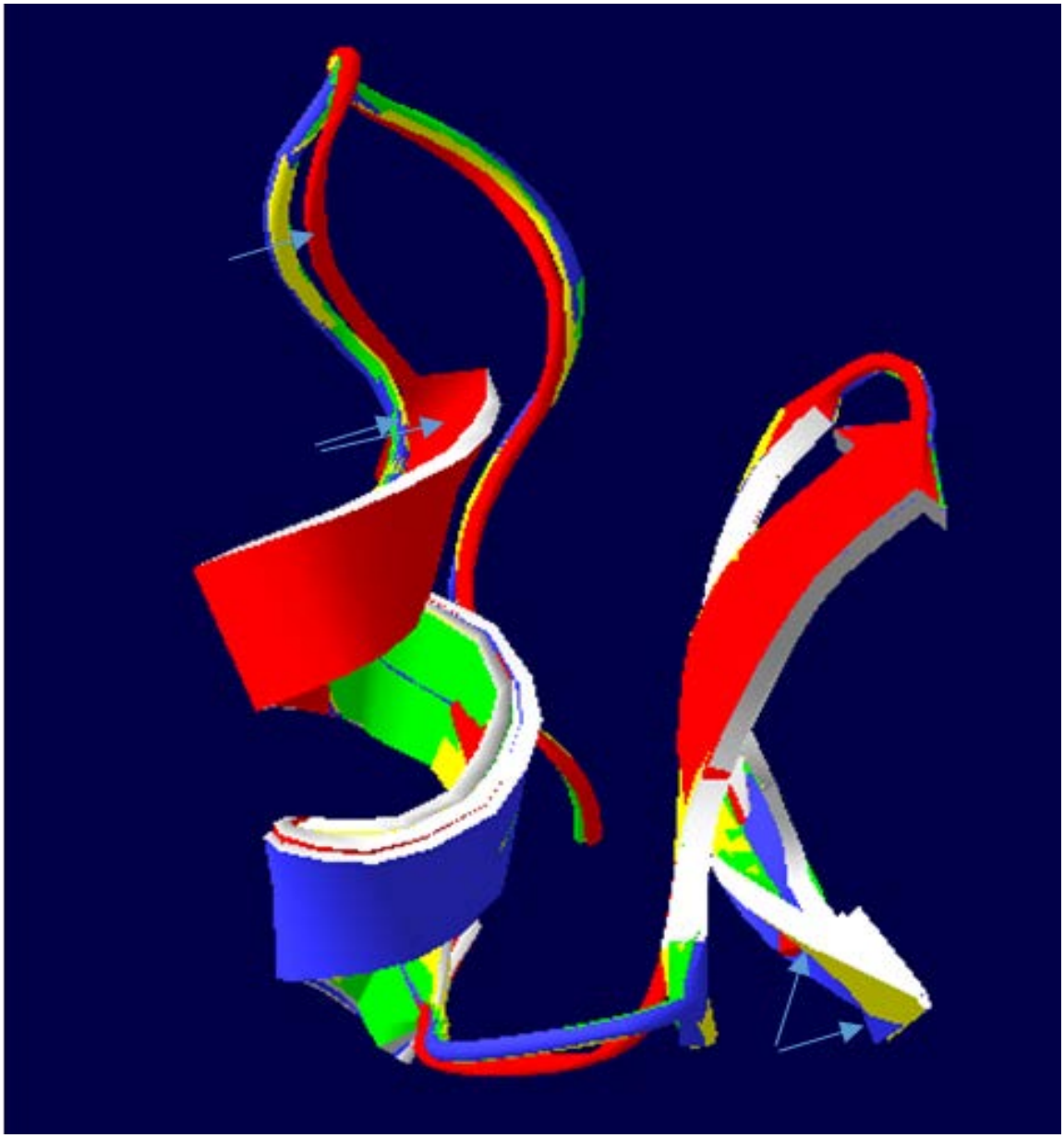

Figure 4.3. Superimposition of Cltx, CA4, CTX23 and Cltx-like peptide AaCTX represented by ribbon display. Superimposed structures of Cltx (yellow), CA4 (blue) and CTX 23 (green) and Cltx-like peptide AaCTX (red) is shown in Figure above. Blue arrows show difference in structures of different toxins.

Stick representation of superimposed models of Cltx-like short toxins is shown in Fig 4.4. Important residues that are conserved in Cltx, CA4, CTX23 and Cltx-like peptide AaCTX are labelled, and location of amino acids overlay accurately. The locations of amino acids differ between the four peptides are indicated with red arrows.

Aromatic residue Phe6 (Phe5 in CA4) overlaid accurately in four short toxins. Cltx-like peptide AaCTX contains Asn9 that is overlaid accurately on Asp9 of Cltx, CTX23 and Asp8 of CA4. A non-polar residue Ala14 and positive residue Lys15 of Cltx-like peptide AaCTX superimposed on Arg14, Lys15 of Cltx, CTX23 and Arg13, Lys14 of CA4 toxin (Fig 4.4). Notably, Arg23 of Cltx-like peptide AaCTX is not overlaid on Lys23 of Cltx, CTX23 or Lys22 of CA4. Negative residues such as Asp16 of CA4 and Asp17 of Cltx are overlaid perfectly, Where, Ser is present at position 17 in CTX23 and Asn is present at position 17 in Cltx-like peptide AaCTX. All residues are overlaid at similar position as labelled by red arrows in Fig 4.4. Cltx, CTX23 and CA4 has a negative residue at position 18 that is superimposed at similar place to each other and Ala18 of Cltx-like peptide AaCTX. Arg25 is superimposed perfectly, but there is slight deviation between aromatic ring of Tyr29 of Cltx, CTX23 and Phe28 of CA4. There is a slight difference between Arg28 of Cltx-like peptide AaCTX and Phe28 of CA4. Overall, amino acids residues of all 3D models superimpose accurately hence representing similar 3D structure.

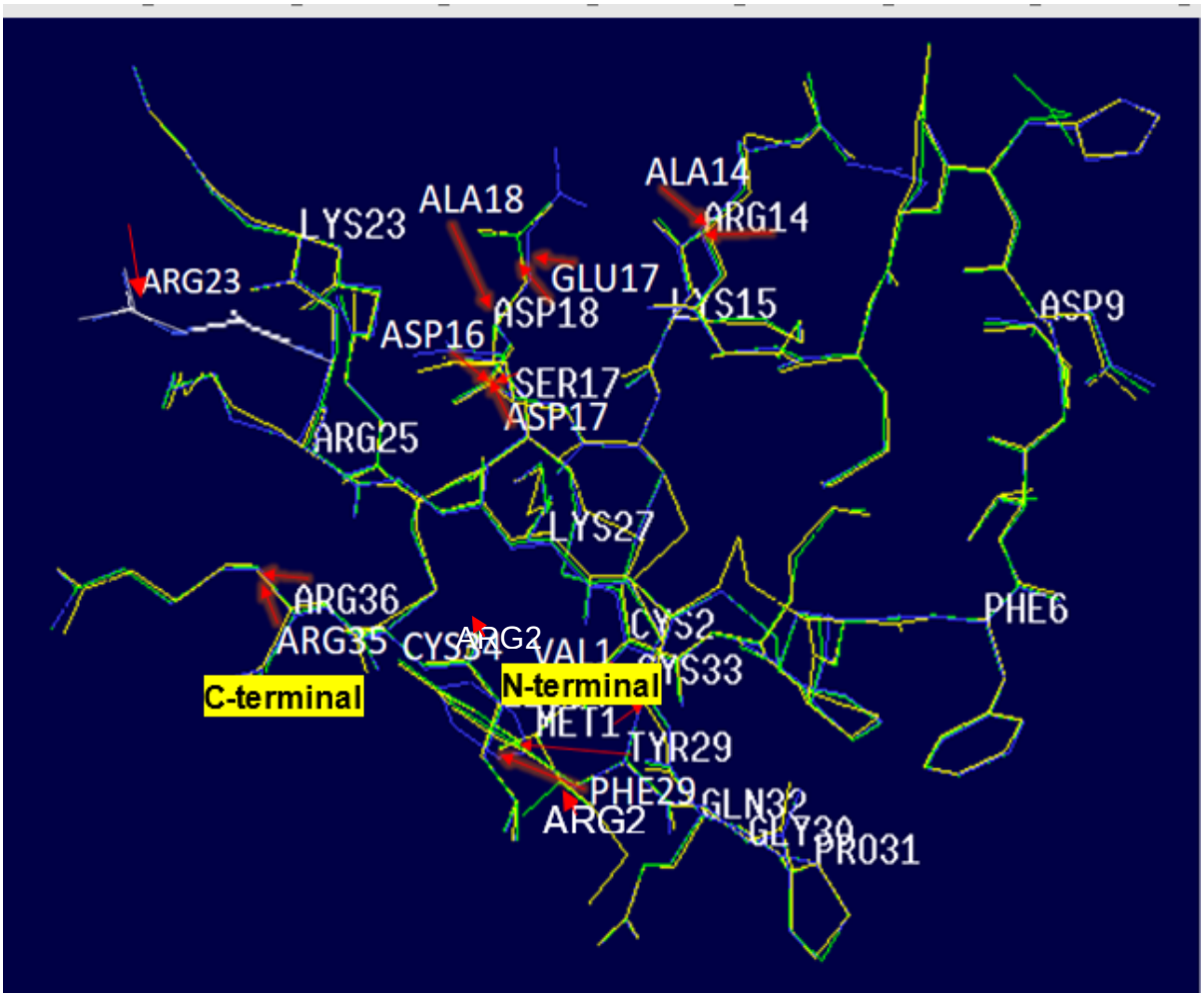

Figure 4.4. **Superimposition of 3D structures of Cltx, CA4, CTX23 and Cltx-like peptide AaCTX by Swiss Deep-View**. 3D structure of Cltx is coloured yellow, 3D models of CA4, CTX23 and Cltx-like peptide AaCTX are coloured green, blue and grey respectively. Important conserved amino acids in all four peptides are labelled, different amino acids of each toxin that are superimposed at similar position identified using arrows red arrows.

The electrostatic surfaces of Cltx, CA4, CTX 23 and Cltx-like peptide AaCTX were calculated using Swiss Deep-View. Analysis of electrostatic surfaces of 3D structure of Cltx show a wide distribution of positive charge on its surface created by positive side chains (Arg, Lys) that possibly interact with ion channels and putative MMP2 receptor. Positive residues such as Arg, Lys exposed on α-helix of Cltx, CA4 and CTX23 giving toxins a strong positive charge. A Positive residue Lys27 is located on second β-sheet that is present on solvent accessible face of toxins that may have involvement in interaction with ion channels. Negative residue Asp9 of Cltx, CTX23 or Asp8 of CA4 is responsible for negatively charged surface around N-terminal loop region. However, Cltx-like peptide AaCTX does not contain this negatively charged surface at loop region as it has a polar residue Asn9 instead of a negative residue.

Hydrophobic surface (neutral) can be visualised near loop region before N-terminal domain of α-helix as well as near 3rd β-sheet of toxins. Overall, all 3 toxins have a strong positive electrostatic surface particularly near C-terminal domain of toxins. Both Cltx and CA4 have strong negative region around α-helix due to Asp17Asp18 and Asp16Glu17 but CTX23 has Ser17 and only one negative residue D18 around α-helix hence weaker negative charged surface in this region. In contrast, Cltx-like peptide AaCTX does not have negative region at all around α-helix as it has Asn17 and Ala18 instead of negative residues. Hence this region contains overall positive charged region possibly due to adjacent Arg23. Moreover, Cltx-like peptide AaCTX does not have a negative residue Asp8 or Asp9 but has a polar residue Asn9 at loop region (encircled red in Figure 4.5). Positive charged surface of the C-terminal domain of Cltx, CA4 and CTX is contributed by Arg35 or Arg36 whereas, Cltx-like peptide AaCTX has Cys34 at C-terminal domain and a slight neutral region around C-terminal region. Cltx, CA4 and CTX23 has an aromatic residue at position 28 or 29 contributing towards a slight neutral region. Cltx-like peptide AaCTX contain Arg28

creating a stronger positive charged surface around β-sheets of the toxin. Second β-sheet of Cltx, CA4 and CTX23 contains Lys26 or Lys27 residue hence contributing to a highly positive charged region around this region. The Cltx-like peptide AaCTX contains a polar residue Ser at position 26 thus does not have a highly positive charged surface around this region as compare to Cltx, CTX23 and CA4 toxins. Overall, all four toxins contain a highly positively charged surface area around peptide surfaces. Cltx, CA4 and CTX23 contain negative and neutral regions particularly around loop regions, while Cltx-like peptide AaCTX as a stronger positive charged surface due to lack of negative residues.

4.3.3 Docking of chlorotoxin-like short toxins

Chlorotoxin is a promising toxin for inhibiting glioma metastasis. It is a potential toxin for targeting MMP2 and AnxA2 that are over-expressed in glioma. Cltx homologues have potential use in glioma as they share sequence homology to Cltx.

Chlorotoxin-like short toxins CA4, CTX23 and Cltx-like peptide AaCTX are relatively important and clinically significant toxins. Biological function of these toxins has been studied and they inhibited glioma metastasis (Xu *et al*., 2016). We docked CA4, CTX23 and Cltx-like peptide AaCTX to their putative binding partners that is MMP2 and Annexin A2 in order to study their interaction and their effects with a view that these toxin peptides would help to understand the mechanism of action of Cltx.

4.3.3.1 Molecular docking of Cltx-like short toxins with MMP2.

MMP2 without activation peptide was docked to CTX23, CA4 and Cltx-like peptide AaCTX in order to study toxin interaction with putative receptor. This generated a top cluster of peptide-MMP2 conformations with the highest number of members bound at similar positions indicating near-native solutions. Top clusters and their energy scoring function were obtained in order to analyse accurate prediction of binding affinity of the peptide-protein complex.

Results show that CTX23 binds to MMP2, the top cluster of CTX23-MMP2 complex shows that CTX23 interacts with the catalytic domain of MMP2 (Fig 4.6.1). Interestingly, bound conformation is similar to that of Cltx. CTX23 is bound near α-helices and loop regions of the catalytic domain. CA4 toxin was docked to MMP2 to study their interaction, results show that CA4 binds to MMP2 as the top complex reveals that CA4 is bound to MMP2 at the catalytic domain. Analysis of the top clusters reveals that CTX23 is bound near α-helices particularly helix-b and extended

peptide region of the catalytic domain (Fig4.6.1). Similar binding orientation is favoured by both CTX23 and CA4 to MMP2 in the top clusters. Cltx also binds to similar locus on activated MMP2 like both CA4 and CTX23 do. Cltx-like peptide AaCTX also interacts with MMP2, top1 cluster of Cltx-like peptide AaCTX and MMP2 shows that peptide is interacting with the catalytic domain of MMP2. Although, Cltx-like peptide AaCTX is docked slightly different to MMP2 catalytic domain than CTX23 and CA4 as α-helix of toxin is bound very closely to all 3 α-helices of the catalytic domain (Fig4.6.1).

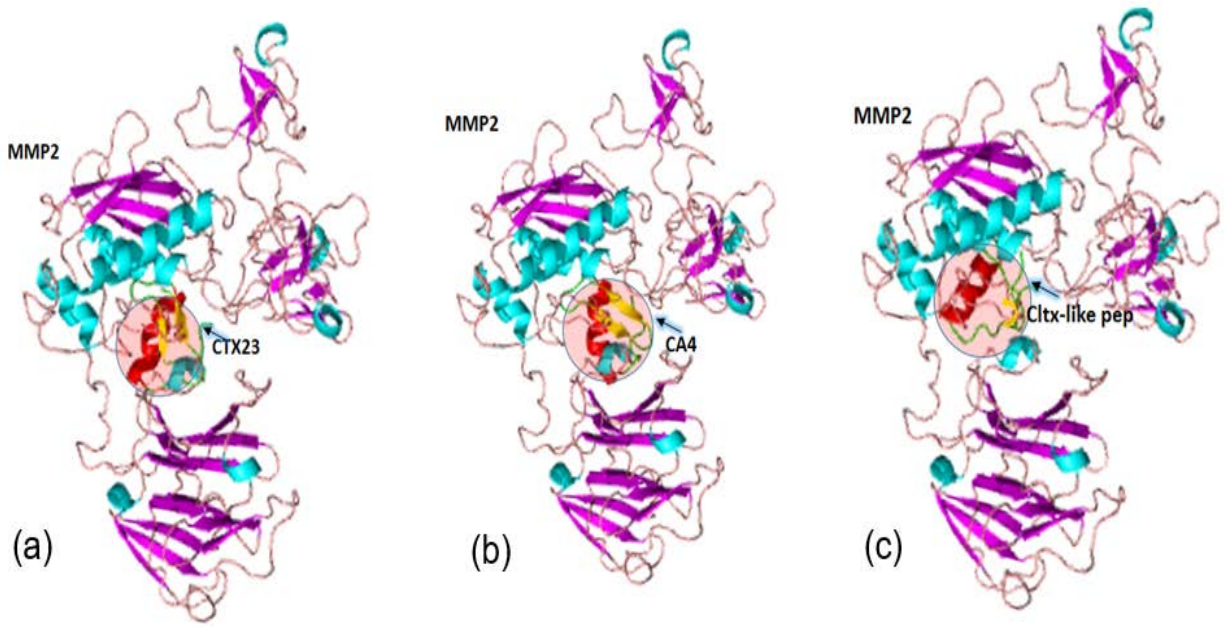

Figure 4.6.1. The top docked clusters of Cltx-like short toxins and MMP2.

MMP2 without activation peptide and point mutation docked to (a) CTX23 toxin peptide, (b) CA4 toxin peptide (c) Cltx-like peptide AaCTX. A-helices of MMP2 are coloured cyan, β-sheets and loops are coloured magenta. Cltx-like short toxins are highlighted by a circle, α-helices of toxins are coloured red while β-sheets are coloured yellow. (a) CTX23 is bound to the catalytic domain of MMP2, interestingly, similar face of CA4 is also bound to the catalytic domain of MMP2. However, orientation of Cltx-like peptide AaCTX is slightly different to both CA4 and CTX23.

Binding between CTX23 and MMP2 the top complex was studied in order to examine amino acids that may have involvement in interaction between peptide and MMP2. The surface constituted by β-sheets and loop region of the C-terminal domain of CTX23 interacts deeply with MMP2. Positive residues of CTX23 particularly, Arg36 is present in close contact with Glu296 of MMP2, Arg25 is in close contact with Asp77. Also, Lys27 of CTX23 is present in close proximity of Glu109 of MMP2. Although, aromatic amino acid Tyr29 of CTX23 is in close contact with Thr318 of MMP2 as shown in Figure below (Fig 4.6.2). Interestingly, Arg14 or Lys15 are situated in free space away from MMP2 target. As shown earlier, CTX23 toxin contains a wide distribution of positive charge around its surface. This indicates the importance of positive residues that possibly have strong binding interaction with negatively charged MMP2. Non-polar residues such as Leu34, Cys35 are also present in proximity of the CAT domain of MMP2.

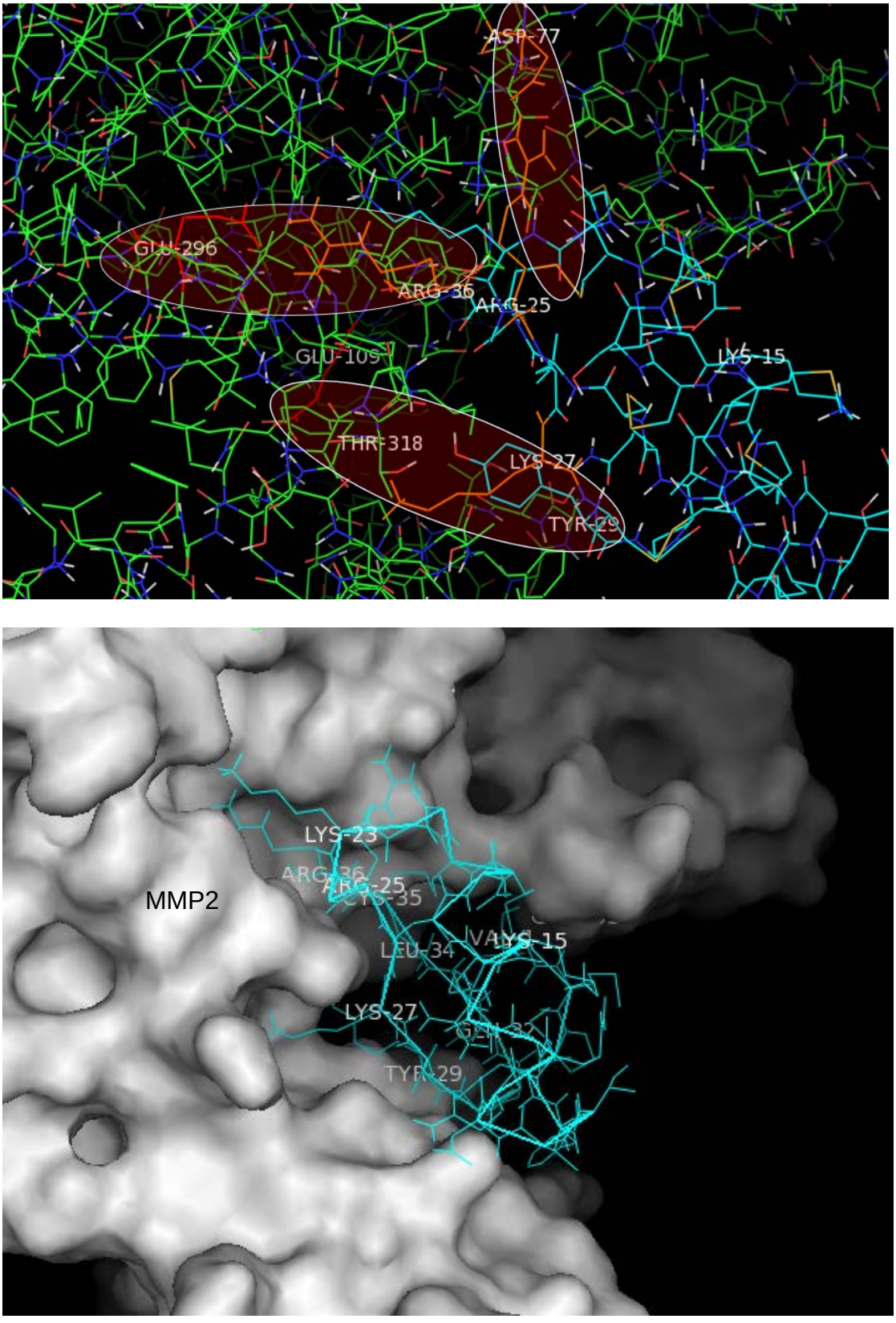

Figure 4.6.2. Interaction the top complex of CTX23 and MMP2 produced by PyMOL. (a) MMP2 is coloured green and CTX23 is coloured turquoise. Possible Interacting residues between CTX23 and MMP2 are represented by circles. Amino acids of CTX23 that are interacting with MMP2 are coloured orange while interacting amino acids of MMP2 are coloured red (b) MMP2 is shown as grey surface and CTX23 is shown as cyan.

Although, positive residues of toxin CTX23 particularly near C-terminal domain are present in close contact with negative residues of MMP2, Arg14 and Lys15 of α-helix does not interact with MMP2 as shown in Fig4.2.2. Majority of residues, particularly positive residues of β-sheets in C-terminal domain are present deeply in pockets of MMP2 interacting with negative residues. Some non-polar residue that are part of C-terminal semi- conserved motif '*GPQCxC'* are also present in close proximity of MMP2.

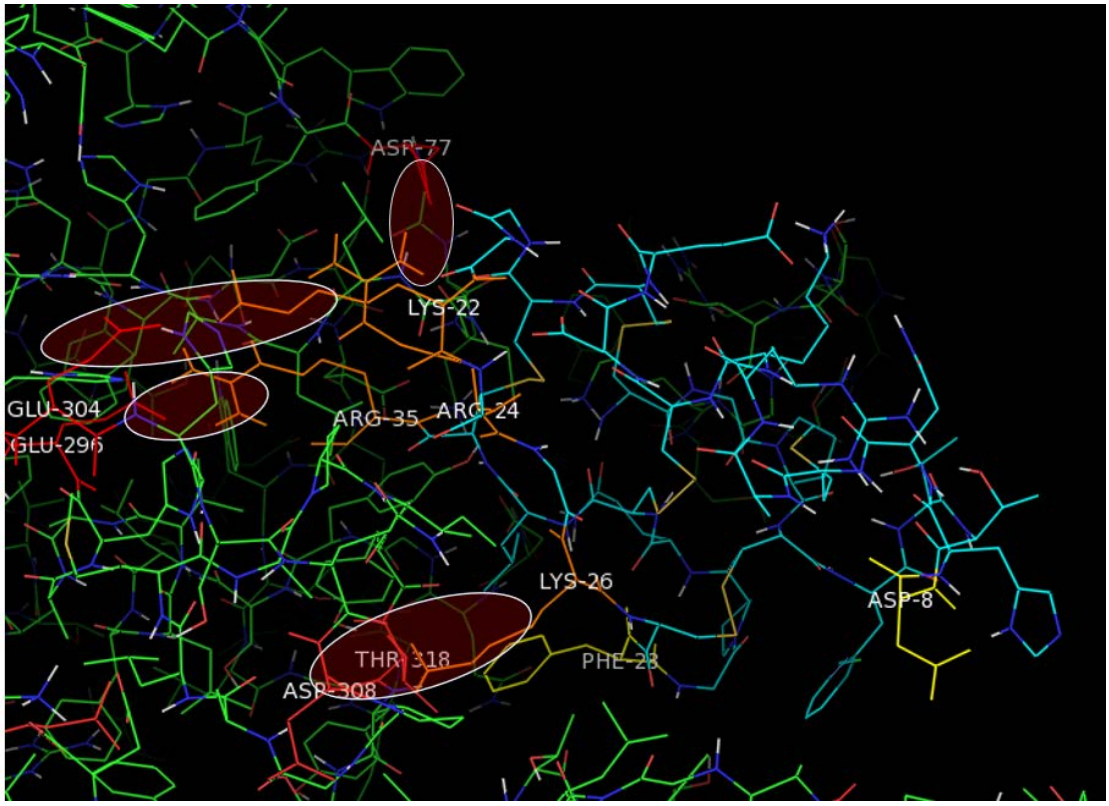

Figure 4.6.3. Show interaction of CA4 peptide with MMP2 target. MMP2 is shown in green and CA4 is shown in cyan. Residues of interest of MMP2 are highlighted in red while residues of toxin are highlighted as orange. Important residues of peptide-protein are highlighted in red circles.

According to the top complex, Arg24 of CA4 is present in close contact of Asp77 of MMP2, Lys22 contacts with GLU304, Arg35 interacts with Glu296 and Lys26 lie close to Asp308 which represent electrostatic attractions between two molecules. Aromatic residue Phe28 of CA4 is close to Thr318 of MMP2. Tyr29 of CTX23 was in close contact with Thr318 of toxin indicating important role of this particular aromatic residue in interaction. Similar to CTX23-MMP2 complex, Arg13 and Lys14 of CA4 that are located in helix region of CA4 does not interact with MMP2.

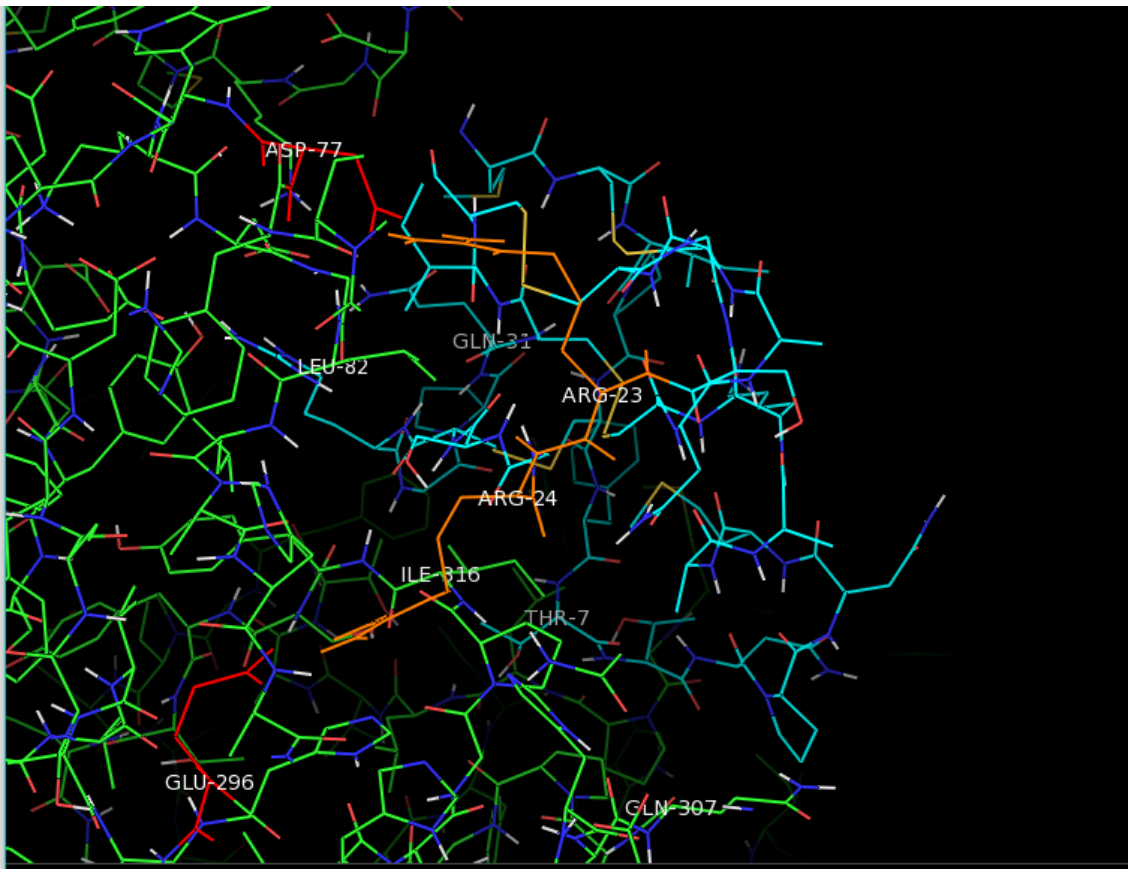

Figure 4.6.4. Interaction of residues between Cltx-like peptide AaCTX and activated MMP2. Cltx-like peptide AaCTX is coloured cyan and residues of interests of the peptide are shown in orange colour. Where, MMP2 target is coloured as green and residues of interests of MMP2 are coloured in red.

The C-terminal domain including β-sheets just before α-helix of the peptide is interacting with MMP2. Although Cltx-like peptide AaCTX does not contain Arg35 or 36 of CA4, CTX23 and Cltx at C-terminus, Arg23 interacting with Asp77 while Arg24 is interacting with Glu296. Hydrophobic residues such as Gly29, Pro32 and Cys30 are buried deep within MMP2 domain. Hence, positive residues do play important part in interaction with MMP2 along with hydrophobic residues on C-terminal domain of the toxin.

4.3.3.2 Energy Scoring of docked complexes of peptide-MMP2.

The energy scoring provides stability and binding affinity of docked complex of peptides and target. Most likely, near-native confirmation is determined by cluster size, that is the highest number of docked neighbour structures within a cluster and with the lowest energy. Lowest energy regions tend to predict bigger cluster of docked complexes. The top cluster contain large cluster energy profile table of top docked cluster is given below. According to data, the top cluster of CA4 has highest number of docked structures that include 261 members and the energy of the cluster centre is -762. CA4/MMP2 docked structure in cluster 1 has the lowest energy of -965.5. Lowest energy of this structure is within similar range of Cltx-MMP2 complex which has lowest energy of -956.50. Although, cluster size of Cltx-MMP2 complex is smaller than cluster size of CA4-MMP2 complex. CA4, like Cltx is bound stably at the CAT domain of MMP2.

Toxin peptide	Cluster	Members	Representative	Weighted Score
CTX23	1	165	Centre	-680
			Lowest energy	-850
CA4	1	261	Centre	-762
			Lowest energy	-965.5
Cltx-like peptide AaCTX	1	127	Centre	-688
			Lowest energy	-802
Cltx	1	183	Centre	-709.5
			Lowest energy	-956.5

Table 4.6.5. Energy score profile of top1 clusters of docked toxin-MMP2 structures. Members represents size of the docked complexes, that is number of docked structures within a cluster, energy of cluster centre represents energy of docked complex that has the highest number of members or neighbours in the cluster and energy of the docked complex in cluster with lowest energy.

The top cluster of the CTX23-MMP2 complex show stable binding affinity with large number of members present in cluster1. Cluster 1 of CTX23-MMP2 contain 165 members with centre energy of -680 and lowest energy of structure is -762. CTX23 also has stable binding to MMP2, though CA4 and Cltx binds strongly to the CAT domain of MMP2.

Although, Cltx-like has a strong binding affinity to MMP2 with -688 energy of cluster centre and lowest energy of structure in the cluster is -802, size of cluster is much small than CA4, CTX23 and Cltx. As cluster has smaller size, Cltx-like peptide AaCTX may not represent a strong binding affinity to the CAT domain of MMP2.

4.3.4 Docking of Chlorotoxin-like short toxins to human Annexin A2.

Cltx was shown to bind to MMP2 and thus inhibits glioma proliferation and metastasis (Deshane *et al*., 2003). Later, Kesavan *et al*., (2010) suggested that Cltx binds to AnxA2 thus inhibits angiogenesis and glioma metastasis. Cltx homologues CTX23, CA4 and Cltx-like peptide AaCTX were docked to human annexin A2 to study their binding interaction with annexin A2. 3D structure of human annexin A2 (PDB code: 2HYW, chain A) was determined by X-ray diffraction at 2.10Å by Shao *et al.,* (2006). 2HYW contains homologous repeats (I-IV) arranged in a near-parallel shape, starting from Asn^{31} to Asp^{338}. All homologous repeats contain 5 α-helices and are structurally similar to each other making a curve-shaped disk structure (Shao *et al.,* 2006). 2HYW share structure similarity to human AnxA2 (PDB code: 1XJL), both 2HYW and 1XJL contain 7 Ca bound ligands.

All three Cltx-like toxins bind to AnxA2. Top docked structure that contains the highest number of docked neighbour structures was examined. The top cluster show CTX23 binds to AnxA2, particularly, beginning of the N-terminal domain, β-sheets and the C-terminal domain of CTX23 are interacting with helices of repeat-I and IV of AnxA2 (Fig 4.7). The α-helix of CTX23 does not interact with AnxA2 and the helix region is not located near the surface of AnxA2. Similar binding pose is observed in that of Cltx-AnxA2 complex. Though, the top cluster of CA4 and Cltx-like peptide AaCTX show that both peptides are bound to repeat I and II of AnxA2, however, with different orientation (Fig4.7).

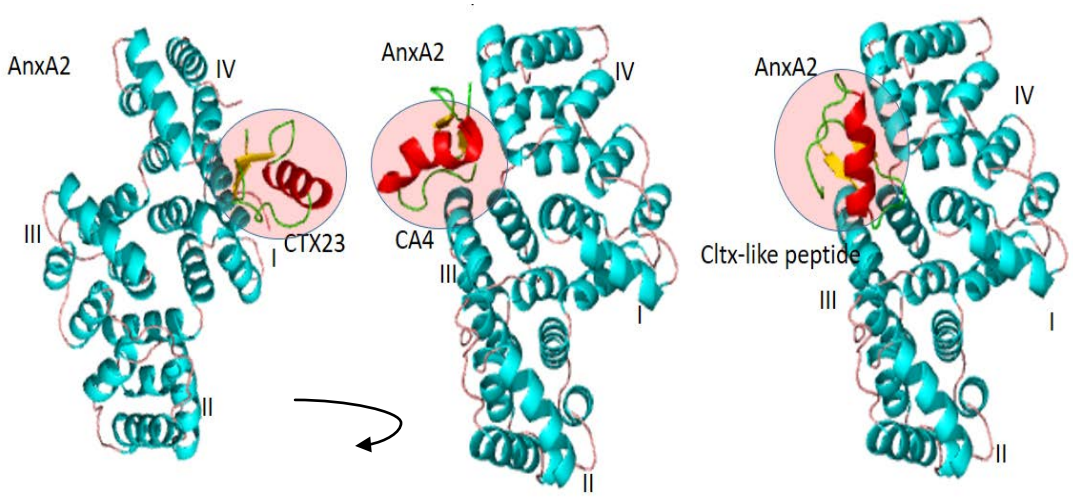

Figure 4.7. Docking of Cltx-like short toxins to human AnxA2. (a) The top bound cluster of CTX23, CA4 and Cltx-like peptide AaCTX with AnxA2, α-helices and loops of AnxA2 are coloured in cyan and magenta respectively. CTX23 is highlighted in a red circle, the α-helix of CTX23 is coloured red and β-sheets are coloured yellow. Black arrow shows that both CA4/AnxA2 and Ctx-like peptide/AnxA2 complexes have a slight orientation difference to CTX23/AnxA2 cluster.

Both CA4 and Cltx-like peptide AaCTX bound to similar region of AnxA2, though bound with different orientation. The CTX23 is bound to repeat I and IV of AnxA2 whereas, CA4 and Cltx-like peptide AaCTX are bound to repeat I and II of AnxA2.

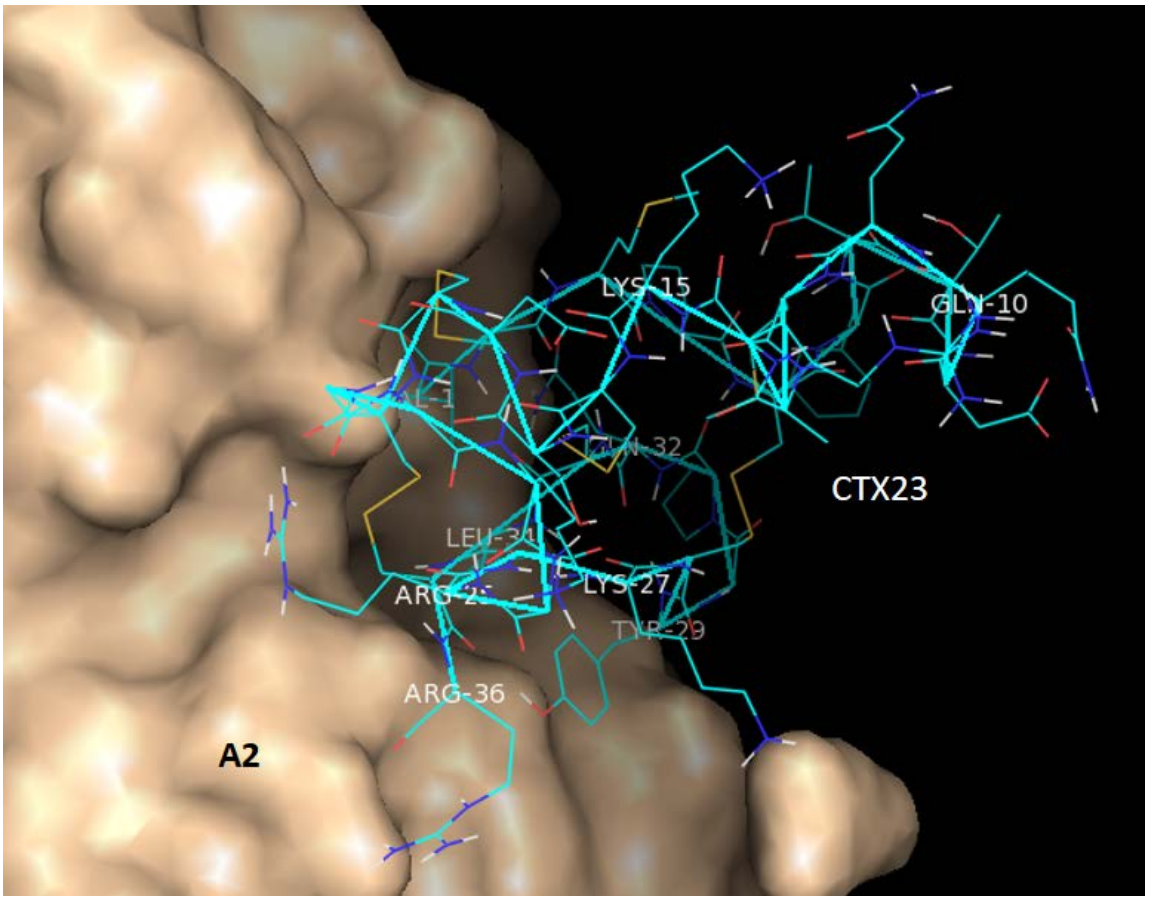

Figure. 4.7.1. CTX23 and AnxA2 interaction in the top complex. Protein surface of AnxA2 is shown in beige and ribbon structure of CTX23 toxin are shown in cyan and amino acids of toxin peptide is labelled.

CTX23 interacts with AnxA2 particularly Met1 of C-terminal domain and hydrophobic residues such as Leu34at Gly30, Pro31 and Leu34 as well as Cys35 of C-terminal domain are interacting deeply with pockets of repeat I and IV of AnxA2. Positive residues such as Arg36 is in close contact with repeat-I and Arg25 and Lys27 are interacting with repeat-IV of AnxA2. Note that Lys15 of toxin does not interact with the peptide, a region before α-helix of CTX23 is freely accessible as it is not interacting with AnxA2 peptide.

Similar docking pose is observed in Cltx-AnxA2 the top confirmation, where Cltx binds to repeat I and IV as well. However, CA4 binds to different region on AnxA2 despite being highly homologous to parent Cltx. Cltx-like peptide AaCTX also binds to AnxA2 at similar repeats as CA4 but both toxins are bound to target with different orientation (Fig4.7.1).

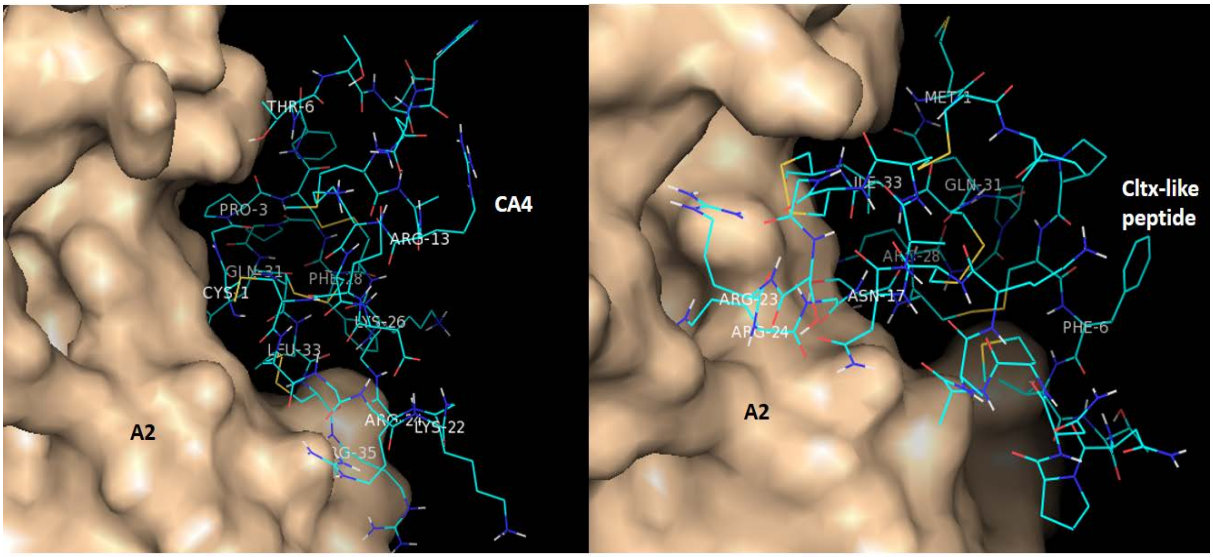

Figure. 4.7.2. Interaction of CA4 and Cltx-like peptide AaCTX to human AnxA2. (a) Surface of AnxA2 is represented by beige colour and stick structure of CA4 is represented by sticks. Amino acids of interests are labelled. (b) Similarly, surface of AnxA2 is shown in beige and stick representation of Cltx-like peptide AaCTX is shown as cyan. Amino acids that are interacting with AnxA2 are labelled.

Both CA4 and Cltx-like peptide AaCTX are docked to AnxA2 with different orientation although both are interacting with repeats I and II of AnxA2. Cys1, Pro3 of CA4 at C-terminal domain are in present in deep cavities of AnxA2, Arg35, Lys26 and Phe28 are also interacting with AnxA2. However, Arg 13 and Lys22 residue of CA4 does not interact with target. In contrast, Arg23 and Arg24 of Cltx-like peptide AaCTX interact with AnxA2 as well as Ile33 and a positive residue Arg28 is interacting with AnxA2.

Binding affinity and stability of these docked confirmation is determined by cluster size and energy scoring. Initial data of energy scoring show that Cltx-like short toxins interacts with AnxA2 although, binding affinity is weaker as compared to that of toxin-MMP2 docked complexes (Fig4.7.2). Similar binding affinity for Cltx-AnxA2 was observed in the top cluster. Weighed energy scores for all the top clusters of short-toxins are shown in table below. Top1 clusters of all short toxins contain largest number of members with centre energy of the centre complex and lowest energy of the structure in the cluster. Results show that CTX23 interacts with AnxA2 with a stable binding complex, the top cluster of CTX23-AnxA2 contain 234 neighbours in cluster centre. Energy of the cluster centre is -507 and lowest energy is -609 of structure complex in the top cluster. Binding stability of CTX23-AnxA2 is in similar range to Cltx-AnxA2 the top cluster which have 289 members and -635.5. However, CA4 has lowest number of member as compared to Cltx and CTX23 when docked to AnxA2, although, all members have stable confirmation with lowest energy of -581. Hence there is no significance difference between energy profiles of the docked complexes. However, there is slight variation in cluster sizes. As mentioned before, the large docked cluster with lowest energy represent near-native structures and have been validated with good acceptable scoring via CAPRI.

Peptide-AnxA2 complex	Cluster	Members	Representative	Weighted Score
CTX23/AnxA2	1	234	Centre	-507
			Lowest energy	-609
CA4-AnxA2	1	168	Centre	-450
			Lowest energy	-581
Cltx-like peptide A AnxA2	1	190	Centre	-475
			Lowest energy	--561.5
Cltx-AnxA2	1	289	Centre	-575.5
			Lowest energy	-635.5

Table 4.8 Energy scoring for Cltx-like short toxins and AnxA2 the top clusters. The top cluster has highest number of members (neighbours docked structures). Members represent size of the docked complexes, that is number of docked structures within a cluster, energy of cluster centre represents energy of docked complex that has the highest number of members or neighbours in the cluster and energy of the docked complex in cluster with lowest energy.

4.5 Discussion

Cltx was initially shown to inhibit small-conductance epithelial chloride channels (DeBin *et al*., 1993). More recently MMP2 and subsequently AnxA2 were proposed potential targets for Cltx (Deshane *et al*., 2003; Kesavan *et al*., 2010). The CA4, CTX23, Cltx-like peptide AaCTX share evolutionary structural relationship and possibly have same molecular targets as there is a high degree of sequence homology between toxins. Although, functional studies of these and other family members are extremely limited or non-existent. Our results of secondary structure analysis show that all toxins share 8 highly conserved cysteine residues that give toxins a compact structure, all toxins consist of three β-sheets packed around α-helix and stabilised by disulphide bridges. Our study shows that electrostatically, all toxins have a highly positive charged surface. Hence, this represent their primary role in interaction with their putative targets. Since, 3D structures of only Cltx and I5A have been determined to date (Arseniev *et al*., 1993; Lippens *et al*., 1995), homology modelling was used to predict 3D models of toxins and then investigated in docking studies. Three toxins were selected and homology modelling show that they share similar structure as majority of residues are conserved in each toxin and residues superimpose well. In the present study, docking results show that structural motif of β-sheets and loop region at the C-terminal domain in CTX23, CA4 as well as Cltx-like peptide AaCTX toxins bound to the targets in majority of top complexes which represents the structural similarities and the way toxin might affect their targets in a similar way *in-vivo.*

Docking results show that CTX23 and CA4 binding with MMP2 is slightly stronger and have large docked members within the top cluster as compared to Cltx-like peptide AaCTX. There is no big difference between binding energy of Cltx-like

peptide AaCTX to other toxins. Although, the number of members in the top cluster is very small as compared other short toxins. Results correlates with the experimental studies performed by Rjeibi *et al*., (2011) as they showed that Cltx-like peptide AaCTX showed less than 10% activity on glioma than Cltx. As mentioned before that the large cluster with lowest energy tend to have a near-native structure in top cluster. Results show that positive residues at the C-terminal domain such as Lys27, Arg36 interact with negative MMP2 target which represent electrostatic interaction between two proteins. Our results show that Cltx-like peptide AaCTX lacks Lys27 and a positive residue at the C-terminal domain which might be involved in weaker binding to its target MMP2 than other Cltx-like short toxins. Hence, lack of some important positive residues in sequence might play a role in weaker binding and smaller cluster size.

Moreover, binding pose of CTX23 and CA4 with MMP2 is similar to Cltx-MMP2 complex which represent a similar binding site on their target protein. Possibility of the evolutionary relationship between these short toxin peptides at the sequence level as well 3D level implies that the interaction model with target MMP2 is conserved.

Our results suggest that the interaction between Cltx-like short toxins and activated MMP2 is much stronger than Cltx-like short toxins interaction with AnxA2. Top clusters of toxin-MMP2 show larger cluster size and very stable complexes within the cluster. There is a difference between energy values of the top clusters of CA4, CTX23 docked to MMP2 than CA4 or CTX23 docked to AnxA2. Soroceanu *et al*., showed that there is a high affinity site on glioma and a lower affinity site and they presumed the presence of more than one cellular targets for MMP2. Based on our results we assume that there is a possibility that the higher affinity binding site is MMP2 and lower affinity site AnxA2. As mentioned before, both MMP2 and AnxA2

are over-expressed in glioblastoma specimen and cells lines (Deshane *et al*., 2003; Lokman *et al*., 2011).

The present study show toxin CTX23 and CA4 interact with their target MMP2, mainly via the β-sheets regions and loops near the C-terminal domain. The second β-sheet contain a highly conserved positive residue Lys27 (conserved in many Cltx-like peptide AaCTX) as well as C-terminal Arg36; they play an important role interaction to MMP2 full-length and the catalytic domain of MMP2 as well as ionic channel (chloride channels). However, positive residues Arg14 and Lys15 do not interact with MMP2 directly. Mutation of these important residues might have significant effects on the way they interact with their targets.

As one can visualise the results in this study, Cltx-like short toxins have a strong binding affinity to activated MMP2. Conserved positive residues especially at the C-terminal domain as well as some hydrophobic residues play an important role in interaction which illustrate importance of electrostatic and hydrophobic interaction between two proteins Our results suggest that Cltx-like short toxins like their parent Cltx has lower binding affinity to AnxA2.

This explains the evolutionarily relationship between these short toxin peptides and significance on their interaction with target MMP2 and AnxA2 is conserved. Both CA4 and CTX23 have tendency to bind to both MMP2 and AnxA2 with strong binding affinity and share similar results as their parent Cltx. Cltx-like peptide AaCTX also has binding effects on MMP2 but lower binding than CA4 and CTX23. CA4 and CTX23 in glioma studies can be further evaluated in order to achieve a peptide toxin with great anti-glioma properties.

5 GENERAL DISCUSSION

Cltx was shown to modulate small-conductance epithelial chloride channels (DeBin *et al*., 1993). Later studies showed that Cltx specifically binds to MMP2 but not MMP1 and MMP9 (Deshane *et al*., 2003). Later, Cltx was shown to facilitate internalisation of the lipid raft-anchored complex which contained MMP2, MT1-MMP, TIMP2 and ClC3 chloride channels thus inhibiting glioma metastasis (Kesavan *et al*., 2010). To complicate matters further, AnxA2 was proposed a putative target for Cltx (Kesavan *et al*., 2010). Although, the mechanism of action of Cltx is unknown. Over the past few years, over 30 Cltx-like short toxins that share sequence homology to Cltx have been identified and they may have therapeutic potential in glioma and other cancers. It is very important to study these Cltx-like short toxins in glioma studies to find a suitable candidate with high potency for effective anti-glioma therapy. Although a few studies have attempted to study role of Cltx in glioma, there is an extremely limited data available about the structure-function relationship of Cltx and Cltx-like short toxins which needs to be investigated in order to understand the mechanism of action of Cltx and achieve a toxin that have higher binding affinity and greater efficacy.

Recently, Cltx homologues CTX23, CA4 and Cltx-like peptide AaCTX have been tested in glioma studies (Rjeibi *et al*., 2010; Xu *et al*., 2016). Although, functional studies of their other family members are extremely limited or non-existent. The present study provides help to understand the mechanism of action of Cltx and provide a help for laboratory studies for a good therapeutic toxin selection.

All Cltx-like short toxins adopted similar 3D structures identical to existing solved structures of Cltx and Insectotoxin I5A, suggesting their evolutionary relationship. Subsequent quality validation of predicted models by QMEAN6, Z-Score and

Ramachandran plots showed that all models had good quality and were comparable to relevant existing solved crystal structures.

The present study show that Cltx has a higher binding affinity to native activated MMP2 than full-length MMP2. Results suggest that Cltx binds to the catalytic domain of MMP2 with a stronger binding interaction which might interfere the enzymatic activity of MMP2. Binding of Cltx to native activated MMP2 generated a top cluster containing 183 members and the cluster had a docked member with the lowest energy of -956.50 and a centre docked structure with the energy of -709.50. In contrast, the top docked cluster of Cltx and latent MMP2 consisted of 135 members which is a smaller cluster as compared to the top cluster of Cltx and native activated MMP2. Similarly, CA4, CTX23 and Cltx-like peptide AaCTX were bound near the active site of native activated MMP2 but with slightly different orientation. Interestingly, binding interaction of CA4 was very stronger than CTX23 and Cltx-like peptide AaCTX. The top cluster of native activated MMP2 and CA4 had a large size compared to Cltx, CTX23 and Cltx-like peptide AaCTX. The top cluster contained 261 members and a member within that top cluster had the lowest energy of -965.50, suggesting a stronger binding interaction and a stable cluster. However, Cltx-like peptide AaCTX showed a weaker binding to activated MMP2 than CA4, CTX23 and Cltx. Our results suggest that Cltx, CA4 and CTX23 might block the catalytic site and thus inhibit enzymatic activity of MMP2.

The present study suggest that Cltx has a higher binding affinity for MMP2 and lower for AnxA2. This supports a study conducted by Sorocenau *et al.,* (1998) who proposed that Cltx has a high and a low affinity binding site on glioma cells. The present study also illustrates that Cltx and Cltx-like short toxins show binding interaction with MMP2 on a similar site but different sites on AnxA2 which suggests a high specificity and selectivity of toxins for MMP2. Thus, future studies can pay attention to the interaction of Cltx with MMP2 rather than AnxA2.

Analysis of residue interactions show that electrostatic interaction plays an important role in binding as the positive residues of Cltx and Cltx-like short toxins are present in close contact with negative residues of MMP2. A wide positively charged surface of Cltx and CTX23, CA4 and Cltx-like peptide AaCTX is created by positive residues such as Arg and Lys that possibly interact with negatively charged MMP2.

Ullrich *et al*., (1993) showed that Cltx block glioma specific Cl^- channels in acute slices of human glioma. Based on the electrostatic surface experiments in this study we can assume that charged residues may play an important role in inhibition of Cl^- channels. Cltx may have interacted with the Cl^- ion channels therefore likely that they would block a chloride channels by making a conformational change to the selectivity filter of a chloride channel rather than plugging it, which is positively charged. A study conducted by Jentsch *et al*., (2002) showed that ChTX plugs K^+ channel thus blocking the activity of the channels.

Our results show that the initial screen of all 34 Cltx-like short toxins docking suggests that they interact with MMP2. CA4 toxin has a strong binding interaction with activated MMP2 and Cltx-like peptide has a weak interaction. Although both MMP2 and AnxA2 have been proposed putative binding partners for Cltx, interaction of Cltx and Cltx-like short toxin is stable and stronger with MMP2 than AnxA2. The present study provides a help in selecting Cltx-like short toxins that have higher binding affinity with its putative target MMP2 and thus selecting toxins with higher efficacy for future laboratory studies.

5.1 Limitation of current study and future directions.

Protein structure is more stable and alters more slowly than its associated amino acid sequence during evolutionary process (Bloom *et al*., 2004). Evolutionary relationship between two proteins at the sequence level produces similar structures (Bloom *et al*., 2004). Studies showed that two completely different functioning

proteins may have similar structure due to evolution (Krieger *et al*., 2003). Homology modelling is an accurate and reliable method to predict 3D models of a protein from its amino acid sequence, using an existing crystal structure (template) for an evolutionary related protein (target) (Moult *et al*., 2014).

Although homology modelling is a very reliable method and validated by CASP (see 1.4), loop modelling in protein structure prediction can be challenging because alterations in loop conformations are hardly predicted in homology modelling (Jamroz and Kolinski, 2010). Loops are involved in structural and biological processes such as protein-protein interactions, signalling cascades and enzyme catalysis (Espadaler *et al*., 2006). Insertions and deletions occurs frequently at loop regions and turns than the conserved α-helix or β-sheets. Even without insertions or deletions, loops may induce a conformational change in a target protein due to involvement of the surface loops in crystal surface contact or exchange of a side chain in loop or either mutation of Pro or Gly residue (Krieger *et al*., 2003). Fortunately, loop modelling was easier in 3D model prediction of Cltx-like short toxins by using existing solved crystal structures of Cltx and Insectotoxin I5A from PDB-databank.

Protein-protein docking is performed to investigate the interaction between two proteins or between a drug and a protein (Coelho *et al*., 2016). Recently, docking has been performed extensively to analyse binding affinities of a protein or a drug on its target molecule which provides useful information of enzyme-inhibitor interactions, drug design screen etc (Kozakov *et al*., 2013; Coelho *et al*., 2016). Although, few problems are generally associated with docking as predicted complexes may give an inaccurate binding sites of the target protein, screening using small molecule database and high binding affinity but failed subsequent MD (Chen *et al*., 2015). A ligand binding to its receptor do not clarify whether it is an agonist or an inhibitor hence some results may be validated by performing bioassays (Chen *et al*., 2015).

Therefore, homology modelling and docking are assessed by CASP and CAPRI experiments as mentioned in introduction section.

ClusPro has been widely used in protein-protein docking in the past few years and results generated by this programme have been used in more than 400 publications (Kozakov *et al*., 2017). The results generated by ClusPro have been validated by experiments such as site-directed mutagenesis with NMR, calorimetry, FRET, or surface plasmon resonance, cross-linking, spectroscopy and X-ray scattering, electron self-exchange reaction, radiolytic protein footprinting with mass spectrometry, hydrogen/deuterium exchange, or intermolecular NOE restraints (Sondermann *et al*., 2005; Guzman *et al*., 2009; Man *et al*., 2010; Kozakov *et al*., 2017). ClusPro performs a global sampling search on a grid to produce highly populated clusters with lowest energy, this approach is useful if there is no prior knowledge of the interacting site or residues available for two proteins. It generates 1000 docked structures, clusters them and refine them by CHARMM minimisation to remove steric clashes and produces near-native structures (Kozakov *et al*., 2013).

One major limitation of the programme is the rigid body assumption that limits the applicability of the programme (Kozakov *et al*., 2004). ClusPro docking programme considers ligand and receptor proteins as rigid body molecules and do not allow a conformational change in active site upon binding (Kozakov *et al*., 2007). MD simulation studies have shown that sometimes binding of a ligand to receptor may introduce a conformational change in the active site (Mobley *et al*., 2007). Protein-protein docking benchmark also uses certain difficult bound complexes that allow conformational change in the backbone geometry upon binding. ClusPro or any other rigid body docking programme cannot provide acceptable solutions if this type of shift occurs upon binding (Kozakov *et al*., 2017). However, rigid body docking is useful in producing good acceptable solutions in many enzyme-inhibitor docking (Kozakov *et al*., 2007). Upon binding, enzyme-inhibitor complexes show a very small

conformational change in backbone geometry and restricted to side chains or loops (Kozakov *et al*., 2007). However, neglecting a small conformational change can lead to errors in binding affinities which may jeopardise the accuracy of binding interactions (Mobley *et al*., 2007). Flexible docking programmes such as ROSETTA and HADDOCK allow conformational change between a ligand and receptor protein upon binding but are very time consuming and requires a *priori* knowledge and list of interacting residues. Since there is a very limited knowledge about Cltx interaction with its proposed putative binding partners, Rosetta and HADDOCK docking programmes are not useful in this study.

Another limitation of the ClusPro is that the docking does not account for non-standard or modified amino acids for example by phosphorylation. Furthermore, the docking may also experience difficulties for certain complexes when it comes to cofactors, which affects protein interactions. Although, continuous work is being carried out to remove these technical shortcomings soon (Xia *et al*., 2016; Kozakov *et al*., 2017). MD simulations are computationally expensive and requires an extensive computational throughput and generally performed at industry level (Mobley *et al*., 2009).

ClusPro results can be used in laboratory validation experiments. Such combination of *in-silico* and laboratory studies can provide most meaningful use of docking.

Enzyme assays can be performed in laboratories to study the interaction of Cltx and Cltx-like short toxins such as CA4 with MMP2. By using this method, the interaction of Cltx with different domains of MMP2 can be studied. Also, the activity of Cltx with the active site of MMP2 can be investigated. In addition, co-localisation studies can be performed to investigate interaction between fluorescently labelled Cltx or Cltx-like short toxins and MMP2 or AnxA2. Furthermore, MMP2 or AnxA2 knockout experiments can be performed to inhibit binding and interaction between Cltx and MMP2 or AnxA2.

Single nucleotide polymorphism (SNP) variations of MMP2 can alter the interaction of Cltx and MMP2 on glioma cells which can affect cancer treatment. SNP variants of MMP2 can be produced *in-silico* and tested against Cltx to find binding affinity and the interaction of two molecules. *In-silico* screening of Cltx and Cltx-like short toxins against MMP2 SNP variants may provide a new therapeutic toxin that have higher binding affinity and greater efficacy. Computational screening of other Cltx-like short toxins may reveal a highly potent and more effective toxin in glioma treatment than Cltx. By using the knowledge of mutation, a better Cltx-like short toxin can be modelled computationally that have stronger binding. The design of the toxin can be used to improve Cltx interactions with MMP2, other MMPs and AnxA2 for effective glioma therapy.

6 REFERENCES

- ALAM, M. *et al*., (2012). Design and synthesis of a peptidyl-FRET substrate for tumor marker enzyme human matrix metalloprotease-2 (hMMP-2). Int. J. *Pept. Res. Ther*. **18** (1) 207–215.
- ALI, A. S. *et al*., (2016). Structure-Activity Relationship of Chlorotoxin-Like Peptides. *Toxins*, **36** (8), 1-18.
- ALLAN, J. A., *et al*., (1995). Binding of gelatinases A and B to type-I collagen and other matrix components. *Biochem J*. **309** (1), 299–306.
- ARSENIEV, A. S. *et al*., (1993). Spatial structure of insectotoxin i5a buthus eupeus by 1h nuclear magnetic resonance spectroscopy (russian). *Bioorg.Khim.* **17** (1), 1613-1632.
- BADIGA, V. A. *et al*., (2011). MMP-2 siRNA Inhibits Radiation-Enhanced Invasiveness in Glioma Cells. *PLoS ONE*, **6** (6), e20614.
- BAO, H., *et al*., (2009). “Overexpression of Annexin II affects the proliferation, apoptosis, invasion and production of proangiogenic factors in multiple myeloma,” *International Journal of Hematology*, **90**, 177–185.
- BAYAT, N., *et al*. (2016). Apoptotic effect of atorvastatin in glioblastoma spheroids tumor cultured in fibrin gel. *Biomedicine & pharmacotherapy = biomedecine & pharmacotherapie*.
- BECKNER, M. E. *et al*., (2005). Proteomic characterization of harvested pseudopodia with differential gel electrophoresis and specific antibodies. *Lab Invest.* **85** (1) 316–327.
- BENKERT, P., *et al*., (2011). Toward the estimation of the absolute quality of individual protein structure model. *Structural Bioinformatics,* **27** (3), 343-350.
- BERNARDO, M. *et al*., (2003). TIMP-2 (tissue inhibitor of metalloproteinase-2) regulates MMP-2 (matrix metalloproteinase-2) activity in the extracellular

environment after pro-MMP-2 activation by MT1 (membrane type 1)-MMP. *Biochem J,* **374** (3), 739-745.

- BESCOND, A. *et al*., (1999). Influence of homocysteine on matrix metalloproteinase-2: activation and activity. *Biochem Biophys Res Commun.* **263** (2) 498-503.
- BIASINI, M., *et al*., (2014). SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acid Research,* **42** (1), 252-258.
- BLOOM, D. J., *et al*., (2004). Stability and the Evolvability of Function in a Model Protein. *Biophys J.* **86** (5), 1304146.
- BLUME, *et al*., (2013). Extracranial glioblastoma with synchronous metastases in the lung, pulmonary lymph nodes, vertebrae, cervical muscles and epidural space in a young patient - case report and review of literature. *BMC Research notes,* **6** (1), 290.
- BONTEMS, F. *et al*., (1992). Analysis of side-chain organization on a refined model of charybdotoxin: structural and functional implications. *Biochemistry,* **31** (1), 7756-7764.
- BORDOLI, L., *et al*. (2009) Protein structure homology modeling using SWISS-MODEL workspace. *Nat. Protocols*, **4** (1), 1-13.
- BRANTLEY, E.C., *et al*., (2008). Loss of protein inhibitors of activated STAT-3 expression in glioblastoma multiforme tumors: Implications for STAT-3 activation and gene expression. *Clin. Cancer Research* 14 (10) 4694–4704.
- BRIKNAROVA, K. *et al*., (1999). THE SECOND TYPE II MODULE FROM HUMAN MATRIX METALLOPROTEINASE 2. *Structure,* **7** (1), 1235-1245.
- BRIKNAROVA, K. *et al*., (2001). The Third Fibronectin Type II Module from Human Matrix Metalloproteinase 2. *J Biol Chem* **276**, 27613-27621.

- BUCKNER, J. C. *et al*. (2003). Factors influencing survival in high-grade gliomas. *Semin Oncol.* **30** (1), 10-14.
- BUTTE, P. V., (2014). Near-infrared imaging of brain tumors using the Tumor Paint BLZ-100 to achieve near-complete resection of brain tumors. *Neurosurg Focus*, **36** E1.
- CAO, J. *et al*., (1995). The C-terminal region of membrane type matrix metalloproteinase is a functional transmembrane domain required for pro-gelatinase A activation. *J Biol Chem.* **270**, 801-805.
- CHAKRABORTI, S. *et al*., (2003). Regulation of matrix metalloproteinases: an overview. *Mol Cell Biochem.* **253** (1) 269-85.
- CHEN, C-Y., (2015). Beware of docking, *Trends in Pharmacol. Sci*. **36** (2) 78-95.
- CHEN, Z. *et al*., (2016). *In vitro* and *in vivo* magnetic resonance imaging with chlorotoxin-conjugated superparamagnetic nanoprobes for targeting hepatocarcinoma. *Oncol Rep.* **35** (5), 3059-67.
- CHOE, G. *et al*., (2002). Active matrix metalloproteinase 9 expression is associated with primary glioblastoma subtype. *Clin Cancer Res.* **8** (9), 2894-901.
- COELHO, E. D *et al*., (2016). Computational Discovery of Putative Leads for Drug Repositioning through Drug-Target Interaction Prediction. *PLoS computational biology,* **12** (11), e1005219.
- DAI, B., *et al*., (2007). Aberrant FoxM1B expression increases matrix metalloproteinase-2 transcription and enhances the invasion of glioma cells. *Oncogene* **26** (1), 6212-6219.
- DAI, B., *et al*., (2010). FoxM1B regulates NEDD4-1 expression, leading to cellular transformation and full malignant phenotype in immortalized human astrocytes. *Cancer Res*. **70** (1), 2951–2961.

- DARDEVET, L., *et al*., (2015). Chlorotoxin: A Helpful Natural Scorpion Peptide to Diagnose Glioma and Fight Tumor Invasion. *Toxins*, **7**, 1079-1101.
- DAS, S. and MARSDEN, P. A. (2013). Angiogenesis in Glioblastoma. *N engl J med,* **369** (16), 1561-1563.
- DEBIN, J. A. *et al*., (1993). Purification and characterization of chlorotoxin, a chloride channel ligand from the venom of the scorpion*. Am J Physiol.* **264** (2), 361-9.
- DESHANE, J., GARNER, C. C. and SONTHEIMER, H. (2003). Chlorotoxin inhibits glioma cell invasion via matrix metalloproteinase-2. *The journal of biological chemistry,* **278** (6), 4135-4144.
- DHANARAJ, V. *et al*., (1999). CRYSTAL STRUCTURE OF GELATINASE A CATALYTIC DOMAIN. *CROATICA CHEMICA ACTA*, **72** (2), 576-591.
- DIAZ, V. M. *et al*., (2004). Specific interaction of tissue-type plasminogen activator (t-PA) with annexin II on the membrane of pancreatic cancer cells activates plasminogen and promotes invasion *in vitro*. *Gut*, **53** (1), 993–1000.
- ESPADALER, J., *et al*., (2006). Identification of function-associated loop motifs and application to protein function prediction. *Bioinformatics,* **22** (1), 2237-2243.
- FARSANI, M. F. *et al*., (2015). Molecular Modeling and Docking Studies on the First Chlorotoxin-Like Peptide from Iranian Scorpion Mesobuthuseupeus (MeICT) and SNP Variants of Matrix Methaloproteinase-2 (MMP-2). *Iranian Journal of Toxicology.* **9** (30), 1368-1376.
- FLOUDAS, C. A., *et al*., (2006). Advances in protein structure prediction and de novo protein design. *Chem Eng Sci.,* **61** (1), 966-88.
- FROY, O. *et al*., (1999). Dynamic diversification from a putative common ancestor of scorpion toxins affecting sodium, potassium, and chloride channels*. J Mol Evol*. **48** (2) 187-96.

- GANAU, L., *et al*. (2015). Management of Gliomas: Overview of the Latest Technological Advancements and Related Behavioral Drawbacks. *Behavioural neurology,* **2015**, 862634.
- GEHRMANN, M. *et al*., (2002). The First Fibronectin Type II Module from Human Matrix Metalloproteinase 2. *J Biol Chem.* **383** (1), 137-148.
- GINGRAS, D. *et al*., (2000). Rapid activation of matrix metalloproteinase-2 by glioma cells occurs through a posttranslational MT1-MMP-dependent mechanism. **1497** (3), 341-350.
- GOHLKE, U. *et al*., (1996). The C-terminal (haemopexin-like) domain structure of human gelatinase A (MMP2): structural implications for its function. *FEBS,* **387** (1), 126-130.
- GOLDBRUNNER, H. R., *et al*., (1999). Cell-extracellular matrix interaction in glioma invasion. *Acta Neurochir,* **141,** 295-305.
- GUZMAN, L*., et al.,* (2009). Blockade of ethanol-induced potentiation of glycine receptors by a peptide that interferes with Gbetagamma binding. *J. Pharmacol. Exp. Ther*. **331,** 933–939.
- HAGEMANN, C., *et al*. (2010). Comparative expression pattern of Matrix-Metalloproteinases in human glioblastoma cell-lines and primary cultures. *BMC Res Notes,* **3,** 293.
- HAN, L. *et al*., (2016). Targeted drug delivery to ischemic stroke via chlorotoxin-anchored, lexiscan-loaded nanoparticles. *Nanomedicine.* **12** (7), 1833-1842.
- HANGAUER, D. G. *et al.,* (1984). An interactive computer graphics study of thermolysin-catalyzed peptide cleavage and inhibition by N-carboxymethyl dipeptides. *Biochemistry,* **23** (2), 5730-41.

- HSIAO, M. H. *et al*., (2015). Hexanoyl-Chitosan-PEG Copolymer Coated Iron Oxide Nanoparticles for Hydrophobic Drug Delivery. *ACS Macro Lett.* **4** (4), 403-407.
- HUR, J. H., *et al*., (2000). Matrix metalloproteinases in human gliomas: activation of matrix metalloproteinase-2 (MMP-2) may be correlated with membrane-type-1 matrix metalloproteinase (MT1-MMP) expression. *J Korean Med Sci.,* **15** (3), 309-314.
- ICHIKAWA, T. *et al*., (2014). Annexin a2 regulates angiogenesis and invasion phenotypes of malignant glioma. *Neuro Oncol*. **16** (3) iii10.
- JAMROZ, M. and KOLINSKI, A. (2010). Modeling of loops in proteins: a multi-method approach*. BMC Structural Biology,* **10** (5), 1-9.
- JENTSCH, T. J., *et al*. (2002). Molecular structure and physiological function of chloride channels. *Physiol Rev.* **82** (2), 503-68.
- JIANG, Y. *et al*., (2016). An expanded evaluation of protein function prediction methods shows an improvement in accuracy. *arXiv preprint arXiv*, **160** (1), 00891.
- JONES, P. G. *et al*., (1992). A nonapeptide to the putative F-actin binding site of annexin-II tetramer inhibits its calcium-dependent activation of actin filament bundling. *J. Biol. Chem*. **267** (1) 13993–13997.
- KANDASAMY, D. A., *et al*., (2010). Matrix metalloproteinase-2 and myocardial oxidative stress injury: beyond the matrix. *Cardiovascular Research,* **85** (3), 413-423.
- KARGIOTIS, O., *et al*. (2008). Adenovirus-mediated transfer of siRNA against MMP-2 mRNA results in impaired invasion and tumor-induced angiogenesis, induces apoptosis in vitro and inhibits tumor growth in vivo in glioblastoma. *Oncogene,* **27,** 4830-4840.

- KESARI, S. *et al*., (2011). Understanding glioblastoma tumor biology: the potential to improve current diagnosis and treatments. *Semin Oncol,* **38** (4) 2-10.
- KESAVAN, K., *et al*. (2010). Annexin A2 is a molecular target for TM601, a peptide with tumor-targeting and anti-angiogenic effects. *J Biol Chem*, **285** (7), 4366–4374.
- KHABIRI, M. *et al*., (2011). Charybdotoxin unbinding from the mKv1.3 potassium channel: a combined computational and experimental study. J Phys Chem B. **115** (39), 11490-500.
- KLING, T. *et al*., (2016). Integrative Modeling Reveals Annexin A2-mediated Epigenetic Control of Mesenchymal Glioblastoma. *EBioMedicine,* **12** (1), 72-85.
- KORNETA, I., *et al*. (2012). Structural bioinformatics of the human spliceosomal proteome. *Nucleic Acid Res.* **40** (15), 7046-7065.
- KOVAR, J.L., *et al*. (2013). Characterization of IRDye 800CW chlorotoxin as a targeting agent for brain tumors. *Anal Biochem*, **440** (1), 212-9.
- KOZAKOV, D. *et al*., (2007). Discrimination of near-native structures in protein–protein docking by testing the stability of local minima. *Proteins*, **72** (1), 993–1004.
- KOZAKOV, D. *et al*., (2013). How good is automated protein docking? *Proteins,* **81** (1), 2159–2166.
- KOZAKOV, D. *et al*., (2017). The ClusPro web server for protein–protein docking. *Nature protocols*, **12** (2), 255-278.
- KRIEGER, E., *et al*., (2003). Homology modelling. *Structural Bioinformatics,* **3,** 507-521.

- KRISHNAMOORTHY, E., *et al*., (2016). Homology modeling of Homo sapiens Lipoic acid Synthase: substrate docking and insights on its binding mode. *Journal of theoretical biology,* **16** (1), 09-005.
- KUMAR, B. R. *et al*., (2015). In silico analysis of potential inhibitors of Ca^{2+} activated K^{+} channel blocker, Charybdotoxin-C from *Leiurus quinquestriatus* hebraeus through molecular docking and dynamics studies. *Indian J Pharmacol.* **47** (3) 280-284.
- KUMAR, B. R., *et al*. (2015). Pharmacophore modeling, *in silico* screening, molecular docking and molecular dynamics approaches for potential α-delta bungarotoxin-4 inhibitors discovery. *Pharmacogn Mag.,* **11** (1), 19-28.
- LAFLEUR, A. M. *et al*., (2001). Activation of pro-(matrix metalloproteinase-2) (pro-MMP-2) by thrombin is membrane-type-MMP-dependent in human umbilical vein endothelial cells and generates a distinct 63 kDa active species. *Biochem. J.* **357**, 107–115.
- LASKOWSKI, R. A., *et al*., (1993). 'PROCHECK: A program to check the stereochemical quality of protein structures'. *J Appl Cryst*, **26** (1), 283-291.
- LAURENT, F. *et al*., (1993). Evaluation of the relaxant effects of SCA40, a novel charybdotoxin-sensitive potassium channel opener, in guinea-pig isolated trachealis". *Br. J. Pharmacol.* **108** (3), 622–6.
- LEE, J. S. *et al*., (2015). Natural killer (NK) cells inhibit systemic metastasis of glioblastoma cells and have therapeutic effects against glioblastomas in the brain. *BMC Cancer,* **15**, 1011.
- LENSINK, M. F. *et al*., (2016). Modeling protein-protein and protein-peptide complexes: CAPRI 6th edition. *Proteins.*
- LEVITT, R. J., *et al*., (2005). PTEN-induction in U251 glioma cells decreases the expression of insulin-like growth factor binding protein-2. *Biochem Biophys Res Commun.* **336** (4), 1056-1061.

- LI, W., *et al*. (2015). The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic acids research,* **43** (W1), W580-4.
- LIOTTA, L. A. *et al*., (1980). Metastatic potential correlates with enzymatic degradation of basement membrane collagen. *Nature*. **284** (5751) 67–8.
- LIPPENS, G., *et al*., (1995). NMR sequential assignments and solution structure of chlorotoxin, a small scorpion toxin that blocks chloride channels. *Biochemistry*, **34** (1) 13-21.
- LOKMAN, A. N. *et al*., (2011). The Role of Annexin A2 in Tumorigenesis and Cancer Progression. *Cancer Microenviron*. **4** (2) 199-208.
- LYONS, S. A, *et al*., (2002). Chlorotoxin, a scorpion-derived peptide, specifically binds to gliomas and tumors of neuroectodermal origin. *Glia,* **39** (2), 162-73.
- MADSEN, H. D. *et al*., (2015). The source of matrix-degrading enzymes in human cancer: Problems of research reproducibility and possible solutions. *JCB,* **209** (2), 195-202.
- MAMELAK, A.N., *et al*. (2007). Targeted delivery of antitumoral therapy to glioma and other malignancies with synthetic chlorotoxin (TM-601). *Expert Opin*. *Drug Deliv*. **4** (1) 175-186.
- MAN, P*., et al.,* (2010). Accessibility changes within diphtheria toxin T domain when in the functional molten globule state, as determined using hydrogen/deuterium exchange measurements. *FEBS J.* **277,** 653–662.
- MAULE, F., *et al*., (2016). Annexin 2A sustains glioblastoma cell dissemination and proliferation. *Oncotarget,* **7** (34), 54632–54649.
- MCFERRIN, M.B., SONTHEIMER H. (2006). A role for ion channels in glioma cell invasion. *Neuron Glia Biol.* **2** (1) 39-49.

- MENDES, O., *et al*., (2016). MMP2 role in breast cancer brain metastasis development and its regulation by TIMP2 and ERK1/2. *Clin Exp Metastasis,* **24** (1), 341-351.
- MISHRA, A. and HOQUE, M. (2015). Three-Dimensional Ideal Gas Reference State based Energy Function. *Curr Bioinformatics Bentham J.* **14** (1), 1-7.
- MOBLEY, D. L., *et al*., (2007). Predicting absolute ligand binding free energies to a simple model site. *J. Mol. Biol.* **371,** 1118-1134.
- MOBLEY, L. D., *et al*., (2009). Binding of small-molecule ligands to proteins; "What you see" is not always "What you get". *Stucture,* **10,** 489-497.
- MORGUNOVA, E. *et al*., (1999). Structure of Human Pro-Matrix Metalloproteinase-2: Activation Mechanism Revealed. *Science,* **284** (1), 1667-1670.
- MOULT, J., *et al*. (2014). Critical assessment of methods of protein structure prediction (CASP)--round x. *Proteins,* **82** (2), 1-6.
- MUNAUT, C., *et al*., (2003). Vascular endothelial growth factor expression correlates with matrix metalloproteinases MT1-MMP, MMP-2 and MMP-9 in human glioblastomas. *International Journal of Cancer*, **106** (1) 848–855.
- MURPHY, G. and KNAUPER, V. (1997). Relating matrix metalloproteinase structure to function: why the "hemopexin" domain? *Matrix Biol.* **15** (1), 511–518.
- NAGASE, H. and WOESSNER, J. F. Jr. (1999). Matrix metalloproteinases. *J Biol*. *Chem.* **274** (1), 21491–21494.
- NAGASE, H., *et al*., (2006). Structure and function of matrix metalloproteinases and TIMPs. *European Society of Cardiology.* **12** (2), 562-573.

- NAKOPOULOU, L. *et al*., (2003). MMP-2 Protein in Invasive Breast Cancer and the Impact of MMP-2/TIMP-2 Phenotype on Overall Survival. *Breast Cancer Res.* **77** (2), 145-155.
- NEWTON, A. K. *et al*., (2007). Mass fingerprinting of toxic fractions from the venom of the Indian red scorpion, Mesobuthus tamulus: biotope-specific variation in the expression of venom peptides. *Rapid Commun. Mass Spectrom*. **21**, 3467–3476.
- ONISHI, M. *et al*., (2015). Annexin A2 regulates angiogenesis and invasion phenotypes of malignant glioma. *Brain Tumor Pathol*. **32** (1), 184–194.
- PETERSEN, T. N. *et al*., (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions*. Nature Methods,* **8**, 785-786.
- RAHME, G. J., *et al*. (2015). Id4 suppresses MMP2-mediated invasion of glioblastoma-derived cells by direct inactivation of Twist1 function. *Oncogene,* **34** (1), 53-62.
- RAMACHANDRAN, K. R. *et al*., (2017). Expression and prognostic impact of matrix metalloproteinase-2 (MMP-2) in astrocytomas. *PLoS ONE*, **12** (2) e0172234.
- RAMADAN, A. M. and HASSANEIN, S. E. (2014). Characterization of P5CS gene in Calotropis procera plant from the de novo assembled transcriptome contigs of the high-throughput sequencing dataset. *Comptes rendus biologies,* **337** (12), 683-690.
- RAMIREZ, P. Y., *et al*., (2013). Glioblastoma Multiforme Therapy and Mechanisms of Resistance. *Pharmaceuticals,* **6** (12), 1475-1506.
- RESENDE, B. F. F., *et al*., (2016). Evaluation of TgH(CX3CR1-EGFP) mice implanted with mCherry-GL261 cells as an in vivo model for morphometrical analysis of glioma-microglia interaction. *BMC Cancer,* **16** (72), 1-14.

- RETY, S. *et al*., (1999). The crystal structure of a complex of p11 with the annexin II *N*-terminal peptide*. Nat. Struct. Biol.* **6** (1) 89–95.
- RJEIBI, I. *et al*., (2010). Purification, synthesis and characterization of AaCtx, the first chlorotoxin-like peptide from Androctonus australis scorpion venom. *Peptides*, **32** (11) 656–663.
- RUPP, P. A., *et al*., (2008). MMP2-Integrin αvβ3 binding is required for mesenchymal cell invasive activity - but not epithelial locomotion: a computational time-lapse study. *Mol Biol Cell,* **19** (1), 5529-5540.
- SAWAYA, R. E *et al*., (1996). Expression and localization of 72 kDa type IV collagenase (MMP-2) in human malignant gliomas in vivo. *Clin Exp Metastasis,* **14** (1) 35-42.
- SELA-PASSWELL, N. *et al*., (2009). Structural and functional bases for allosteric control of MMP activities: Can it pave the path for selective inhibition? *Biochimica et Biophysica Acta*, **1803** (1), 29-38·
- SHAO, C. *et al*., (2006). Human Annexin A2 with Calcium bound. *J Biol Chem***.** **281** (1), 31689-31695.
- SHARMA, M. C. and SHARMA, M. (2007). The role of annexin II in angiogenesis and tumor progression: a potential therapeutic target. *Curr Pharm Des.* **13** (35), 3568-75.
- SINGH, S. K., *et al*. (2004). Identification of human brain tumour initiating cells. *Nature.* **432** (7015), 396-401.
- SODA, Y., *et al*., (2011). Transdifferentiation of glioblastoma cells into vascular endothelial cells*. Proc Natl Acad Sci USA.* **108,** (11), 4274–4280.
- SOEDA, A., *et al*., (2015). The evidence of glioblastoma heterogeneity. *Science reports,* **5** (10), 7979.
- SONDERMANN, H., *et al*., (2005). Computational docking and solution x-ray scattering predict a membrane-interacting role for the histone domain of the

Ras activator son of sevenless. *Proc. Natl. Acad. Sci. USA*, **102,** 16632–16637.

- SONG, H., *et al*., (2009). Low-density lipoprotein receptor-related protein 1 promotes cancer cell migration and invasion by inducing the expression of matrix metalloproteinases 2 and 9. *Cancer Res,* **69** (1), 879-886.
- SOROCEANU, L. *et al*., (1998). Use of chlorotoxin for targeting of primary brain tumors. *Cancer Res.* **58** (21), 4871-9.
- SOROCEANU, L. *et al*., (1999). Modulation of Glioma Cell Migration and Invasion Using Cl^- and K^+ Ion Channel Blockers. *J. Neurosci*. **19** (1), 5942–5954.
- SPRINGMAN, E. L. *et al*., (1990). *Proc Natl Acad Sci USA***. 87** (1), 364–368.
- STRONGIN, A. Y. *et al*., (1995). Mechanism of cell surface activation of 72-kDa type IV collagenase. Isolation of the activated form of the membrane metalloprotease. *J. Biol. Chem.* **270** (1) 5331–5338.
- STROUD, M. R. *et al*., (2011). *In vivo* bio-imaging using chlorotoxin-based conjugates. *Curr. Pharm. Des.* **17** (1) 4362–4371.
- STUPP, R., *et al*., (2014). High-grade glioma: ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Annals of Oncology,* **25** (3), 93-101.
- SUN, S., *et al*. (2012). Protein alterations associated with temozolomide resistance in subclones of human glioblastoma cell lines. *Journal of neuro-oncology,* **107** (1), 89-100.
- TAMBORINI, M., *et al*. (2016). A combined approach employing chlorotoxin-nanovectors and low dose radiation to reach infiltrating tumor Niches in Glioblastoma. *ACS nano*, **10** (2), 2509-2520.

- TAO, P. *et al*., (2009). Matrix Metalloproteinase 2 (MMP2) Inhibition: QM/MM Studies of the Inhibition Mechanism of SB-3CT and its Analog. *Biochemistry,* **48** (41), 9839.
- TARASSISHIN, L., *et al*., (2014). Aberrant Expression of Interleukin-1β and Inflammasome Activation in Human Malignant Gliomas. *PLoS One,* **9** (1), 103432.
- TARUI, T., *et al*., (2002). Plasmin-induced Migration of Endothelial Cells. A potential target for the anti-angiogenic action of angiostatin Ss. *The Journal of Biological Chemistry*, **277**, 33564-33570.
- TATENHORST, L., *et al*. (2006). Knockdown of annexin 2 decreases migration of human glioma cells in vitro. *Neuropathol Appl Neurobiol,* **32** (3), 271–277.
- THOMAS, N. P., *et al*. (2011). SignalP 4.0: discriminating signal peptides from transmembrane regions. *Nature Methods*, **8** (1), 785-786.
- TURNER, L. K. *et al*., (2014). Cl^- and K^+ channels and their role in primary brain tumour biology*. Philos Trans R Soc Lond B Biol Sci*. **369** (1638) 20130095.
- ULLRICH, N. *et al*., (1998). Expression of voltage-activated chloride currents in acute slices of human gliomas. *Neuroscience.* **83** (4), 1161-73.
- VALAPALA, M. *et al*., (2011). A competitive hexapeptide inhibitor of annexin A2 prevents hypoxia-induced angiogenic events. *J Cell Sci.* **124** (9), 1453-1464.
- VASQUEZ-DUNDDEL, D., *et al*., (2013). STAT3 regulates arginase-I in myeloid-derived suppressor cells from cancer patients. J Clin Invest. 123(4) 1580-9.

- VEISEH, M. *et al*., (2007). Tumor paint: a chlorotoxin:Cy5.5 bioconjugate for intraoperative visualization of cancer foci. *Cancer Research,* **67** (14), 6882-6888.
- VEISEH, O., *et al*. (2009). Inhibition of tumor-cell invasion with chlorotoxin-bound superparamagnetic nanoparticles. *Small,* **5** (1) 256–264.
- VISSE, R. and NAGASE, H. (2003). Matrix metalloproteinases and tissue inhibitors of metalloproteinases: structure, function, and biochemistry. *Circ Res.* **92** (8), 827-39.
- VYAS, K. V. *et al*., (2012). Homology Modeling a Fast Tool for Drug Discovery: Current Perspectives. *Indian J Pharm Sci.* **74** (1) 1-17.
- WANG, H., *et al*., (2003). Insulin-like growth factor binding protein 2 enhances glioblastoma invasion by activating invasion-enhancing genes. *Cancer Res* **63** (1), 4315-4321.
- WANG, H., *et al*., (2015). The Challenges and the Promise of Molecular Targeted Therapy in Malignant Gliomas. *Neoplasia,* **17** (3), 239-255.
- WANG, M. *et al*., (2003). The expression of matrix metalloproteinase-2 and -9 in human gliomas of different pathological grades. *Brain Tumor Pathol.* **20** (2), 65-72.
- WUDAYAGIRI, R. *et al*., (2005). Functional expression of lepidopteran-selective neurotoxin in baculovirus: Potential for effective pest management. *Biochimica et Biophysica Acta*. **1760** (6) 158–163.
- XIA, B. *et al*., (2016). Accounting for pairwise distance restraints in FFT-based protein–protein docking. *Bioinformatics.* **2016,** 1-3.
- XU, T. *et al*., (2016). Identification of two novel Chlorotoxin derivatives CA4 and CTX-23 with chemotherapeutic and anti-angiogenic potential. *Scientific Reports* **6** (1), 1-14.

- YAN, W., *et al*., (2011). Oncogene addiction in gliomas: implications for molecular targeted therapy. *J Exp Clin Cancer Res*. **11** (30), 58.
- YI, Y., *et al*., (2016). Glioblastoma Stem-Like Cells: Characteristics, Microenvironment, and Therapy. Front Pharmacol. **7,** 477.
- YIN, L. T., *et al*. (2007). Potential biochemical therapy of glioma cancer. *Biochemical and biophysical research communications,* **362** (2), 225-229.
- YONG, V. W. *et al*., (2001). Metalloproteinases in biology and pathology of the nervous system. *Nature reviews Neuroscience*. **2** (7) 502–11.
- ZHAI, H., *et al*. (2011). Annexin A2 Promotes Glioma Cell Invasion and Tumor Progression. *J Neurosci.* **31** (40), 14346-14360.
- ZHANG, Y., *et al*. (2016). Functionalized magnetic nanochains with enhanced MR imaging: Anovel nanosystem for targeting and inhibition of early glioma. *Colloids and Surfaces B: Biointerfaces,* **140** (16), 437-445.
- ZHAO, J. *et al*., (2011). Preparation and *In-vitro* Evaluation of 131 I-BmK CT as a Glioma-Targeted Agent. *Cancer Biotherapy & Radiopharmaceuticals*, **25** (3), 353-9.
- ZHAO, S. H., *et al*., (2010). Annexin A2 promotes choroidal neovascularization by increasing vascular endothelial growth factor expression in a rat model of argon laser coagulation-induced choroidal neovascularization. *Chin Med J* (*Engl)*, **123** (6), 713–721.
- ZHOU, H. *et al*., (2002). Distance-scaled, finite ideal-gas reference state improves structure-derived potentials of mean force for structure selection and stability prediction. *Protein Science*, **11** (1), 27142726.

7 APPENDIX.

7.1 Quality of MMP2 Isoforms

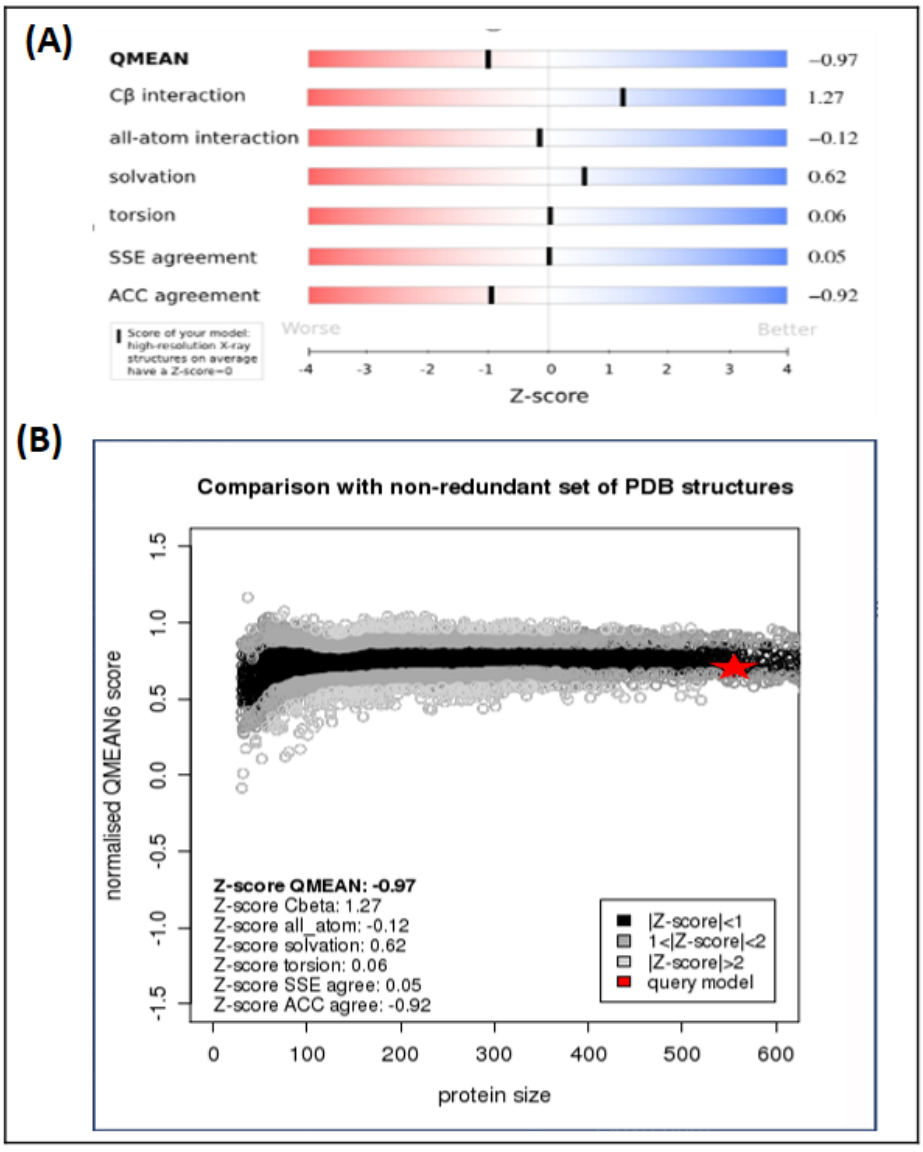

Fig 7.1.1 The quality assessment of 1CK7-A-. (a) QMEAN6 score of the native activated MMP2 (1CK7-A-) and identification of the score of individual geometrical features. (b) Z-Score of 1CK7-A- different shades of grey as black dots represent Z-Scores of the reference structures located in the PDB-databank. The red star in the figure represents 1CK7-A- model.

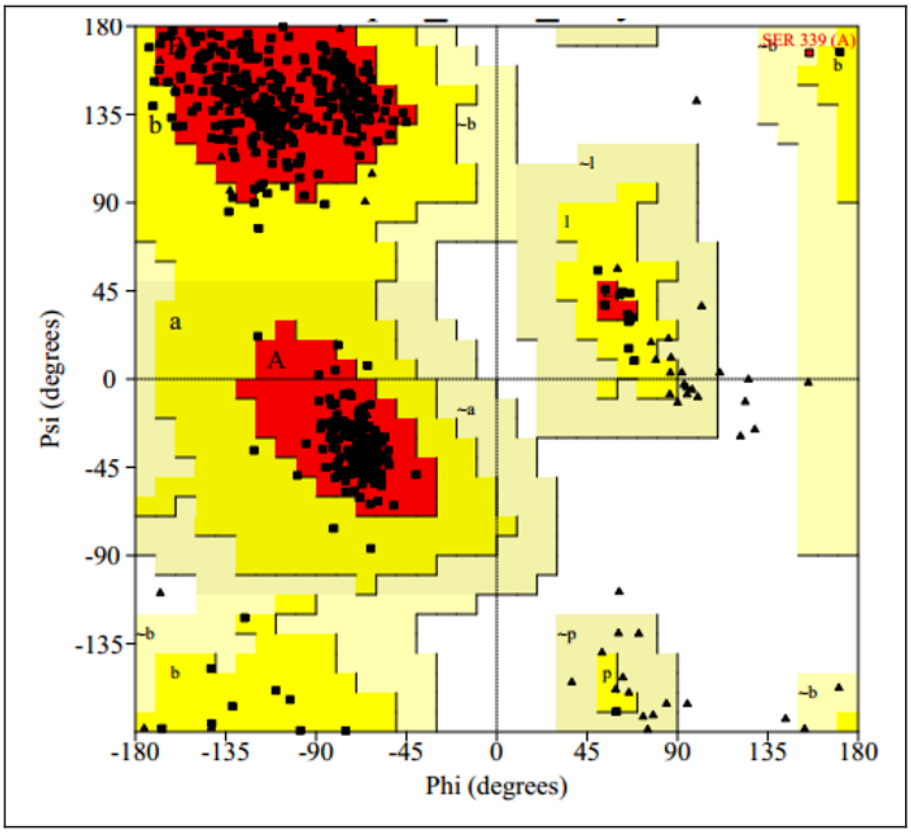

Fig 7.1.2 PROCHECK Ramachandran plot of 1CK7˙A˙. Ramachandran plot shows resdiues in the most favoured, yellow region show residues in the allowed region, grey region show residues in the disallowed region, white region shows residues in the disallowed region. Glycine can be located in white regions as it does not contain a side chain.

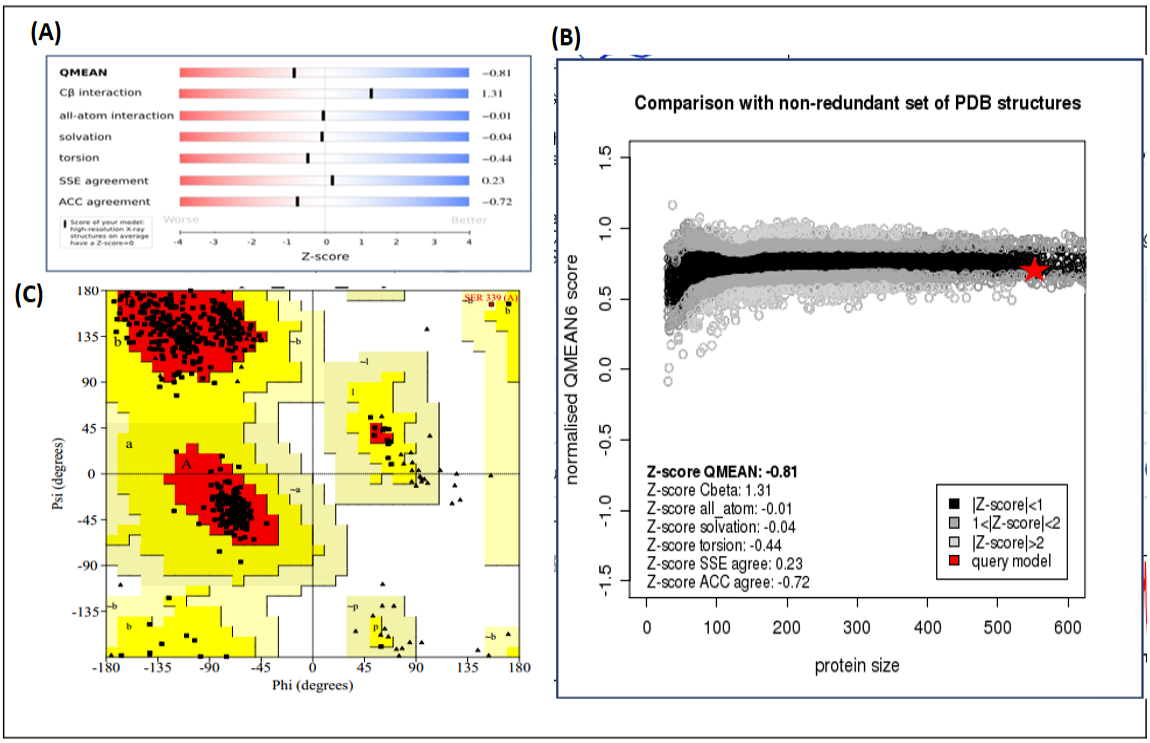

Fig 7.1.3 The quality assessment of 1CK7 A-. (A) QMEAN6 score of the activated MMP2 (1CK7 A-) and identification of the score of individual geometrical features. (B) Z-Score of 1CK7 A- different shades of grey as black dots represent Z-Scores of the reference structures located in the PDB-databank and the red star represents 1CK7 A- model. (C) Ramachandran plot shows resdiues in the most favoured, yellow region show residues in the allowed region, grey region show residues in the disallowed region, white region shows residues in the disallowed region. Glycine can be located in white regions as it does not contain a side chain.

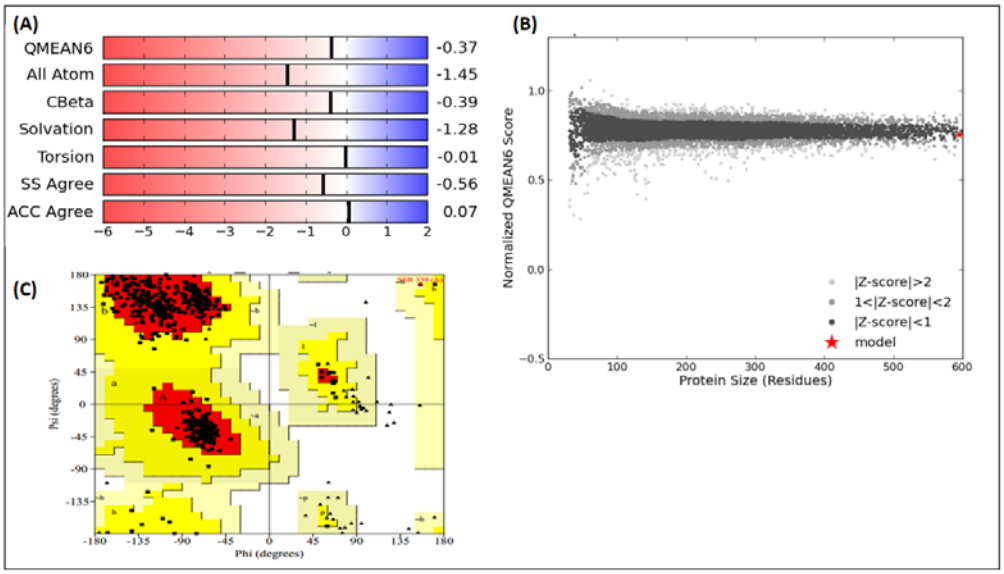

Fig 7.1.4 The quality assessment of 1CK7-. (A) QMEAN6 score of the native MMP2 (1CK7-) and identification of the score of individual geometrical features. (B) Z-Score of 1CK7- different shades of grey as black dots represent Z-Scores of the reference structures located in the PDB-databank and the red star represents 1CK7 - . (C) Ramachandran plot shows resdiues in the most favoured, yellow region show residues in the allowed region, grey region show residues in the disallowed region, white region shows residues in the disallowed region. Glycine can be located in white regions as it does not contain a side chain.

7.2 Quality of Chlorotoxin

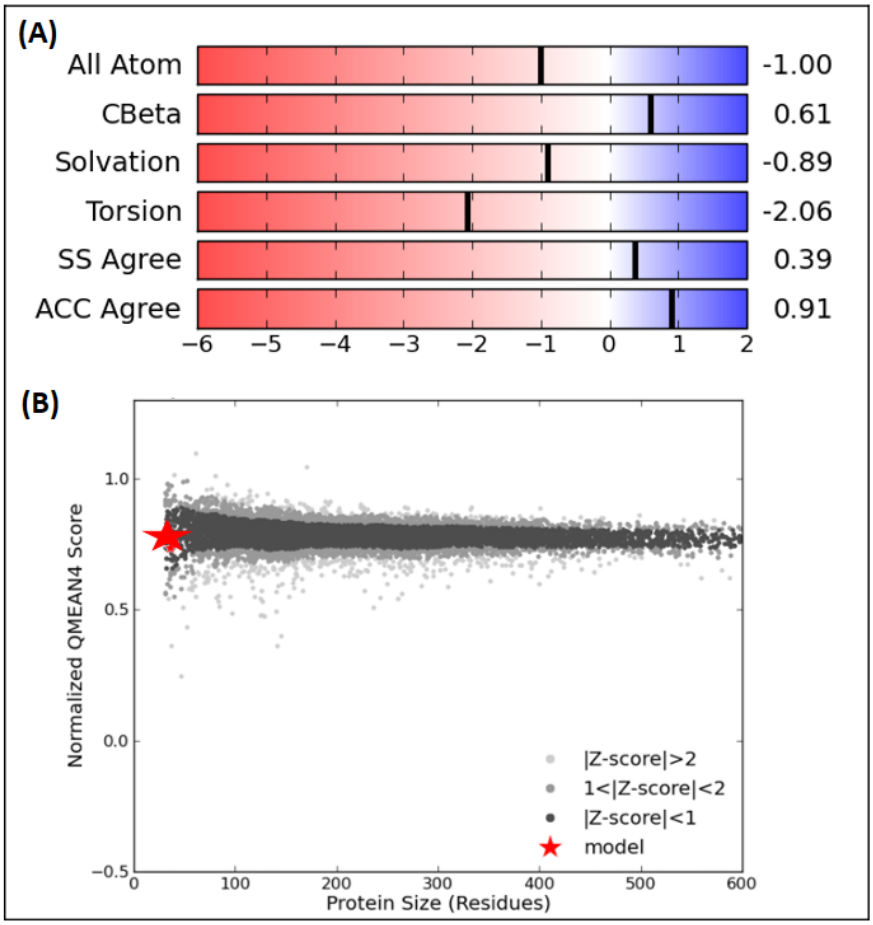

Fig 7.2.1 The quality assessment of chlorotoxin. (a) QMEAN6 score of chlorotoxin and identification of the score of individual geometrical features. QMEAN6 value of Cltx is 0.33 (b) Z-Score of Cltx, different shades of grey as black dots represent Z-Scores of the reference structures located in the PDB-databank and the red star represent Cltx.

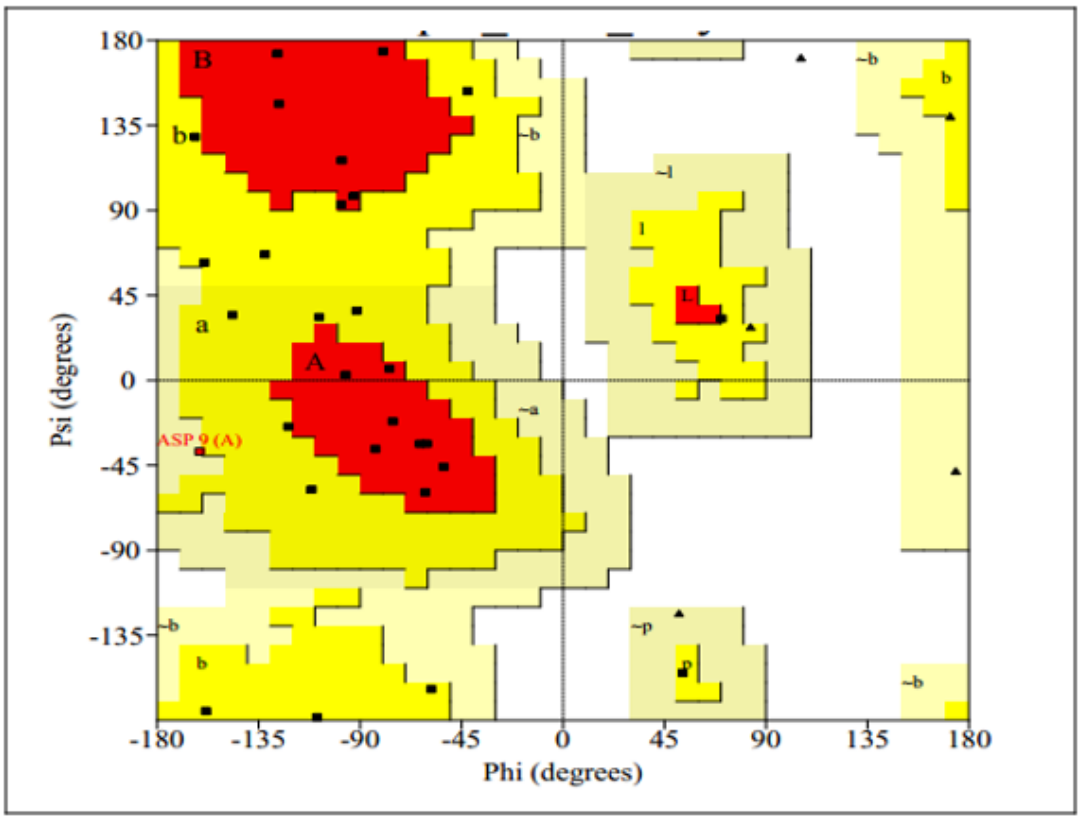

Fig 7.2.2 PROCHECK Ramachandran plot of chlorotoxin 3D strcutrue. Ramachandran plot shows resdiues in the most favoured, yellow region show residues in the allowed region, grey region show residues in the disallowed region, white region shows residues in the disallowed region. Glycine can be located in white regions as it does not contain a side chain.

Toxin	QMEAN 6	Z-Score	DFIRE	% Residues in disallowed region
Cltx	0.33	-1.561	-32.55	0
CTX23	0.45	-0.916	-33	0
CA4	0.61	-0.054	-32.37	0
Cltx-like pep AaCTX	0.75	0.706	-32	0

Table 7.2.3 The quality assessment of 3D models of Cltx-like short toxins. QMEAN6 and Z-Score values of all four models are within an acceptable QMEAN range. All models have a low energy hence less steric clashes. All models are good structures and comparable to a set of experimentally determined 3D structures of similar sizes from PDB-databank. None of residues are in disallowed region in Ramachandran plot.

Acknowledgement

I would like to thank Dr. Kirsti Newton and Prof. Peter Strong, my thesis supervisors, for their academic supervision and support throughout all my years in Biomolecular Research Centre, as well as, Dr. Anne-Marie Harrison for her supervision at various stages during this project, for her invaluable advice and support.

I wish to thank my parents and family for their support, encouragement and motivation during the work on my thesis.

I also thank everyone in BMRC who have helped me during this research study.