Stony Brook University

Identification of Small Molecules that Bind to the Hemopexin Domain of Matrix Metalloproteinase-9

A Thesis Presented

by

Zhu Li

to

The Graduate School

in Partial Fulfillment of the

Requirements

for the Degree of

Master of Science

in

Chemistry

Stony Brook University

August 2012

Stony Brook University

The Graduate School

Zhu Li

We, the thesis committee for the above candidate for the Master of Science degree, hereby recommend acceptance of this thesis.

Professor Nicole S. Sampson, Advisor, Department of Chemistry

Professor Robert C. Rizzo, Chair, Department of Chemistry

Professor Elizabeth Boon, Third member, Department of Chemistry

This thesis is accepted by the Graduate School

Charles Taber
Interim Dean of the Graduate School

Abstract of the Thesis

Identification of Small Molecules that Bind to the Hemopexin Domain of Matrix Metalloproteinase-9

by

Zhu Li

Master of Science

in

Chemistry

Stony Brook University

2012

Matrix metalloproteinases (MMPs) play an important role in cancer progression by degrading extracellular matrix components, promoting tumor cell migration, and thus enhancing tumor metastasis. Several generations of anticancer drug were designed to inhibit the proteolytic functions of MMPs, but these drugs failed in clinic trails due to lack of selectivity. This study focused on the hemopexin domain of MMP-9, a nonproteolytic domain, which is essential for MMP-9 regulated tumor cell migration. Compared with the catalytic domain of MMPs, fewer amino acids are conserved between the hemopexin domains of MMP family members. This difference provides the principle basis for identification of inhibitors targeting the hemopexin domain with high selectivity and specificity. An *in silico* docking approach was employed to screen for novel small-molecule compound that bind to the hemopexin domain of MMP-9. Seven computational hits were evaluated in biochemical binding assays. Of these, 4 compounds were identified that have micromolar affinities for the hemopexin domain of MMP-9

Table of Contents

List of Figures

List of Tables

Acknowledgement

It is a pleasure to thank many people who made this thesis possible.

First of all, I would like to express my sincere gratitude to my advisor Prof. Nicole S. Sampson for her continuous support to my master's study and research. With her patience, enthusiasm and immense knowledge, she helped me to overcome difficulties in my study. Throughout my thesis-writing period, she provided encouragement, sound advice, good teaching, and lots of good ideas. I would not finish this thesis without her.

Besides my advisor, I would like to thank my thesis committee members: Prof. Robert C. Rizzo and Prof. Elizabeth Boon, for their encouragement, insightful comments, and patience. In Prof. Rizzo's classes, I learned how to use the computational tool, docking, in drug design and discovery, which is really meaningful and novel to me. I would not finish the computational part of my study without his instruction. A lot of important knowledge about chemical biology was learned from Prof. Boon in CHE 541. Her guidance enlightened my first glance of biochemistry

I would like to thank Prof. Jian Cao and Prof. Jin Wang. They gave a lot of support and instruction to this project. The protein preparation was finished in Prof. Cao's lab. He also generously provided the NCI compounds tested in the fluorescence assay. My sincere thanks also go to Prof. Wang, because the computational hits all came from his lab.

I am indebted to many labmates for providing a stimulating and fun environment in which to learn and grow. I am especially grateful to Suzanne Thomas, Eunjung Lee, Siyeon Lee, Linghui Wu, Benson Su, Xiaoqian Jin, Rui Lu, Meng Yang, He Huang, and Sudipto Mukherjee. Sudipto Mukherjee was particularly helpful when I was doing computational part in this project.

Last but not least, I would like to thank my family and friends for providing a loving environment for me.

Chapter 1. Introduction

1.1. The role of matrix metalloproteinases (MMPs) in cancer

The metastasis of tumors is caused by the uncontrolled interaction between malignant cancer cells and surrounding nonmalignant extracellular matrix (ECM) consisting of interstitial matrix and basement membrane proteins. Understanding and control of such interactions is one of the significant challenges in cancer research [1].

Matrix metalloproteinases (MMPs) are a major family of enzymes that regulate the molecular communication between cells and surrounding extracellular matrix by modulating the composition and integrity of the extracellular matrix. They are capable of cleaving extracellular matrix (ECM) components, such as college and gelatin, and several nonmatrix proteins. By degrading ECM components, MMPs release bioactive chemicals located in ECM, such as cell growth factors and cytokines. MMPs regulate the signaling pathway in various normal physiological processes, such as inflammation, cell proliferation, and angiogenesis [2]. But the uncontrolled proteolysis caused by MMPs leads to tumor cell invasion and metastasis in pathological conditions. As a result, MMPs have been an attractive target for anticancer drug discovery [3].

The MMPs are zinc-dependent endopeptidases, which belong to a larger family of protease known as the metzincin superfamily. The biological activity of MMPs is complicated, and can be highly mediated at different levels: gene expression, compartmentalization, activation of zymogen as well as the presence of tissue inhibitors of metalloproteinases (TIMPs). TIMPs are a family of specific MMP protein inhibitors and are also commonly expressed in tumor sites. The uncontrolled proteolysis caused by MMPs is

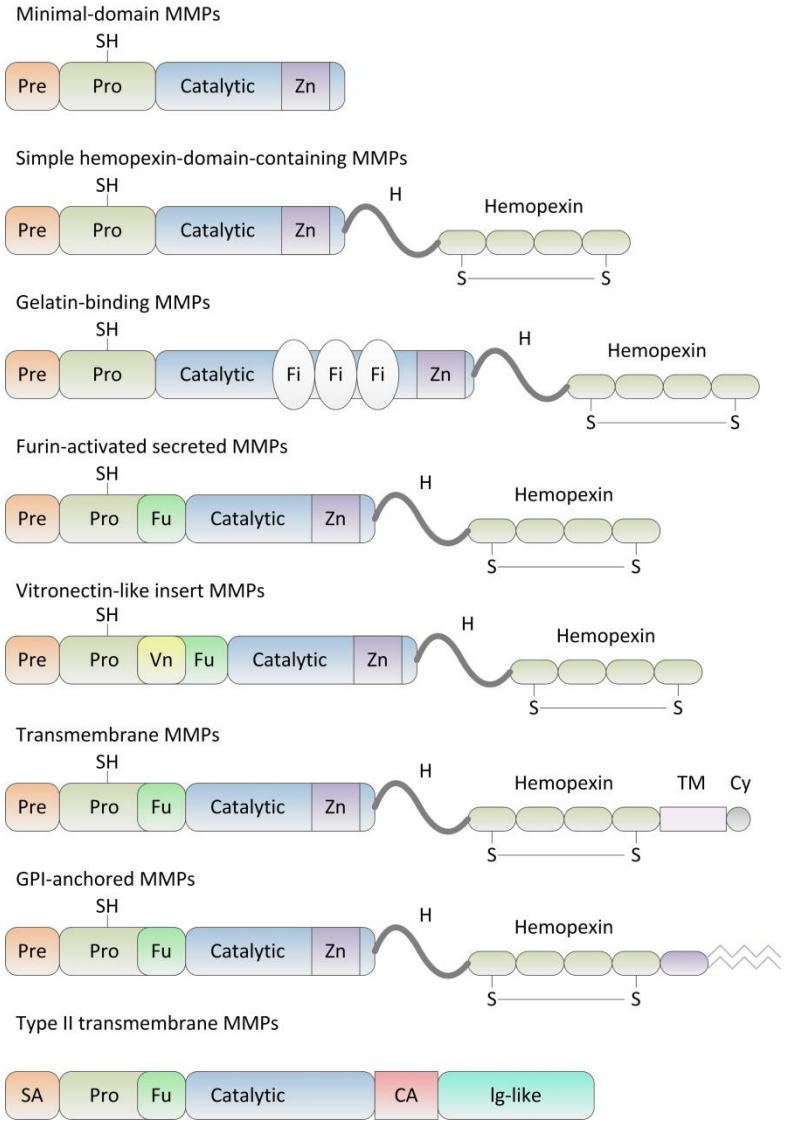
characterized as excessive or deficient degradation of ECM composition [4]. In cancer, such proteolysis results in unregulated tumor growth, invasion and metastasis. Moreover, MMPs can even modulate signaling pathways in a nonproteolytic manner [5].

1.2. Members of the MMP family

There are 24 MMPs in humans which have broad and overlapping substrate specificities. Historically, according to their specificity for ECM components, the members in MMP family were divided into four subfamilies: collagenase, gelatinases, stromelysins and matrilysins. With the increasing number of substrates, a new classification system was employed: members are grouped based on their structure. According to the new criteria, there are 8 structural classes: four of them are secreted and the rest are membrane-type matrix proteinases (Figure 1) [5]. Most MMP members share three common and distinctive domains: an amino-terminal propeptide region, a catalytic domain; and a hemopexin domain at the carboxy-terminal (Figure 2). The propeptide region is composed of 80-90 amino acids containing a cysteine residue. MMPs are initially secreted as a latent form which is caused by the interaction between the cysteine residue on propeptide region and catalytic zinc ion on catalytic domain (Figure 2 and 3) [6]. The disruption of such interaction is called cysteine switch mechanism, resulting in MMPs zymogen activation (Figure 3). The cysteine switch mechanism is usually regulated by the proteolytic removal of propeptide region or chemical modification of cysteine residue [7]. In catalytic domain, there are two zinc ions. The first zinc ion is located at the active site of MMPs, and is called catalytic zinc. It is essential for the proteolytic activity of all MMPs. This catalytic zinc ion coordinates with three histidine residues, and this structural feature is conserved among

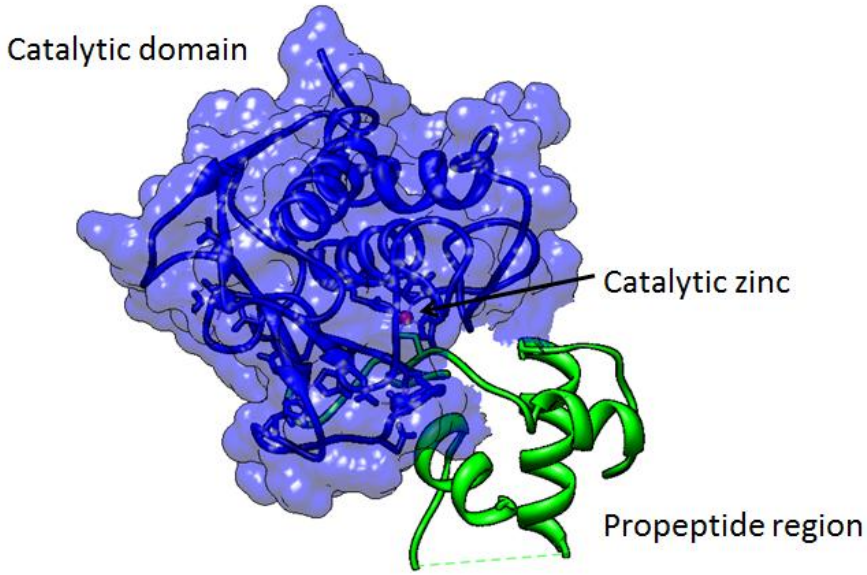
all members in MMP family (Figure 1 and 2) [6]. The second zinc ion is called structural zinc, and is approximately 12 Å away from the catalytic zinc. The function of this zinc ion is still not clear. Although it does not participate in the catalytic process directly, the hemopexin domain plays an important role in the biological function of MMPs. The hemopexin domain is responsible for substrate recognition and binding, it also interacts with tissue inhibitors of metalloproteinases (TIMPs).

Figure 1. Schematic representation of domain structures for 8 MMP subgroups.

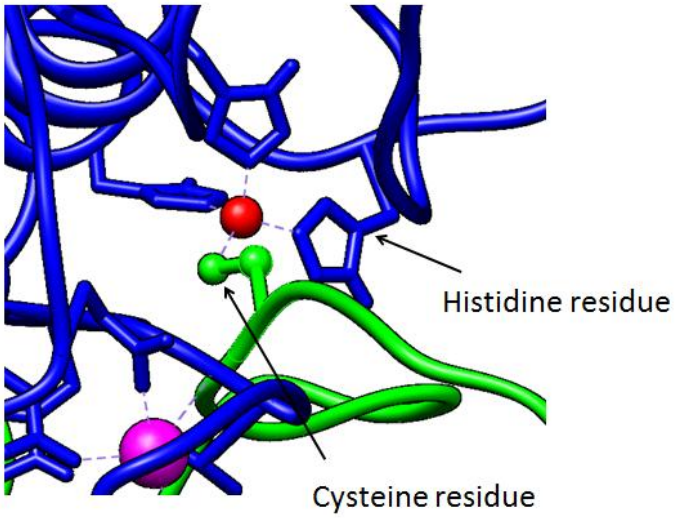


According to their structural differences, MMPs are divided into 8 subgroups. The first five subgroups are secreted MMPs; the rest three are membrane type MMPs. Except from Type II transmembrane MMPs, all MMPs contain an amino-terminal signal sequence (Pre) and a propeptide region (Pro) which has a cysteine residue (SH). In most of MMPs, there is a hemopexin domain with a disulphide bond (S-S) which connects its first and fourth blade. The catalytic domain is shared by all MMPs and connected to the hemopexin domain by a hinge (H). The gelatin binding MMPs, MMP-2 and MMP-9, contain a gelatin binding site (Fi), which is similar to collagen-binding type-II repeats in fibronectin. In furin-activated secreted MMPs, there is an intracellular furin-like serine proteinase recognition motif (Fu). This Fu motif is also shared by vitronecctin-like insert (Vn) MMPs and membrane type MMPs. The membrane type MMPs have a transmembrane domain (TM) and a cytoplasmic domain (Cy). In type II transmembrane MMP, MMP-23, there is an N-terminal signal anchor (SA), a unique cyctein array (Ca), as well as an immunoglobulin-like domain (lg-like).

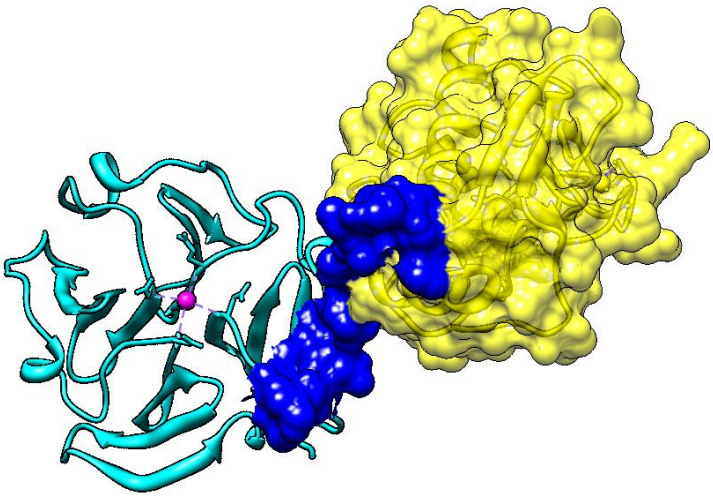
Figure 2. Domain structures of MMPs



***a**. The crystallographic structures of propeptide region and catalytic domain (PDB code: 1SLM [8]). The structure shown here is the inactive form of MMPs. The catalytic domain is shown by blue surface, the propeptide region as green ribbon; and the catalytic zinc is represented by red sphere.*

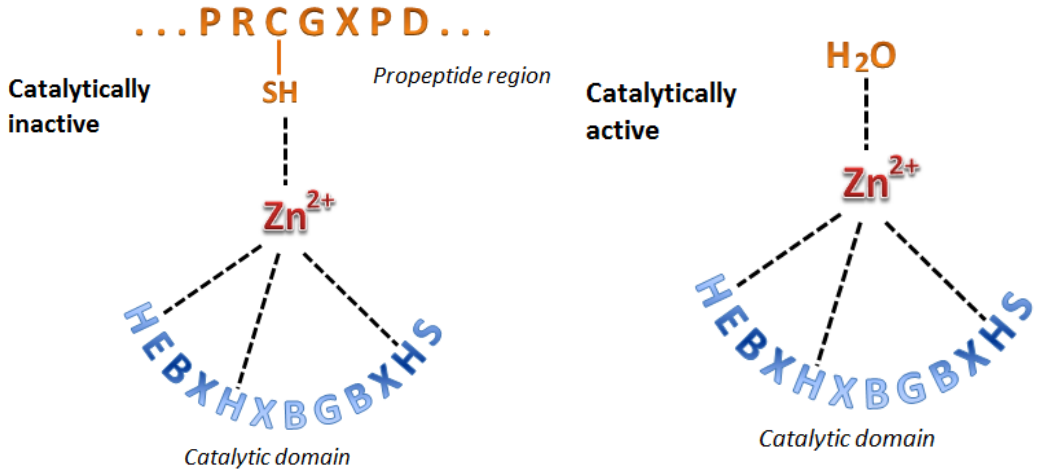


***b**. The structure of active site in the catalytic domain of MMPs. The catalytic zinc ion coordinates with three histindine residues from the catalytic domain and a cysteine residue from the propeptide region. Such conformation keeps MMPs in latent form. The catalytic domain is in blue, and the propeptie region is in green. The catalytic zinc ion is shown by red sphere, and the calcium ion is represented by magenta sphere.*



c. *The crystallographic structure of the catalytic domain, the hinge region and the hemopexin domain (PDB code: 1FBL, [9]). The structure shown here is the active form of MMPs, due to the removal of the propeptide region. The catalytic domain is shown by yellow surface, the hinge region as blue surface, and the hemopexin domain is represented by cyan ribbon.*

Figure 3. The cysteine switch mechanism



The PRCGXPD sequence is conserved among propeptide regions of MMP family. The thiol group from this conserved sequence interacts with the catalytic zinc ion in active sites of MMPs, thus keeping the enzyme in inactive form. MMPs are activated by proteolytic removal of the propeptide region or chemical modulation of the cysteine residue. Therefore, the thiol group is replaced by a water group.

1.3. The development of anticancer drugs design: inhibitors of MMPs

Due to the roles of MMPs in tumor progression, MMPs are important targets for anticancer drug discovery. Several types of agent have been designed to inhibit bioactivity of MMPs, based on four normal strategies: 1. inhibiting MMP synthesis; 2. disturbing the interaction between MMPs and other proteins; 3. introducing cytotoxic agents activated by MMPs; 4. blocking enzymatic activities of MMPs. Even if the previous MMP inhibitors (MMPI) are designed in terms of different concepts, none of them show selectivity towards an individual member in the MMP family.

Peptide mimics are the first generation of MMP inhibitors. They mimic the amino acid sequences in cleavage sites of collagen. These peptide mimics serve as competitive inhibitors that chelate the catalytic zinc ion in MMPs. The second generation of MMP inhibitors is non-peptide mimics. These inhibitors are designed based on the three-dimensional structure of active sites in the MMP catalytic domains [10].

Although a large quantity of preclinical studies showed that these MMP inhibitors were promising, most of them failed in cancer clinical trials. There are several possible reasons responsible for the failure of previous MMP inhibitors [11]. The first reason was the lack of selectivity for an individual member in the MMP family. The highly conserved structure in the catalytic domain of all MMPs makes it difficult to design inhibitors which can discriminate between different members of the MMP family. Secondly, the action mode of MMPs is complicated: recent research found that some MMPs have tumor suppressing functions, in addition to their tumor promoting functions [3]. Some MMPs can suppress cancer cell growth by activating transforming growth factor-β (TGF-β). MMPs can also stimulate cell death by changing ECM components, indirectly affecting integrin signaling

pathways. Moreover, by degrading extracellular matrix components, MMPs can release some chemicals with angiogenesis-inhibiting activities [12]. For instance, MMPs release anti-angiogenic factors, such as angiostatin, endostatin and tumstatin, by cleaving plasminogen, collagen type VI. MMPs also inhibit angiogenesis by degrading urokinase-type plasminogen activator receptor (uRAP) which is essential to endothelial cell invasion *in vitro* [5]. For example, increased tumor progression and pathological vascularization were also observed in MMP-9 knock-out mice. This pathology is caused by the lack of MMP-9 which degrades collagen Ivα and thus releases tumstatin, an angiogenesis suppressor. As a result, MMPs influence both tumor metastasis promotion and metastasis suppression, depending on when they are expressed and what substrates are available [7]. Thirdly, some MMPs elicit function in a nonproteolytic manner. For example, the cell migration in transwell chamber assay is promoted by the proteolytically inactive form of MMP-9. In this case, the functional domain of MMP-9 is the hemopexin domain, rather than the catalytic domain [13]. The lack of abilities to inhibit the nonproteolytic functions of MMPs is another reason they failed in clinical trials.

The development of MMP inhibitors with high selectivity for a given member in MMP family is pressingly needed. First, selective MMP inhibitors in cancer treatment are necessary. Broad spectrum inhibitors suppress proteolytic activity at a system wide level, which is beneficial to the treatment of cancer in later stages when most MMPs have been upregulated. However, these broad spectrum MMP inhibitors also inhibit the tumor suppressing functions of some special MMPs. In contrast, the development of selective MMP inhibitors will enable cancer therapy to target individual MMP, according to cancer stages and individual MMP upregulated level. The combinatorial employment of selective

MMP inhibitors will allow an attack on crucial MMPs which promote cancer progression, and have no influence on the MMPs which have cancer suppressing functions. Therefore, the therapy with selective MMP inhibitors can effectively reach the treatment goal and avoid the side effects at the same time.

1.4. The matrix metalloproteinase-9 (MMP-9)

In recent years, large quantities of evidence verified the association between the deregulated MMPs and tumor metastasis in human cancers, particularly for the matrix metalloproteinase-9 (MMP-9) which is also called gelatinase-B.

1.4.1. The biological functions of MMP-9 in physiological and pathologic conditions

Due to its highly regulated biological activity, MMP-9 is involved in a variety of physiological and pathological processes. The main function of MMP-9 in physiology is remodeling the extracellular matrix components (EMC); meanwhile the uncontrolled degradation of ECM often leads to pathologic problems.

MMP-9 is secreted as a latent zymogen, followed by the removal of the propeptide region which converts a 92 kDa proMMP-9 into an 82 kDa active MMP-9. MMP-9 can be activated by plasmin, trypsin and other activated MMP members, such as MMP-2, MMP-3 and MMP-13 [6]. The catalytic activity of MMP-9 also can be activated by modification of a cysteine residue in the propeptide region. For example, the binding of MMP-9 to type IV collagen-coated surface can disrupt the interaction between the cysteine residue in the propeptide region and the catalytic zinc ion in the catalytic domain, enhancing the catalytic

activity of MMP-9 even in the presence of the propeptide region [14]. Activated MMP-9 is capable of degrading various ECM components and non-matrix proteins (Table 1) [14].

Table 1. MMP-9 substrates [14]

Substrates type	*Substrates*
ECM	*Collagens III, IV and V; Gelatin; Elastin; Vitronectin; Entactin*
Non-matrix	*ProTGF-β; proTNF-α; IL-2Rα; ICAM-1; EGRF-1; Kit ligand; CXCL1/GRO-α; CXCL4/PF4; CXCL8/IL-8; CXCL9/MIG; CXCL11/ITAC; CXCL12/SDF-1; α1 proteinase inhibitor; Plasminogen; KISS-1; IFN-β*

MMP-9 plays an essential role in the physiological process of inflammation. MMP-9 is produced by both monocytes and T cells, and is also an important component of human neutrophils. MMP-9 is thought to be essential to host defense, because monocytes, T cells and neutrophils are involved in the inflammatory process. In five years, immunodeficiency, such as spontaneous tumor development and opportunistic infections, was observed in the MMP-9 deficient mice which were under a conventional animal house condition rather than specific pathogen-free (SPF) environment [15].

MMP-9 is also important in the wound repair process under physiological conditions. In wound healing, the epithelium and granulation tissue express a large quantity of MMP-9. This phenomenon indicates that MMP-9 supports cell movement in wound matrix and remodel granulation tissue matrix by separating keratinocytes from the basement membrane. For example, in fluid of burn blisters, MMP-9 expression was observed within four to eight hours and obviously increased in the first 2 days after injury. This increase suggests that MMP-9 promotes remodeling of denatured collagen.

In addition to its biological function under normal conditions, MMP-9 is believed to be a cancer biomarker. All members of the MMP family are thought to participate in tumor development, invasion and metastasis, while special attention has focused on MMP-9 for two reasons. First, sharply increased MMP-9 levels are detected in various malignant tumors. Second, tumor aggressiveness and poor prognosis are highly related to overexpression and activity of MMP-9. Increased MMP-9 levels are observed in melanoma, hematological malignancies, as well as lung, breast and pancreas cancers [16, 17].

The degradation of ECM enables MMP-9 to produce or release bioactive molecules which can promote tumor development. The cryptic information stored in the ECM is also released by MMP-9, resulting in angiogenesis or tumor cell spreading. For instance, MMP-9 is able to expose cryptic sites by proteolytic degradation of collagen IV, leading to angiogenesis [18]. At the same time, MMP-9 also releases growth factors and angiogenic factors, such as FGF-2, VEGF AND TGF-β, promoting endothelial cells production and migration, and further enhancing tumor angiogenesis and invasion [19].

Recent research shows that MMP-9 can also suppress tumor metastasis. For example, by degrading collagen Ivα, MMP-9 can release tumstatin, an angiogenesis suppressor. This can explain why there was increased tumor progression and pathological vascularization detected in MMP-9 knock-out mice.

1.4.2. The structure of MMP-9

MMP-9 belongs to gelatin-binding subgroup of the MMP family. This subgroup also includes MMP-2 (gelatinase-A). MMP-9 shares a common domain structure with other members in this subgroup consisting of a propeptide, a catalytic and a hemopexin domain.

In the catalytic domain, there is a conserved sequence AHEXGHXXGXXH, which includes three histidine residues coordinated with the catalytic zinc ion. This catalytic zinc ion has a fourth ligand when MMP-9 is initially synthesized and secreted. The fourth ligand is a cysteine residue from the conserved sequence PRCGXPD in the propeptide region [15]. The interaction between the cysteine residue and the catalytic zinc maintains the nascent MMP-9 in an inactive form. The disruption of this interaction, such as the removal of propeptide region or modification of the cysteine residue in the propeptide region, triggers the zymogen activation. The hemopexin domain is connected to the catalytic domain by a hinge region. It is responsible for substrate recognition and specificity. The hemopexin domain of MMP-9 is also suggested to participate in the interactions with tissue inhibitors of metalloproteinases (TIMPs), which is major mechanism to regulate the biological activity of MMP-9. The amino-terminal domains of TIMPs bind to the catalytic sites of MMPs, and thus inhibit the proteolytic activities. In addition, some proMMPs, including proMMP-1, proMMP-2, proMMP-9 and proMMP-13, also interact with TIMPs through the hemopexin domains which bind to the carboxyl-terminal of TIMPs [15, 20].

In addition to their three conserved domains, MMP-9 and MMP-2 have a unique domain called the fibronectin type II-like domain or gelatin binding domain. This domain contains three tandem fibronectin type II-like modules. Because of this special domain structure, MMP-2 and MMP-9 are grouped into a MMP subfamily which is called gelatinases. The fibronectin type II-like modules each consists of 58 amino acid residues, including four cysteine residues. Among these cysteine residues, the first residue forms a disulfide bonds with the third one; the second cysteine residue bonds to the fourth one to form another

disulfide bridge. In this manner, a hydrophobic pocket is generated which probably explains the interaction with gelatin [4].

MMP-9 is initially synthesized and secreted as an inactive monomer. In many cell types, it may form homodimers after being secreted, and it also can form a covalent complex with neutrophil gelatinase B associated lipocalin (NGAL). With non-reducing SDS-PAGE, these various forms of MMP-9 can be observed, but only monomer is visualized on a reducing SDS-PAGE gel. This phenomenon indicates the dimerization is stabilized by a covalent disulfide bond.

A three-dimensional structure of full-length MMP-9 has not been elucidated. The possible reason is that the flexibility of MMP-9 hinders the crystallization. Recombinant catalytic domain of MMP-9 and various mutants have been crystallized and their structures have been determined by X-ray crystallography, e.g., PDB 1GKC, PDB 1GKD and 1L6G. The crystallographic structure of the hemopexin domain of MMP-9 has also been determined by Hyunju Cha in 2002, the PDB code is 1ITV [21].

1.4.3. The hemopexin domain of MMP-9

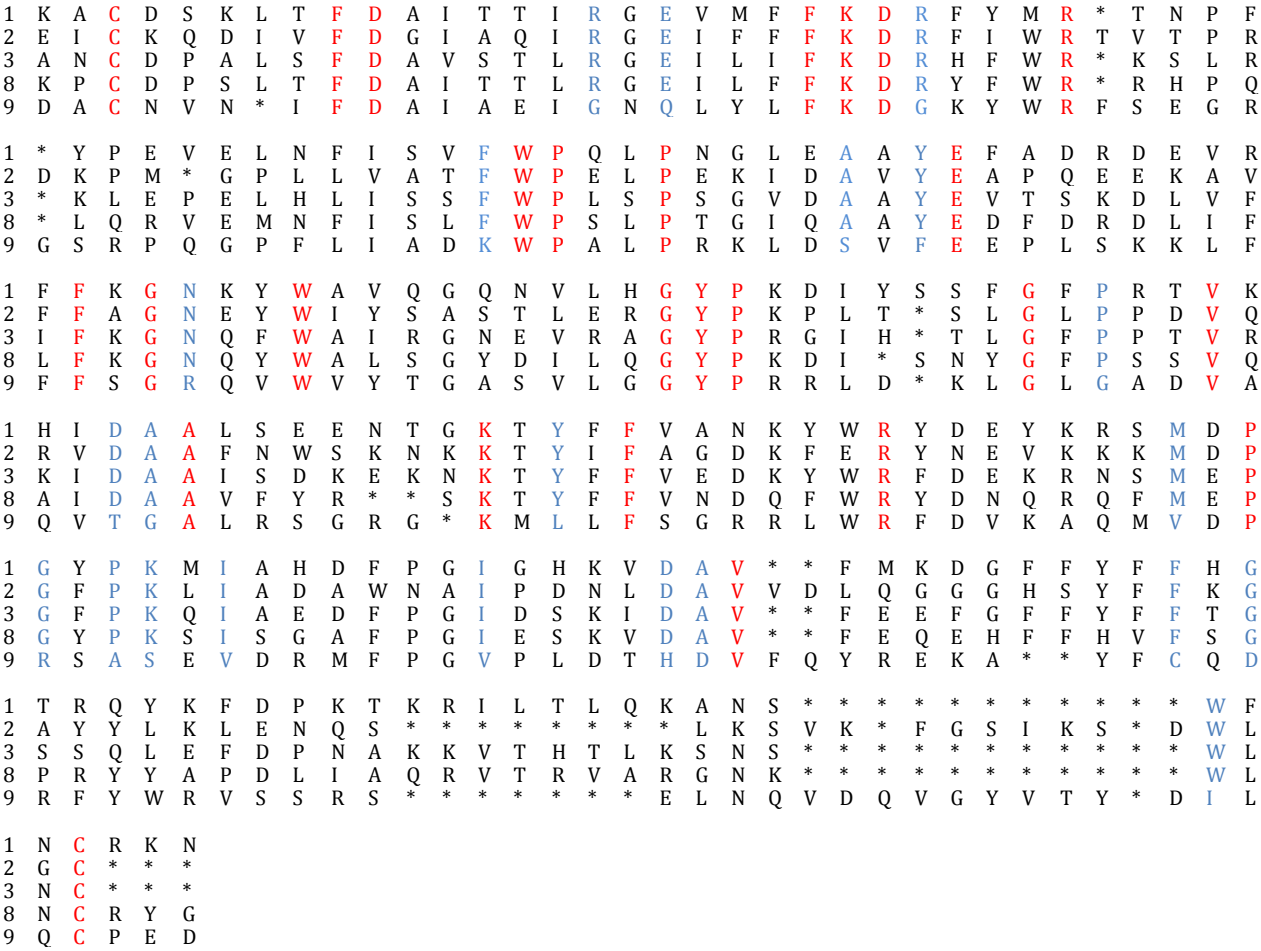
In addition to the proteolytic function of MMPs, increasing attention is being paid on their non-proteolytic functions which are affected by the hemopexin domain. In this study, we are screening for small molecules that bind to the hemopexin domain of MMP-9 with high selectivity. The hemopexin domain of MMP-9 was chosen as a target for two reasons.

First, the hemopexin domain of MMP-9 is thought to play an important role in cancer cell migration, independent of the proteolytic activity of the catalytic domain. In ovarian cancer, the increased level of proMMP-9, an inactive form of MMP-9, is related to

poor survival [22]. This phenomenon indicates the function of noncatalytic domain of MMP-9 in cancer progression. Recent research further shows that the hemopexin domain of proMMP-9 is capable of enhancing epithelial cell migration by proMMP-9 – CDD4 – EGFR – MAPK pathway [13], and the formation of homodimer of MMP-9 through hemopexin domain is the prerequisite condition [23].

Second, the structural features of the MMP-9 hemopexin domain provide a basis for identification of inhibitors with high selectivity. Most previous MMP inhibitors do not have the ability to discriminate different members in MMP family, resulting in many side effects. This is because they target the catalytic domain of MMPs where the amino acid sequences are highly conserved among members in MMP family (43%-65%) [24]. In contrast, less amino acid sequences are shared between hemopexin domains of MMPs (25%-33% [24]; Figure 4). The structural difference between hemopexin domains achieve their various biological functions, such as substrate recognition and selectively interacting with tissue inhibitors of metalloproteinases (TIMPs) [25].

Figure 4. Amino acid sequences alignment of the hemopexin domains

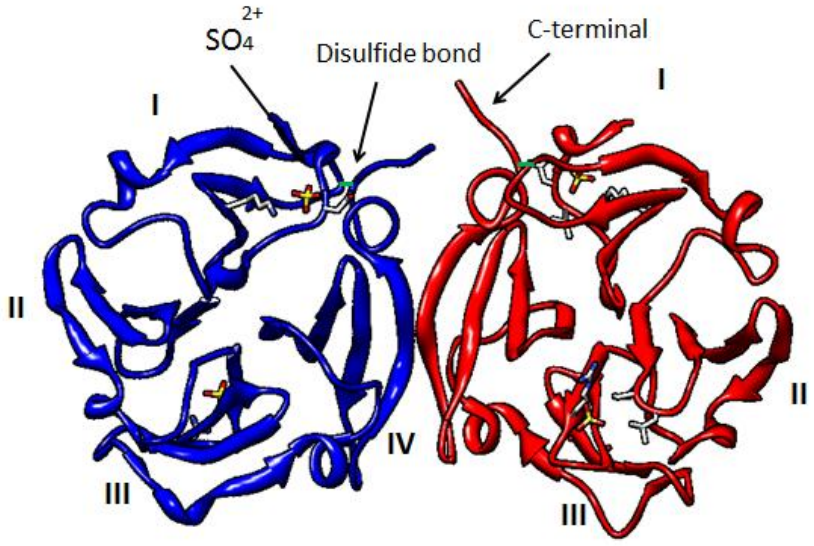


The first column contains the names of the aligned proteins. They are the hemopexin domains of MMP-1, MMP-2, MMP-3, MMP-8, and MMP-9 respectively. The consensus sequences which contain the same amino acid residue at a particular alignment position are colored by red. The positions where all aligned proteins share the same amino acid residues except MMP-9 are colored by blue.

The crystallographic structure of the hemopexin domain of MMP-9 was published by Hyunju Cha and his colleagues in 2002 (Figure 5 [21]). The hemopexin domain of MMP-9 exhibits a four-bladed-β propeller structure, similar to the hemopexin domain structures of all other MMPs. There are four antiparallel β-strands in each blade and all blades are connected to each other by a peptide loop. Blade I and IV are connected by a disulfide bond between cysteine 516 and cysteine 704, maintaining the structural integrity for the disc-like shape. This intramolecular disulfide bond is highly conserved among all previous

determined hemopexin domain structures. The hemopexin domain of MMP-9 exists in both a monomeric and a dimeric state, and the homodimerization is reduction sensitive. Therefore, the homodimerization is believed to be disulfide bond linked. However, Hyunju Cha obtained the crystallographic structure of the dimer of the hemopexin domain of MMP-9 and shown that dimerization is not caused by intermolecular disulfide linkages. The crystallographic structure shows that the homodimerization is caused by hydrophobic interaction between blade IV of each subunit (Figure 5). The reduction sensitivity of dimerization is due to intramolecular disulfide bond between blade I and IV. This intramolecular covalent linkage is important to the structural integrity of four-bladed-β-propeller. The reducing agents cleave this disulfide bond and disrupt the β-propeller structure, resulting in the dissociation of the MMP-9 hemopexin domain dimer.

Figure 5. Structure of the homodimeric hemopexin domain of MMP-9



Recently, a small molecule, N-[4-(difluoromethoxy) phenyl]-2-[(4-oxo-6-propyl-1H-pyrimidin-2-yl) sulfanyl]-acetamide (Zinc ID 8580066, Figure 6) was illustrated to be able to selectively bind to hemopexin-9, disrupting MMP-9-regulated cancer cell migration

independent of catalytic activity of MMP-9. The mechanism is that the binding of small molecule to hemopexin domain of MMP-9 disrupted MMP-9 homodimerization which is prerequisite condition for initiating downstream signaling pathway [24]. The discovery of this small molecule validated the possibility of hemopexin domain of MMP-9 as anticancer drug target directly.

1.5. Computational method: Docking

Docking is an efficient computational method to implement structure-based drug design by simulating molecular recognition which involves both geometric fitting and chemical fitting. The aim of docking is to search the most energetically favorable binding modes of a small drug-like molecule, or a ligand, within a binding pocket of a known receptor [26]. There are three important and practical applications of docking: 1. if there is crystallographic structure of receptor and ligand, it will be easy to find the most energetically favorable binding poses. 2. If there is only structure of receptor, best ligands will be found by virtually screening database of small molecules against active site of the receptor. In docking, the best ligands are considered to be the ones which bind to receptor with lowest binding energy. 3. With the crystallographic structure of protein and its substrate, docking can be used to look for possible binding sites on this protein. The second application is widely used in structure-based drug design, and accumulating evidence has shown this computational approach is effective and promising [24].

DOCK 6.0, as an extension of previous Dock version, adapts a flexible docking algorithm to efficiently treat the flexibility of ligands [27]. The flexibility docking addresses

the trickiest problem in virtual screening and enhances the accuracy of forecasting the binding affinity of ligands with receptor. In flexible docking, the anchor, the largest functional group, is selected and the rest parts of ligand are divided into different layers. The anchor is oriented to binding site of receptor, and then each layer grows from the anchor until the full molecule is formed. In this process, a large number of anchor positions are generated, which are then pruned to retain a certain number of conformation with lowest energy score and highest positional diversity. When the whole is completed, all strains generated during the growth procedure are released and all conformations are reached the local optimization [26].

In this study, docking simulation software, DOCK 6.0, was used to screen a subset of Zinc database against the hemopexin domain of MMP-9. The emergence of ZINC database facilitates the development of structure-based drug design [28]. ZINC database is an exhaustive library of small drug-like molecules which are commercially available. Molecules in ZINC database can be directly used for virtual screening, because they are added with hydrogens and charges, and are stored in mol2 format. In addition to well-prepared molecular format, ZINC database also provide necessary information about molecule property, such as the number of rotatable bonds. Some of molecule file even includes multiple tautomeric forms.

1.6. Project overview

Due to the failure of most previous MMP inhibitors, the long-term goal of our project is to identify novel compounds which selectively inhibit biological function of MMP-9. The

specific aim of this study is to identify small molecules that bind to the hemopexin domain of MMP-9.

We chose the hemopexin domain of MMP-9 as the target for two reasons. First, the hemopexin domain of MMP-9 promotes cancer cell migration, independent of the proteolytic activity of the catalytic domain [13]. In ovarian cancer, the increased level of proMMP-9, inactive form of MMP-9, is related to poor survival [15]. Recent research further shows that the hemopexin homodimer is essential to pro-MMP-9 enhanced cell migration; and TIMP-1 (tissue inhibitor of metalloproteinase-1) inhibit cell migration by interacting with the hemopexin domain of proMMP-9 [23]. Second, the structural features of the MMP-9 hemopexin domain provide a basis for the selecting of inhibitors with high selectivity. Previous MMP inhibitors cannot discriminate individual members of the MMP family, because they target the catalytic domain of MMPs. The amino acids sequence in the catalytic domain is highly conserved among MMPs. In contrast, less amino acid sequences are shared among the hemopexin domain of MMPs. In addition, some MMPs do not have the hemopexin domain: MMP-7, -23, and -26. Based on low similarity between the hemopexin domains, inhibitors that target the hemopexin domain of MMP-9 are proposed to have selectivity for MMP-9.

We utilized a computational method, docking, to screen a subset of ZINC database against the hemopexin domain of MMP-9. Dr. Jin Wang also used docking to screen the NCI database against the hemopexin domain of MMP-9. We then employed a biochemical fluorescence assay to evaluate the binding affinity of computational hits. In fluorescence

assay, we tested the computational hits form Dr. Jin Wang's lab. We tested Dr. Wang's results, because we could obtain free compound samples from the National Cancer Institute.

Chapter 2. Experimental section

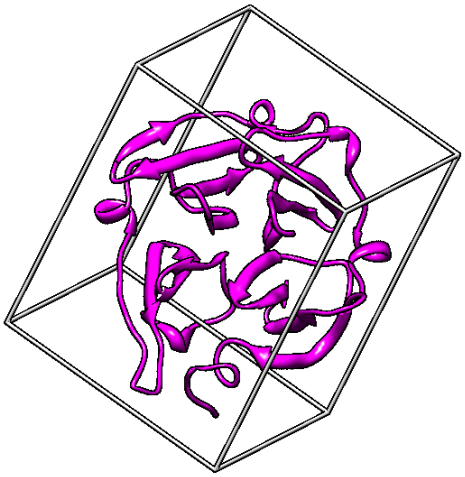
2.1. Docking

2.1.1. Program setups

We utilized DOCK 6.0 to identify the potential ligand binding sites in the hemopexin domain of MMP-9 and the potential ligands. The X-ray crystallographic structures of proteins were downloaded from Protein Data Bank and prepared in software Chimera. The receptors with hydrogens were converted to a mol2 file and the charges were calculated using AM1-BCC force.

The grid size was set to 100 x 100 x 100 points with a grid spacing of 0.3 Å. The box was centered on the original ligand position in crystal structure complex. The grid box included the whole protein and in order to provide enough space for ligand translational and rotational walk (Figure 6). Step sizes of 1.0 Å for translation and the cluster RMSD threshold was set to 2.0 Å. The number of ligands per cluster was set to 100.

Figure 6. Image of box and receptor



The box includes the whole protein to provide enough for ligand translational and rotational walk.

2.1.2. Control experiments

Before running the virtual screening against the hemopexin domain of MMP-9, four control experiments were conducted to verify the reliability of docking approach and program setup.

Only one crystallographic structure of hemopexin domain of MMP-9 is in the Protein Data Bank, 1ITV [21], which has a sulfate ion as ligand. So the first control experiment was re-docking this sulfate ion into its binding site. The second control was done by using an inhibitor-catalytic domain of MMP-9 complex (PDB code: 1GKC; resolution: 2.30 Å [29]). The sample in third control experiment is an inhibitor-catalytic domain of MMP-3 (PDB code: 1HY7; resolution: 1.50 Å [30]). By separating ligands from receptors and re-docking them back into the binding site, we got the root mean square deviation (RMSD) which can show differences between the original binding pose and the new binding pose generated by our simulation. Detailed information about *three crystal structures used in control experiments* is shown in Table 2.

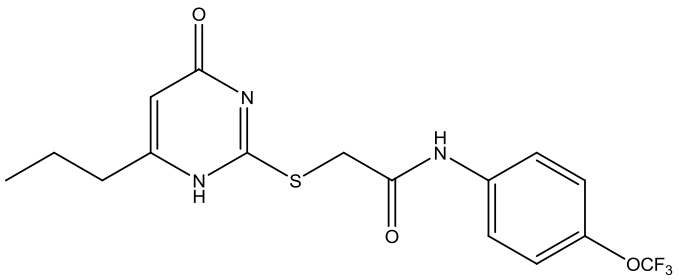
Table 2. Three crystal structures used in control experiments

PDB code	Protein/ Ligand	Resolution (Å)
1ITV	Hemopexin domain of MMP-9/ sulfate ion	1.95
1GKC	Catalytic domain of MMP-9/ BUM+STN	2.30
1HY7	Catalytic domain of MMP-3/ MBS	1.50

BUM : 2-AMINO-N,3,3-TRIMETHYLBUTANAMIDE
STN : 2-{[FORMYL(HYDROXY)AMINO]METHYL}-4-METHYLPENTANOIC ACID
MBS: R-2-{[4'-METHOXY-(1,1'-BIPHENYL)-4-YL]-SULFONYL}-AMINO-6-METHOXY-HEX-4-YNOIC ACID

The fourth control experiment was docking one small drug-like molecule (Zinc Code: 8580066, Figure 7) into the hemopexin domain of MMP-9. This molecule comes from paper published by Antoine Dufour in 2011 [24].

Figure 7. Structure of compound ZINC 8580066



2.1.3. Virtual screening

After the verification of our docking algorithm by control experiments, we applied the same computational parameters to screen for potential ligands against hemopexin domain of MMP-9. In previous study, hemodimerization is observed under X-ray crystallographic conditions (PDB code: 1ITV; [21]) and in cell culture. The subunits are similar, but not completely symmetric. In this study, we focused on chain A of MMP-9 hemopexin domain for docking. Finally, molecules bound to the PEX domain of MMP-9 were ranked based on their cluster size, grid score (energies).

2.2. Preparation of recombinant hemopexin domain of MMP-9

2.2.1. Protein expression

The pET-16b vector was employed to prepare the construct which placed the 6xHis-tag at the N-terminus of the hemopexin domain of MMP-9. The 6xHis-tagged hemopexin domain of MMP-9 was expressed in BL21 *E. coli strain*. The *E. coli* was cultured in 20 ml Luria broth media (LB media) with 100 mg/ml carboxicillin 14 h at 37 ^{o}C. Then the preculture was transferred into 1 L LB media. Once the OD_{600} reached 0.6 to 0.8, *E. coli* was induced at least 4 h with 1 mM IPTG. After spinning down (40000 rpm, 35 minutes), the *E. coli* cell pellet was suspended in 5 mL lysis buffer (50 mM Tris-HCl, 0.1 mM EDTA, 100 mM NaCl, pH=8.0) with 1 mM PMSF as serine protease inhibitor and 1 mg/ml lysozyme at RT. After 20 min, the *E. coli* sample was sonicated on ice (three 30 seconds bursts). DNase I (5 μg/ml) was added into lysate, in order to degrade and remove unwanted DNA from samples. After another 20 min, the lysate was spun down (15000 x g, 15 minutes, 4 ^{o}C). Both 6 mL supernatant and pellet were saved for SDS-PAGE analysis.

2.2.2. Protein purification

The recombinant 6xHis-tagged hemopexin domain of MMP-9 was purified under denaturing conditions by using a nickel-nitrilotriacetic acid (Ni-NTA) metal-affinity chromatography matrix. The insoluble lysate pellet which contained recombinant hemopexin domain of MMP-9 was suspended in 5 mL guanidine hydrochloride buffer (100 mM NaH_2PO_4, 10 mM Tris-HCl, 6 M GnHCl, pH=8.0), and incubated for 1 h at RT. After spinning down (15000 x g, 15 minutes, 4 ^{o}C), the supernatant containing solubilized

hemopexin domain of MMP-9 was added into 1 mL 50% Ni-NTA beads and incubated for 1 h at RT, to allow sufficient binding between Ni-NTA between 6xHis-tagged beads. The bead-protein complex was then washed with 4 mL wash buffer (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH=6.3) 3 times, and each wash sample was saved for SDS-PAGE analysis.

After washing, the bead-protein complex was eluted with 1 mL elution buffer 1 (100 mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH=5.9) and 500 uL elution buffer 2 (100mM NaH_2PO_4, 10 mM Tris-HCl, 8 M urea, pH=4.0) 4 times each at RT. The 8 mL eluate was analyzed by SDS-PAGE.

2.2.3. Protein refolding

The native hemopexin domain of MMP-9 is required in this study, thus a refolding step followed. In order to remove urea, a denaturant, the purified protein (8 mL, 1 mg/mL) was dialyzed against 800 mL dialysis buffer (10 mM Tris-HCl, 250 mM NaCl, 40 mM arginine) with decreasing concentration of urea (from 8 M to 0 M) at 4 ^{o}C. The first dialysis buffer (800 mL) contained 7 M urea; the second one contained 6 M urea. The buffer was changed each 4 h, until the concentration of urea was decreased to 0 M.

2.2.4. Dimerization analysis

The refolded protein was analyzed by gel filtration chromatograph (Superdex 200 10/300 GL high performance column). The analysis buffer contained 10 mM Tris-HCl, 250

mM NaCl, 40 mM arginine. The protein (250 nmol/L) was analyzed at 4 ^{o}C, pH 7.0 and a flow rate of 0.5 ml/min.

2.3. Binding assay

The binding of compounds to the hemopexin domain of MMP-9 was tested by monitoring the change of tryptophan fluorescence emission from the hemopexin domain of MMP-9.

2.3.1. Positive control experiment

Before testing the computational hits, we chose compound 8580066 as positive control to setup experimental conditions. We titrated compound 8580066 (Figure 6) which is published by Antoine Dufour in 2011 [24] into the purified hemopexin domain of MMP-9 at RT. The purified recombinant hemopexin domain of MMP-9 was diluted in buffer (10 mM Tris-HCl , 250 mM NaCl, 40 mM arginine, pH 7.0) to a concentration of 100 nM in the presence or absence of compound 8580066. The protein sample was excited at 280 nm. The emission scans were collected from 310 to 400 nm with slit width of 2.0nm. The fluorescence emission was measured three times at each concentration. The K_d was caculated by software Origin 8.0, fitting to the Equation 1, in which λ_{max} is the wavelength at which maximal fluorescence of the protein was observed.

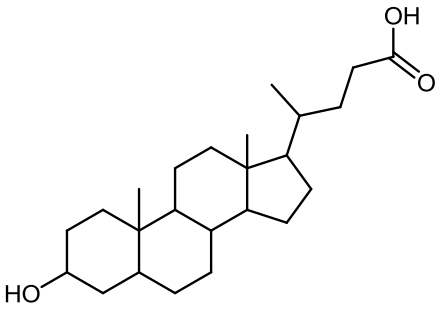
$$\Delta\lambda max = \frac{capacity[compound]}{K_d + [compound]} \quad \textit{(Equation 1)}$$

2.3.2. Test of computational hits

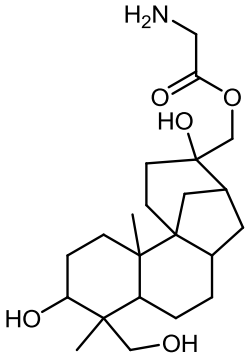
Seven compounds with high calculated binding affinity for hemopexin domain of MMP-9 were tested by fluorescence assay (Figure 8) at RT. These compounds were docked by Xiliang Zheng in Dr. Jin Wang's lab. We also tested 5 compounds in NCI diversity set which did not show high calculated binding affinity. The change of fluorescence emission intensity was measured by multiple-detection plate reader which is mounted in Gen5 Microplate software. All these compounds were diluted in buffer (1 mM DMSO). The protein was diluted in buffer (10 mM Tris-HCl , 250 mM NaCl, 40 mM arginine, pH 7.0) to a concentration of 100 nM. The excitation filter and emission filter were 284/10 nm and 340/30 nm respectively. The fluorescence emission was measured three times at each concentration of compound. The K_d was calculated by OriginPro 8.0 software, fitting to Equation 2. In equation 2, Δ intensity is the change of fluorescence intensity at the presence of compounds.

$$\Delta intensity = \frac{capacity[compound]}{K_d + [compound]} \quad \text{(Equation 2)}$$

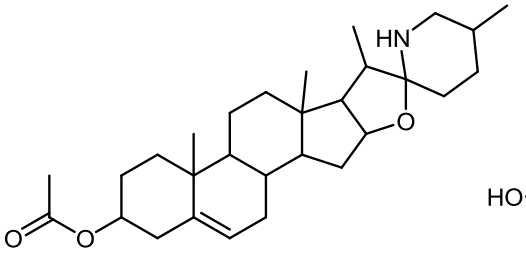
Figure 8. Structures of computational hits



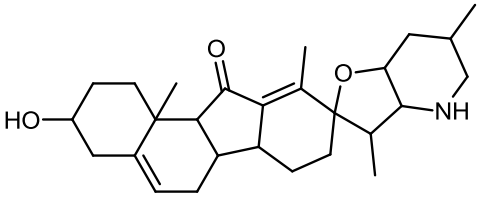
NCI 683770



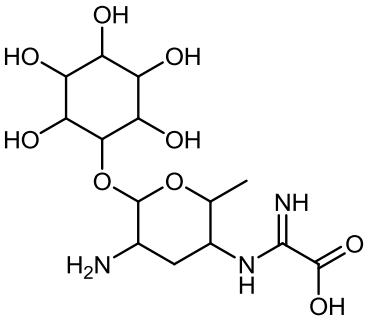
NCI 303812



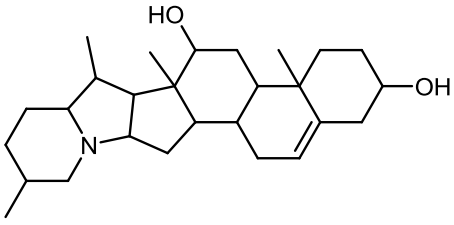
NCI 96021



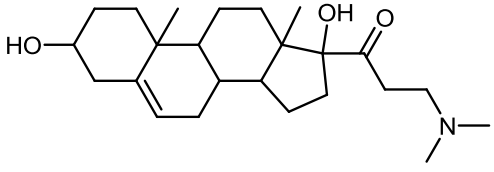
NCI 7520



NCI 100585



NCI 76026



NCI 48630

Chapter 3.Results and discussion section

3.1. Results

3.1.1. Docking

3.1.1.1 Results in Control experiments

The hemopexin domain is a novel target for anticancer drug design, thus seldom studies have been done to look for inhibitor targeting it. In addition, the hemopexin domain of MMP-9 is a disc-like shaped structure composed mostly of hydrophobic surfaces making this domain difficult to target with drugs [31]. Therefore, there no crystallographic structure of inhibitor-hemopexin domain of MMP-9 complex is available. The only one crystallographic structure of hemopexin domain of MMP-9 in Protein Data Bank is 1ITV, which has a sulfate ion as ligand. The first control experiment was re-docking this sulfate ion into its binding site. However, it is hard to reproduce the same or correct binding pose for sulfate ion, because the sulfate ion is too small to provide enough information. Taking the deficiency of sulfate ion into consideration, we did other three control experiments.

In control experiments, a good result is usually expected be that the binding pose with lowest grid score should have lowest RMSD. The RMSD and the binding poses obtained in control experiments are shown in Table 3 and Figures 9 to 12. The root mean square deviations (RMSD) of three control experiments are less than 2 Å, which are considered to be positive results. However, the result in docking sulfate ion in hemopexin domain of MMP-9 is problematic. The pose with lowest RMSD (1.87029 Å) is not the one

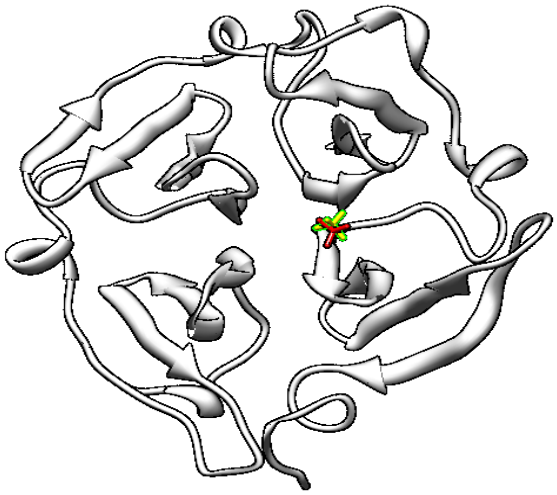
has lowest grid score (Figure 9). This situation is caused by the size of sulfate ion: the sulfate ion is too small to provide enough information for computation, resulting in the bad reproducibility. The results in the rest of experiments are reasonable: all binding poses with lowest RMSD have lowest grid score; and all RMSD are less than 2 Å. These positive results confirmed the feasibility of this docking approach. The result in fourth control experiment is also positive (Figure 15). The new pose is consistent with the original one. Comparing with previous control experiments, the fourth one provides solid support to our docking algorithm directly, because the docking was conducted by using the hemopexin domain of MMP-9 and a known ligand.

Table 3. RMSD obtained in control experiments

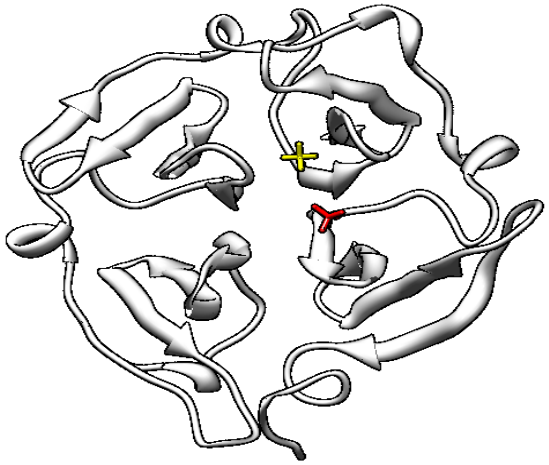
PDB code	*Resolution (Å)*	*the top ranked RMSD (Å)*
1ITV	*1.95*	*1.87*
1GKC	*2.30*	*1.16*
1HY7	*1.50*	*0.98*

Figure 9. The first control experiment: docking sulfate ion into the hemopexin domain of MMP-9

a

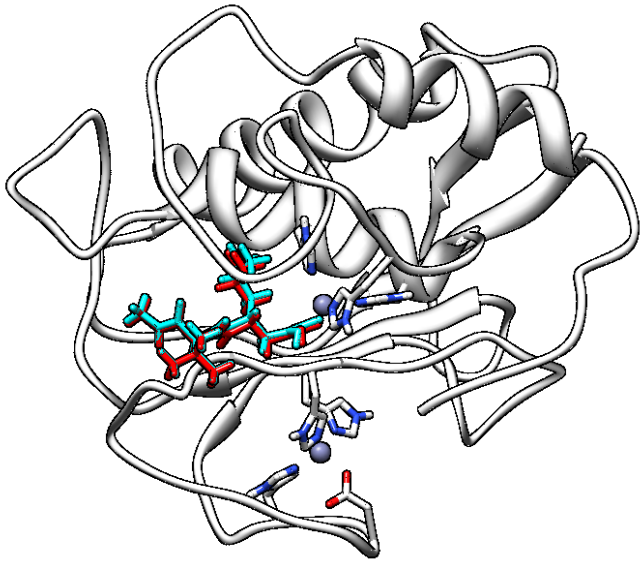


b



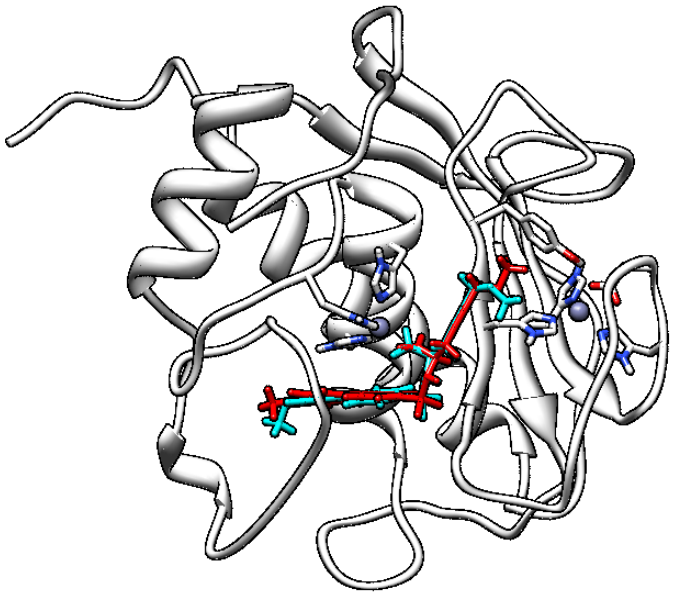
The original pose is in red, and the docking result is in yellow. ***a****. the pose with lowest RMSD (RMSD: 1.87; Grid score: -27.91)* ***b****. the pose with best grid score (RMSD: 6.89; Grid score: -36.89)*

Figure 10. The second control experiment: docking ligand into catalytic domain of mmp-9



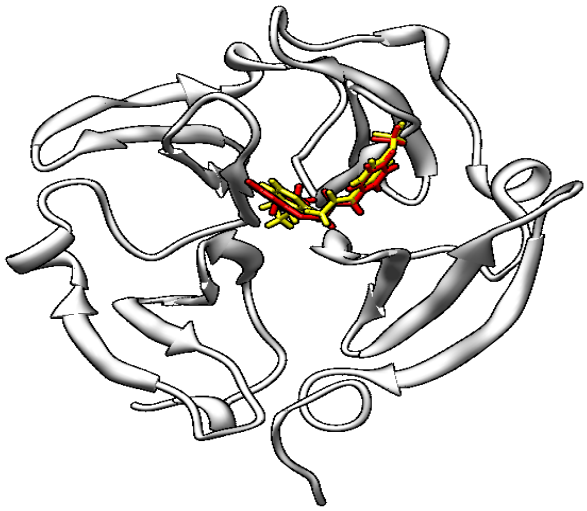
The original pose is in cyan; the docking result is in red (RMSD: 1.16; Grid score: -77.84).

Figure 11. The third control experiment: docking ligand into the catalytic domain of mmp-3.



The original pose is in red; the docking result is in cyan. (RMSD: 0.98; Grid score: -96.29)

Figure 12. The fourth control experiment: docking ligand into hemopexin domain of MMP-9

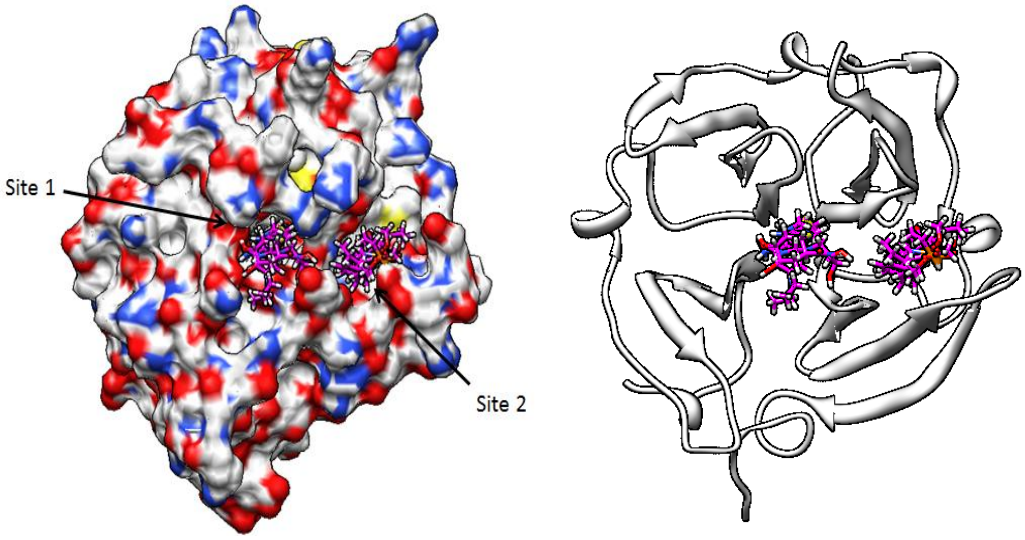


The original binding pose is red; the docking result is in yellow.

3.1.1.2. Results in virtual screening

Molecules were docked to two sites: Site 1 is located at the interface of four blades; site 2 is on the interface of blade III and IV (Figure 13).

Figure 13. Two possible binding sites in the hemopexin domain of MMP-9

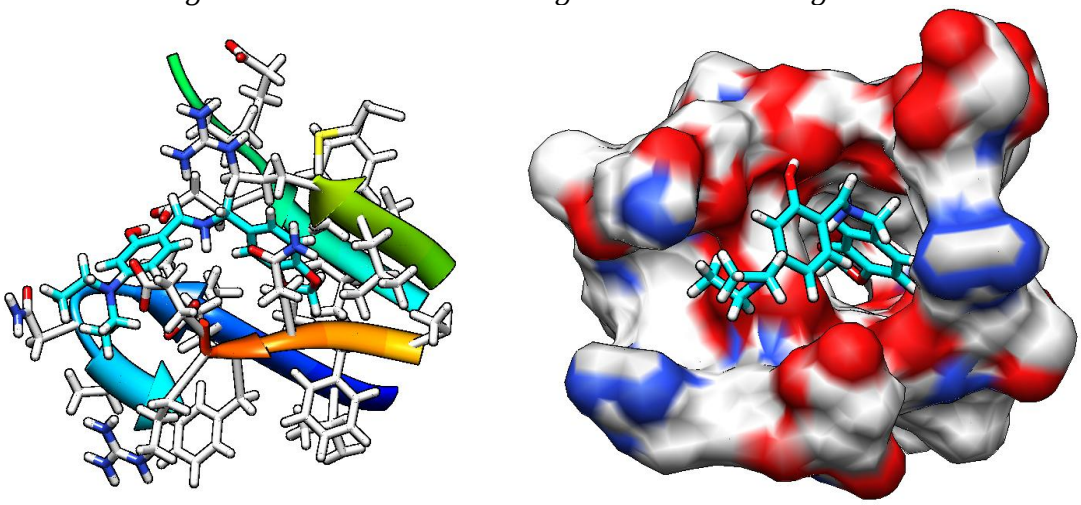


Left figure is shown in electronic surface; right one is shown in ribbon.

3.1.1.2.1. Binding site 1

The crystallographic structure of the hemopexin domain of MMP-9 shows a large cavity in the center of the top face of barrel which is formed by the innermost strands of all 4 blades (Figure 13). This is a compacted and deep cavity. The molecule with best grid score is shown in Figure 14 and all residues within 5 Å from the ligand were retained.

Figure 14. Molecule with best grid score in binding site 1



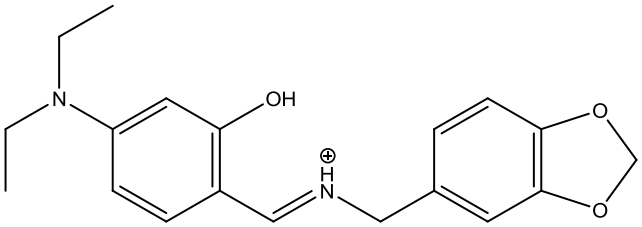
Left figure is shown in ribbon; right one is shown in electronic surface.

Detailed information of three top ranked molecules binding to site 1 is shown in Table 4. Their structures are shown in Figure 15.

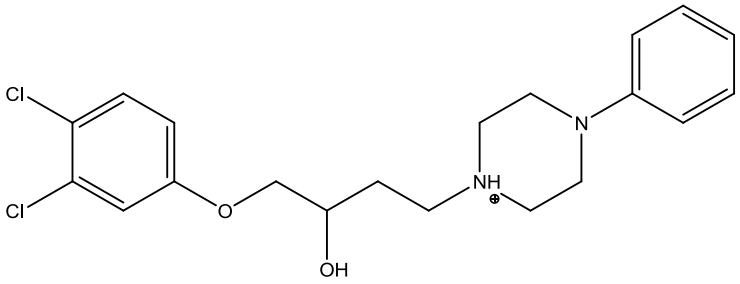
Table 4. Three top ranked molecules in binding site 1

Zinc code	*Grid score energy (-Kcal/mol)*	*Van der waals energy (-Kcal/mol)*	*Electrostatic energy (-Kcal/mol)*
22726619	*-77.25*	*-49.86*	*-27.40*
24981806	*-73.12*	*-50.57*	*-22.55*
20504103	*-72.14*	*-53.19*	*-18.95*

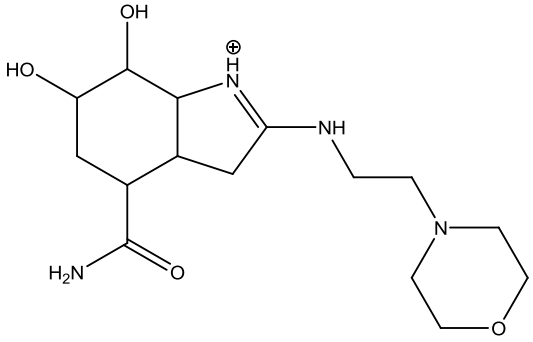
Figure 15. Structures of molecules in binding site 1



Zinc No. 22726619 (Grid score: -72.14)



Zinc No. 24981806 (Grid score: -73.12)



Zinc No. 20504103 (Gird score: -77.25)

3.1.1.2.2. Binding site 2

Binding site 2 is a loose and shallow cavity compared with site 1. The molecules with the lowest grid score are shown in Figure 16. All residues within 8 Å from the ligand were retained. Detailed information of three top ranked molecules which bind to site 2 is shown in Table 5. Their structures are shown in Figure 17. The van der waals energy of

molecule 03958322 and molecule 05478300 is 14.85 and 19.28 respectively, which is not energetically favorable. Their interactions with receptor are dominated by electrostatic interaction, which is consistent with structure features: they have negative formal charge.

Figure 16. *Molecules with best grid score in binding site 2*

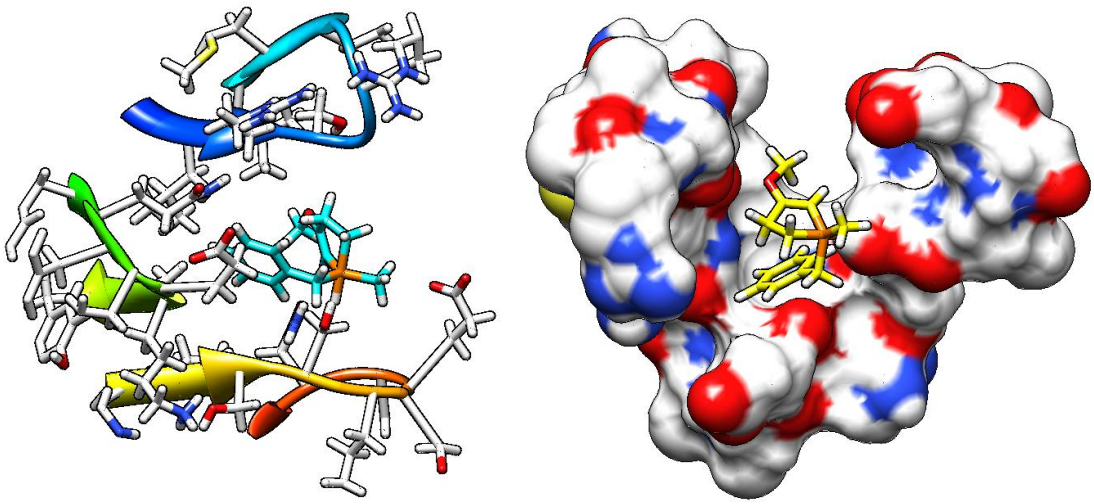
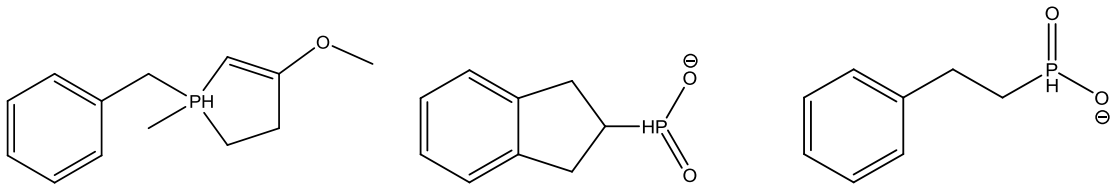


Table 5. Three top ranked molecules in binding site 2

Zinc code	Grid score energy (-Kcal/mol)	Van der waals energy (-Kcal/mol)	Electrostatic energy (-Kcal/mol)
6482820	-155.63	-8.35	-147.28
03958322	-133.47	14.85	-148.32
05478300	-129.08	19.28	-148.36

Figure 17. Structures of molecules in binding site 2



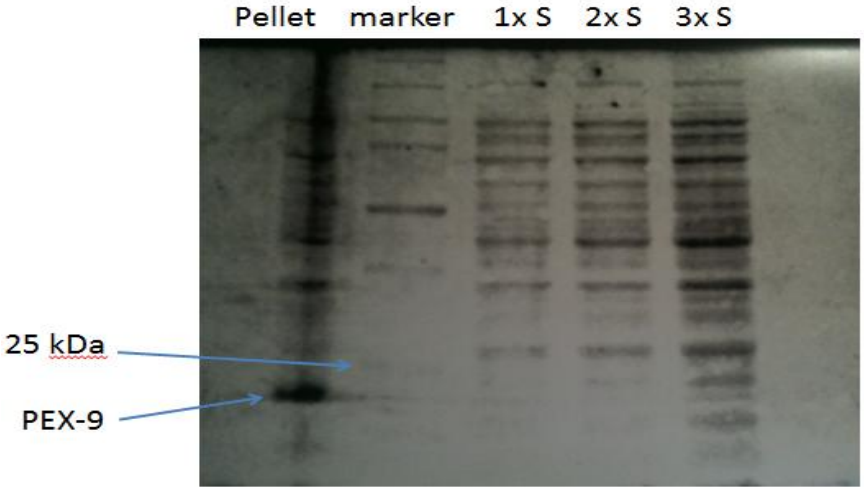
Zinc No. 6482820 *Zinc No. 03958322* *Zinc No. 05478300*

Grid score: -133.47 *Grid score: -155.63* *Grid score: -129.08*

3.1.2. Protein preparation

The 6xHis-tagged hemopexin domain of MMP-9 was expressed in *E. coli*. We chose *E. coli*, rather than mammalian cell system, because of its high protein yield. After expression, the *E. coli* lysate was spun down. Cell pellet and supernatant were analyzed by SDS-PAGE gel (Figure 18).

Figure 18. SDS-PAGE analysis for E.coli lysate



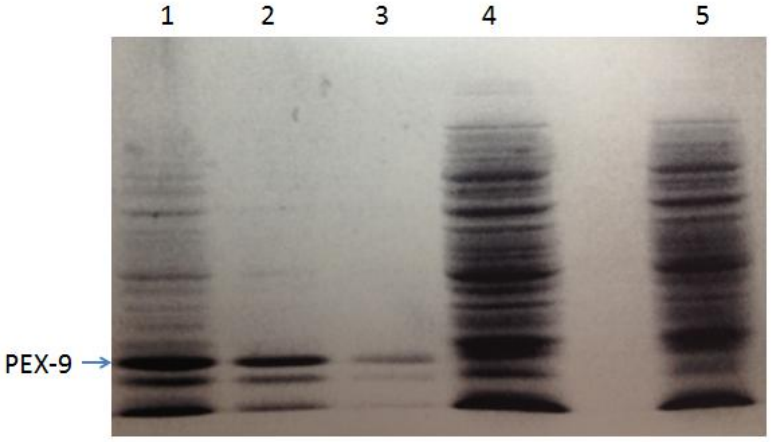
Well 1 is lysate pellet. Well 3-5 are different concentrations of lysate supernatant. S ---- Supernatant of lysate.

The calculated molecular weight of MMP-9 hemopexin domain is 21208 Dalton. Almost all the recombinant protein resided in the pellet as inclusion body, so the

supernatant of lysate was discarded and the protein was solubilized with the guanidine hydrochloride buffer (100 mM NaH_2PO_4, 10 mM Tris-Cl, 6 M GnHCl, pH 8.0).

In the protein purification step, the Ni-NTA matrix-protein complex was washed with wash buffer, and each wash sample was saved for SDS-PAGE analysis (Figure 19). According to SDS-PAGE analysis, most of the MMP-9 hemopexin domain was bound to Ni-NTA beads, and 3 washes are necessary to remove non-specific binding.

Figure 19. SDS-PAGE analysis for samples in purification step



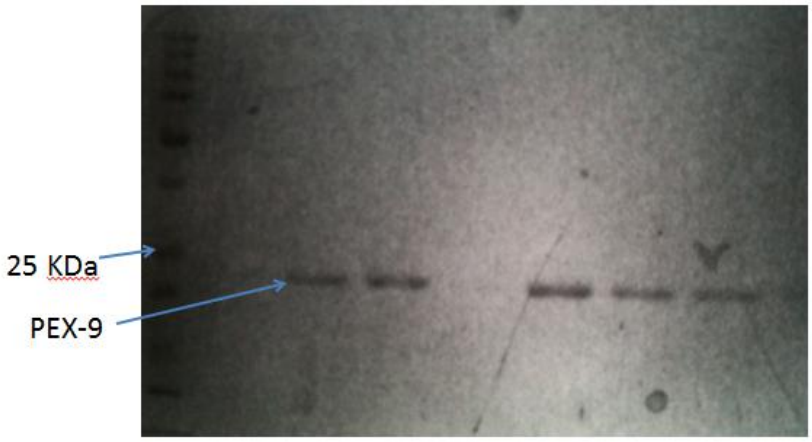
Well 1: the first wash sample; Well 2: the second wash sample; Well 3, the third wash sample; Well 4: intact supernatant of lysate; Well 5: the lysate supernatant after adding beads.

In the purification step, 8 M urea was used in the elution buffer, which is denaturing condition. There are two advantages to purify the protein under denaturing conditions. First, it is a good method to deal with overexpression of the protein. Overexpression of the protein results in the formation of inclusion bodies, which localize in the cell pellet. The employment of denaturant helps to dissolve the protein. Second, the purification process is simple under denaturing conditions. The denaturing conditions allow the employment of low pH. By simply lowering the pH of wash buffer and elution buffers, the histidine residues chelating with nickel ion in matrix will obtain proton from the buffer, and thus

detach from the matrix. In contrast, imidazole is required under native condition purification. In this case, an extra chemical, imidazole, is introduced.

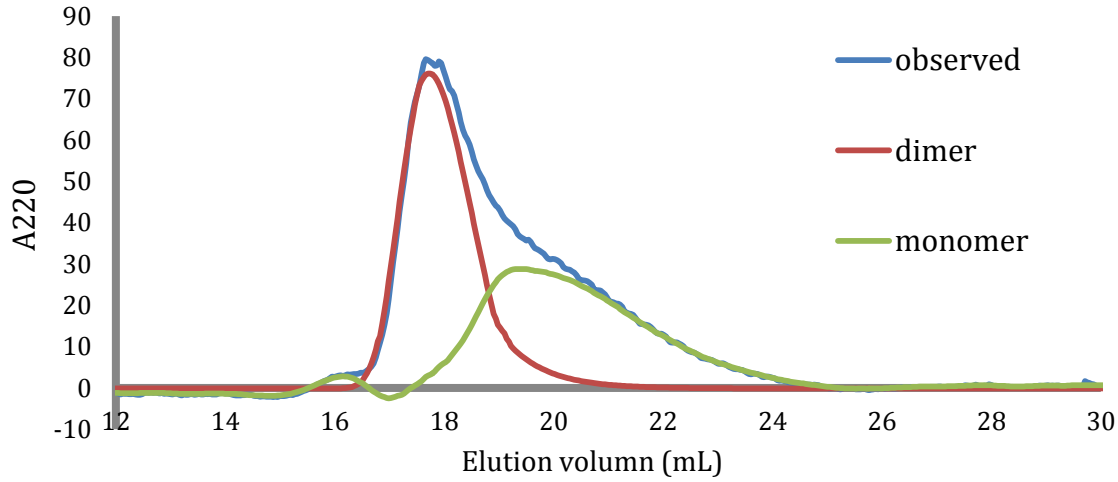
After purification and refolding, the eluate was analyzed by SDS-PAGE (Figure 20). Finally, 8 mg purified protein was obtained from a 1 L culture.

Figure 20. SDS-PAGE analysis for eluate



The dimerization of MMP-9 is prerequisite for MMP-9 promoted cell migration, and its dimerization is achieved through the hemopexin domain [23]. Therefore, it was necessary to check the oligomer form of the purified hemopexin domain of MMP-9. The refolded protein was analyzed by gel filtration chromatography, in order to determine the oligomerization state of the refolded hemopexin domain of MMP-9 (Figure 21).

Figure 21. Gel filtration chromatograph for purified hemopexin domain of MMP-9



The elution volume of the observed peak (blue) is 17.60 mL, and the corresponding molecular weight is 42268 Dalton. The calculated molecular weight of a hemopexin domain of MMP-9 subunit is 21208 Dalton. Therefore, the protein is a mixture of dimer and nomomer. In order to estimate the relative amounts of dimer and monomer, deconvolution calculations were performed on the mixture peak by OriginPro 8.0 software. The deconvolution process is able to decompose a complex peak in an ensemble of simple Gaussian peaks. The mixture peak (blue) was deconvoluted in two peaks related to the dimer (red) and monomer (green). According to the area under each peak, the ratio of dimer to monomer is 3: 5. The concentration of protein was 250 nM. Therefore, the dissociation constant of dimerization is 520 nM.

3.1.3. Binding assay

Because of the five tryptophan residue in the hemopexin domain of MMP-9, the binding between compounds and the MMP-9 hemopexin domain can be tested by a biophysical approach: fluorescence assay. The binding of compound to protein can result in the modification of protein conformation and the change of tryptophan fluorescence emission. As a result, we can test the binding by monitoring the change of tryptophan fluorescence emission by the hemopexin domain of MMP-9.

3.1.3.1. Positive control experiment

Before testing the computational hits, we chose compound 85880066 as positive control to setup experimental conditions. We titrated compound 8580066 into the hemopexin domain of MMP-9 (100 nM) and monitored the change of tryptophan fluorescence emission (Figure 23; Table 7). At the same time, corresponding amount of compound buffer was added into the protein as negative control (Figure 22, Table 6).

Table 6 and Figure 22. Negative control: adding compound buffer in protein

[c] of compounds (μM)	λmax
0	346.9
0.1	346.9
0.2	345.9
0.5	346.9
1	346.9
2	347.9
5	346.9
10	346.9
20	347.9
50	346.9
100	346.9

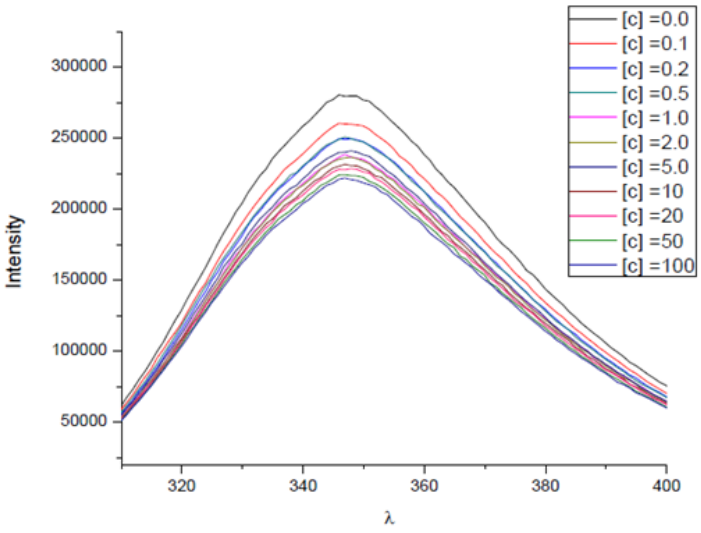
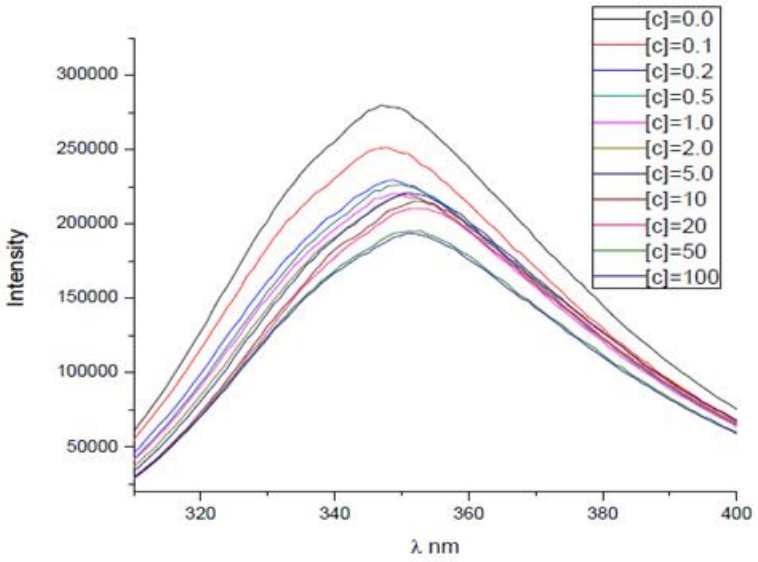


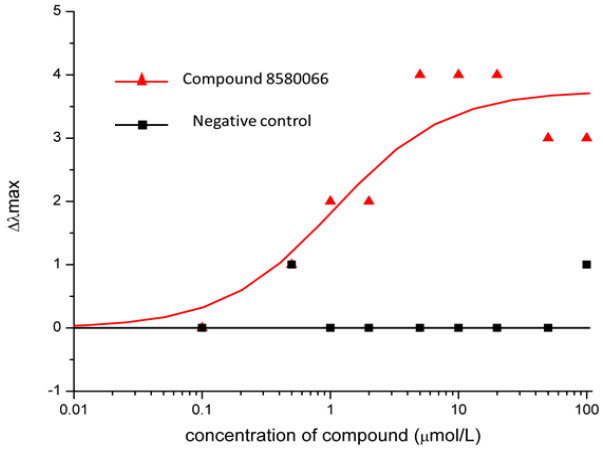
Table 7 and Figure 23. Experimental group: adding compound 8580066 into protein

[c] of compounds (μM)	λmax
0.0	346.9
0.1	346.9
0.2	347.9
0.5	348.9
1	348.9
2	350.9
5	350.9
10	350.9
20	350.9
50	349.9
100	349.9



There is no change on protein fluorescence emission generated in negative control. In contrast, the saturation of the purified hemopexin domain of MMP-9 by compound 8580066 resulted in 4 nm red shift in λ_{max} and quenching of fluorescence emission. The K_d for compound 8580066 binding to hemopexin domain of MMP-9 is 1.1 ± 0.30 uM (Figure 24).

Figure 24. Compound 8580066 binds to the hemopexin domain of MMP-9



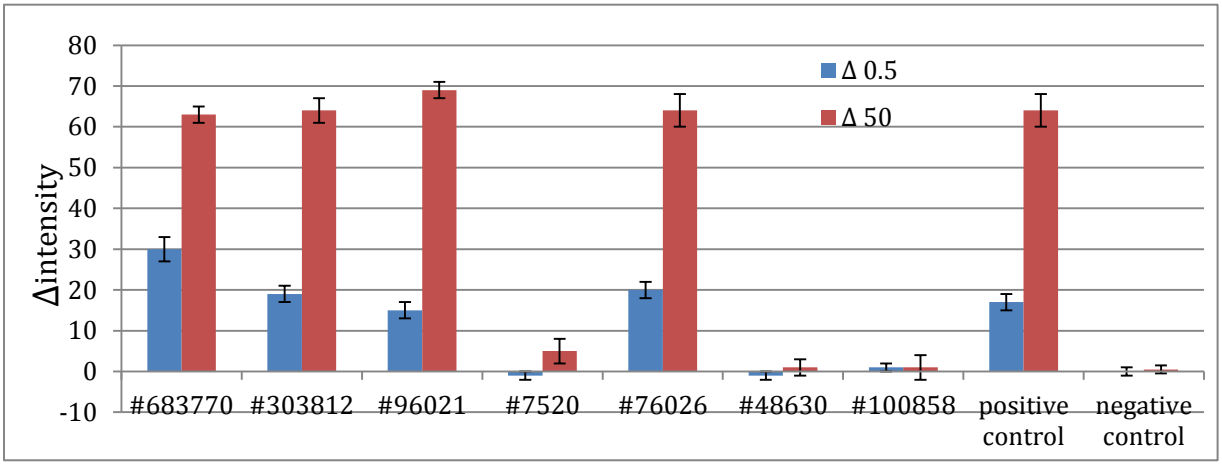
The λ_{max} was monitored upon titrating compound 8580066 into the purified recombinant hemopexin (100 nM) at pH 7.0, RT. The excitation wavelength was 280 nm.

3.1.3.2. Test of computational hits

Taking the compound 8580066 as positive control, we tested 7 computational hits (Figure 11) from Dr. Jin Wang's lab by measuring the intensity change of fluorescence emission.

The intensity of tryptophan fluorescence emission was measured by plate reader at compound concentration of 0.5 μM and 50 mM respectively. Corresponding amount of compound buffer solution was added into the protein as negative control experiment. The results are shown in Figure 25.

Figure 25. The fluorescence change of MMP-9 hemopexin domain in the presence of NCI compounds



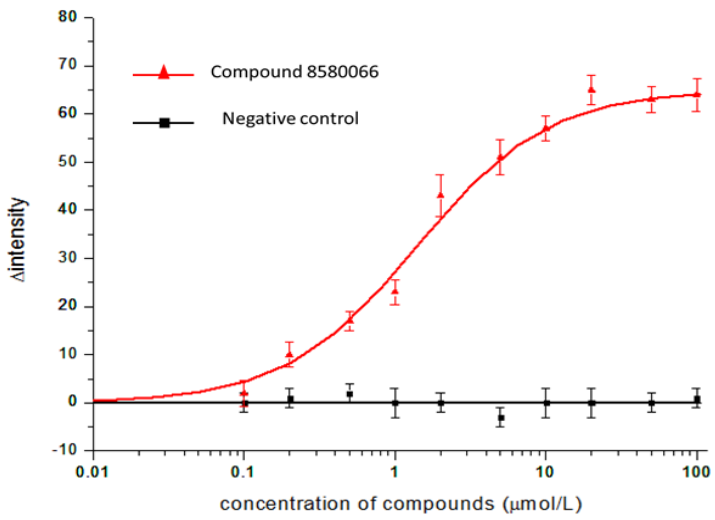
Δ 0.5 represents the fluorescence intensity change of MMP-9 hemopexin domain (100 nM, pH 7.0, RT) in the presence of 0.5 μM compounds.
Δ 50 represents the fluorescence intensity change of MMP-9 hemopexin domain (100 nM, pH 7.0, RT) in the presence of 50 μM compounds.

The results show that there are 4 compounds that have potential binding to the hemopexin domain of MMP-9. The fluorescence assay validated the feasibility of computational method docking.

Further, the binding affinity of these four compounds for the hemopexin domain was measured. The binding was monitored by a plate reader, rather than fluorometer, because the plate reader requires a smaller sample signal. Therefore, the fluorescence intensity change of the hemopexin domain served as indicator of binding. We reconstructed the positive control experiment first (Figure 26 and Table 8). The K_d is 1.4 ± 0.03 μM, which is consistent with previous experiment.

Table 8 and Figure 26. The binding of compound 8580066 to the hemopexin domain of MMP-9

[c] of compounds (μM)	*Δ intensity*
0	*0*
0.1	*2*
0.2	*10*
0.5	*17*
1	*23*
2	*42*
5	*51*
10	*57*
20	*65*
50	*63*
100	*64*



The intensity of tryptophan fluorescence emission was monitored upon titrating compound 8580066 into the purified recombinant hemopexin (100 nM) at pH 7.0, RT. The excitation and emission filters were 284/10 nm and 340/30 nm respectively.

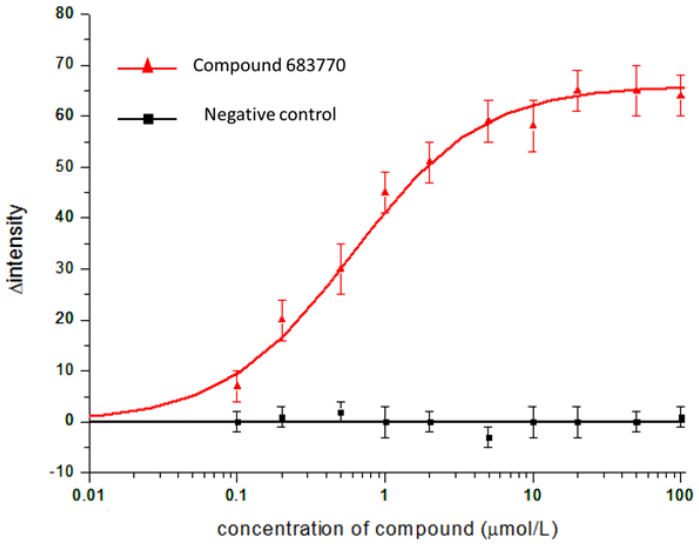
After validating the feasibility of our experimental methods, we measured the binding affinity of these four compounds respectively. The K_d's of these compound are listed in Table 9. The detailed information about fluorescence intensity change of these compounds is shown in Table 10 -13, Figure 27-30.

Table 9. The K_d's of compounds that bind to the hemopexin domain of MMP-9

compound No.	*K_d (μM)*
8580066	*1.4 ± 0.11*
683770	*0.6 ± 0.09*
303812	*1.1 ± 0.13*
96021	*1.6 ± 0.15*
76026	*1.1 ± 0.17*

Table 10 and Figure 27. The binding of compound 683770 to the hemopexin domain of MMP-9

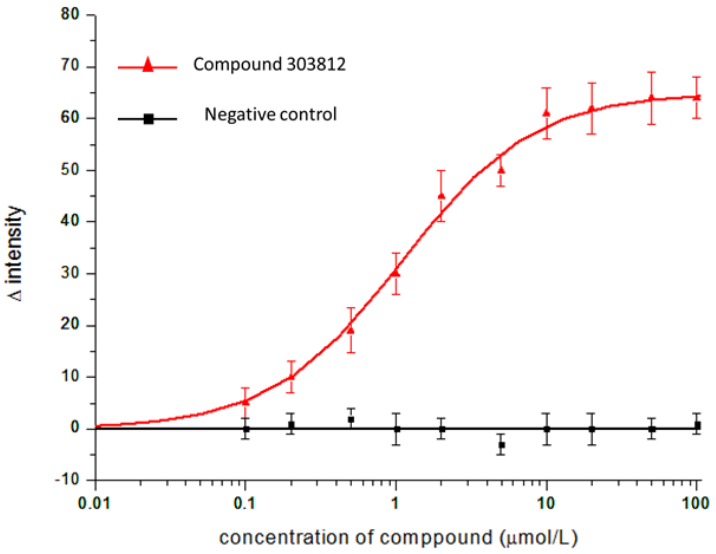
[c] of compounds (μM)	Δ intensity
0	0
0.1	7
0.2	20
0.5	30
1	45
2	51
5	59
10	58
20	65
50	65
100	63



The intensity of tryptophan fluorescence emission was monitored upon titrating compound 683770 into the purified recombinant hemopexin (100 nM) at pH 7.0, RT. The excitation and emission filters were 284/10 nm and 340/30 nm respectively.

Table 11 and Figure 28. The binding of compound 303812 to the hemopexin domain of MMP-9

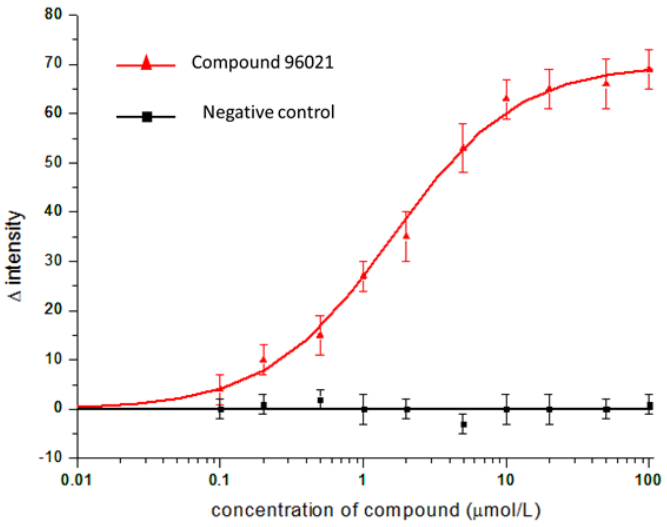
[c] of compounds (μM)	Δ intensity
0	0
0.1	5
0.2	10
0.5	19
1	30
2	45
5	50
10	61
20	62
50	64
100	64



The intensity of tryptophan fluorescence emission was monitored upon titrating compound 303812 into the purified recombinant hemopexin (100 nM) at pH 7.0, RT. The excitation and emission filters were 284/10 nm and 340/30 nm respectively.

Table 12 and Figure 29. The binding of compound 96021 to the hemopexin domain of MMP-9

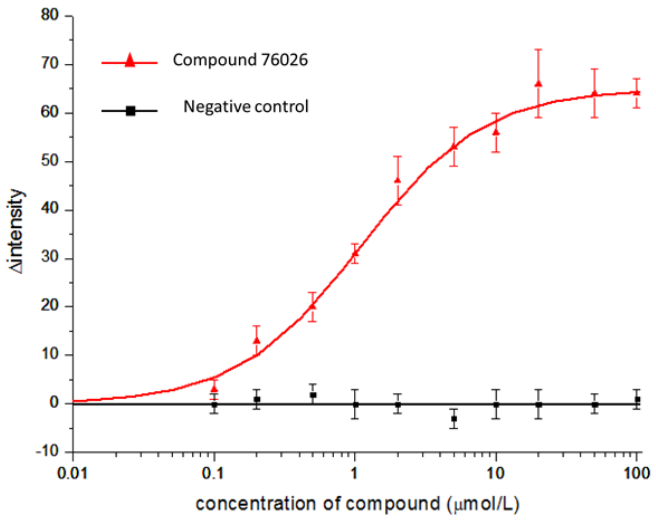
[c] of compounds (μM)	Δ intensity
0	0
0.1	4
0.2	10
0.5	15
1	27
2	35
5	53
10	63
20	65
50	66
100	69



The intensity of tryptophan fluorescence emission was monitored upon titrating compound 96021into the purified recombinant hemopexin (100 nM) at pH 7.0, RT. The excitation and emission filters were 284/10 nm and 340/30 nm respectively.

Table 13 and Figure 30. The binding of compound 76026 to the hemopexin domain of MMP-9

[c] of compounds (μM)	Δ intensity
0	0
0.1	3
0.2	13
0.5	20
1	31
2	46
5	53
10	56
20	66
50	64
100	64

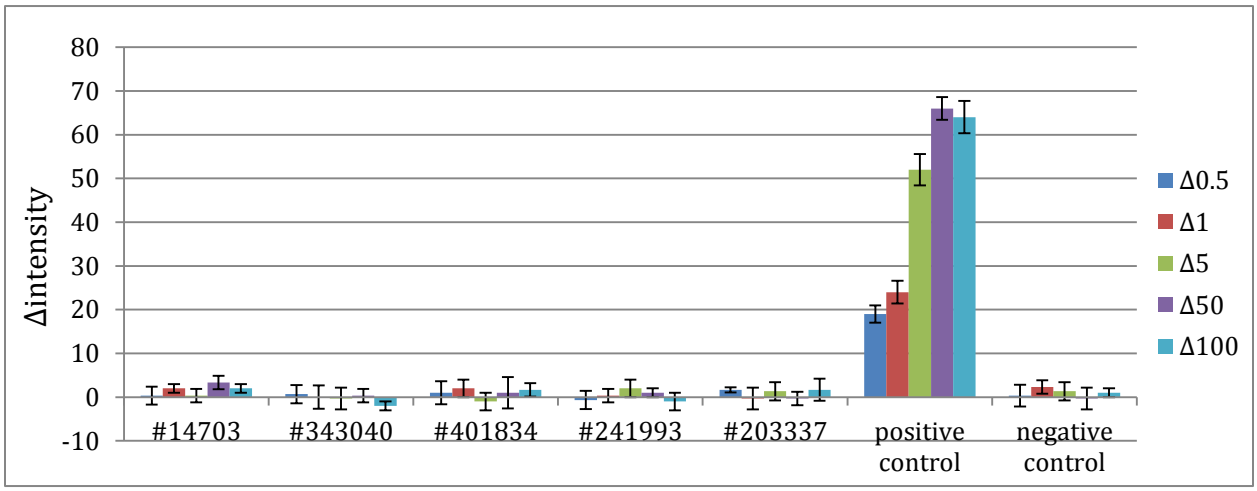


The intensity of tryptophan fluorescence emission was monitored upon titrating compound 76026into the purified recombinant hemopexin (100 nM) at pH 7.0, RT. The excitation and emission filters were 284/10 nm and 340/30 nm respectively.

We also randomly chose 5 compounds which were not computational hits in NCI diversity set and tested them by fluorescence assay. The structures of these compounds are shown in Figure 32.

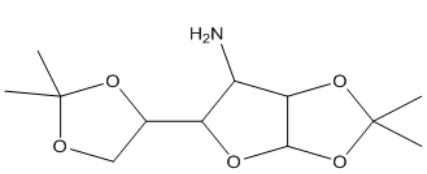
The intensity of tryptophan fluorescence emission was measured by plate reader at compound concentration of 0.5, 1, 5, 50 and 100 μM respectively. The protein concentration was 100 nM. Corresponding amount of compound buffer solution was added into the protein as negative control. The results are shown in Figure 31. None of these compounds results in fluorescence intensity change of the hemopexin domain of MMP-9 as predicted by docking.

Figure 31. Fluorescence assay for inactive NCI compounds

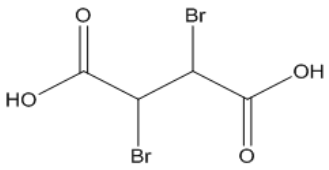


Δ 0.5, 1, 5, 50, 100 represents the fluorescence intensity change of MMP-9 hemopexin domain (100 nM, pH 7.0, RT) in the presence of 0.5, 1, 5, 50, 100 μM compounds. The excitation and emission filters were 284/10 nm and 340/30 nm respectively.

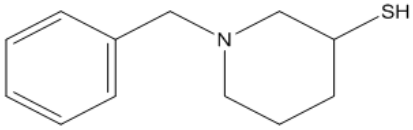
Figure 32. Structures of inactive NCI compounds



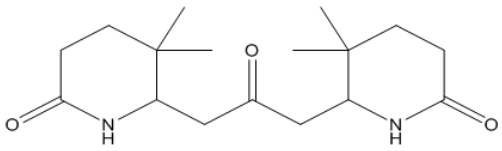
NCI 401834



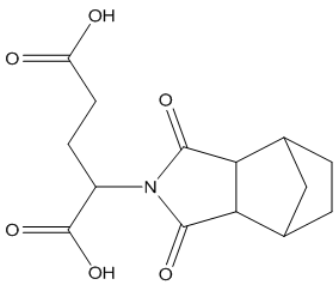
NCI 203337



NCI 351112



NCI 14703



NCI 241993

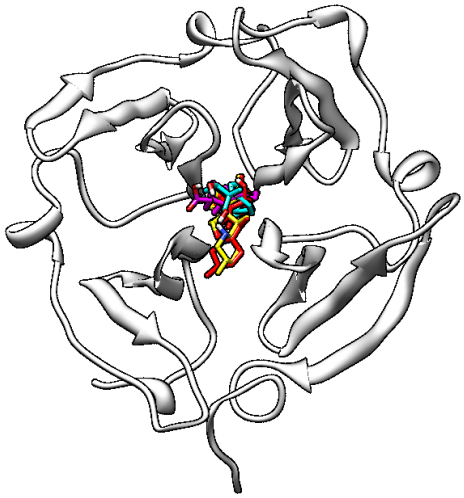
3.2. Discussion

3.2.1. Docking in Dr. Jin Wang's lab and properties of active molecules

Xiliang Zheng in Dr. Jin Wang's lab used program AutoDock Version 4.2 [32] to screen the NCI diversity set against the hemopexin domain of MMP-9. The NCI diversity set contains 1990 small drug-like molecules which were selected from the almost 140,000 available compounds. These compounds in NCI diversity set satisfied two requirements. First, they should be easy to obtain. Second, they should have diverse pharmacophores.

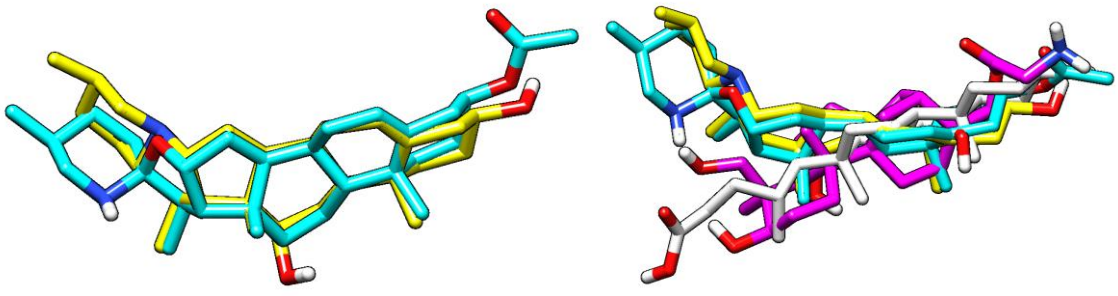
We tested a total of twelve NCI compounds. Seven of these compounds were from the top ranked fifty molecules, and the remaining five compounds were randomly chosen from NCI diversity set. Four compounds from the top ranked fifty molecules show activity. All four molecules were predicted to bind to the large cavity in the center of the top face of the barrel (Figure 33). Molecule 96021 and 76026 share a 5-ring core with various substituents, and their binding poses are highly overlapping. However, the last two molecules adopt different binding poses (Figure 34).

Figure 33. Four active molecules bound to the hemopexin domain of MMP-9



Molecule 683770, 303812, 96021 and 76026 are represented in magenta, cyan, yellow and red respectively.

Figure 34. Four active molecules in their predicted binding orientation.



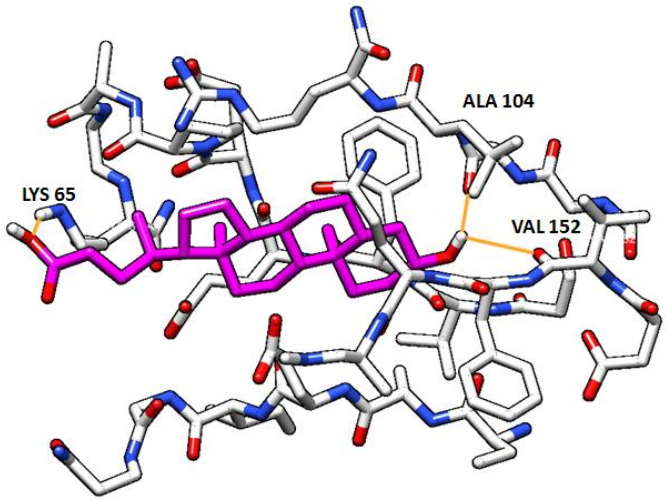
Molecule 683770, 303812, 96021 and 76026 are represented in white, magenta, cyan, and yellow respectively.

3.2.1.1. Molecule 683770

Molecule 683770 forms three hydrogen bonds with the hemopexin domain of MMP-9 (Figure 35). The hydrogen bonds were identified by Chimera software; and the hydrogen bond constraints were relaxed by 0.4 angstroms and 20 degrees. The first hydrogen bond is formed by the oxygen atom on the backbone of ALA 104 and H atom of hydroxyl group on C_9 of molecule 683770. The second hydrogen bond is formed by same H atom and the

oxygen atom on the backbone of VAL 152. The third one is formed by H atom on LYS 65 and O atom on carboxyl group linking to C_1 of molecule 683770.

Figure 35. Interaction between molecule 683770 and the hemopexin domain of MMP-9



The orange bonds represent hydrogen bonds; and the red and blue region represent oxygen and nitrogen respectively. All hydrogen atoms were hidden except the ones which involved in hydrogen bonds formation.

The PubChem BioAssay Database contains bioactivity screens of chemical substances hosted by the US National Institutes of Health (NIH). It provides descriptions of each bioassay, including descriptions of the conditions and readouts specific to that procedure [33]. The compounds in NCI diversity set were mainly tested by three biological assays. The first biological assay is NCI human tumor cell line growth inhibition assay. In this assay, anticancer activity of compound was measured in the presence of 60 different human tumor cell lines. The anticancer activity of compound was evaluated by the value of $GI_{50,}$ concentration of compound required for 50% inhibition of growth. Compounds with log (GI_{50}) (unit M) less than -6 were considered as active. The second biological assay is NCI yeast anticancer drug screen. In NCI yeast anticancer drug screen, the yeast strains with

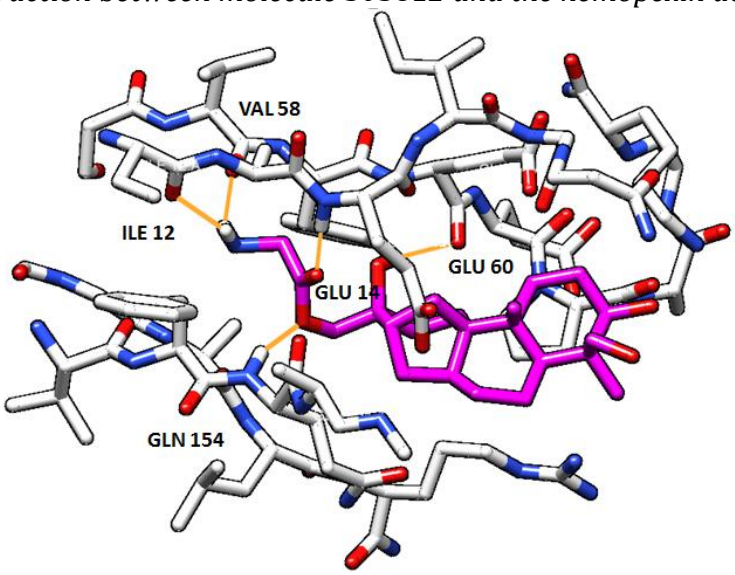
defined genetic mutation were exposed to the compounds. The growth inhibition of yeast strains that was caused by the compounds was measured relative to vehicle treated controls. Compounds with inhibition of growth more than 70% were considered active. The third biological assay is NCI in vivo anticancer drug screen. In this screen, the antitumor activity of compounds was measured in mice bearing transplantable tumors. Survival or tumor size was measured. The results were expressed as the measured made in the treated group (T) divided by the measurement made in the vehicle treated control group (C). Compounds with T/C larger than 150% were considered active.

Molecule 683770 did not show any activity in NCI human tumor cell line growth inhibition assay. Its anticancer activity was measured in the presence of 59 different human tumor cell lines. Molecule 683770 did not show any activity in NCI yeast anticancer drug screen, in which molecule 683770 was tested against 6 yeast strains. The molecule 683770 was not tested by NCI in vivo anticancer drug screen.

3.2.1.2. Molecule 303812

Molecule 303812 forms five hydrogen bonds with the hemopexin domain of MMP-9 (Figure 46). All five hydrogen bonds are formed by the substituent linking to C_1 of molecule 303812.

Figure 36. Interaction between molecule 303812 and the hemopexin domain of MMP-9



The orange bonds represent hydrogen bonds; and the red and blue region represent oxygen and nitrogen respectively. All hydrogen atoms were hidden except the ones which involved in hydrogen bonds formation.

Molecule 303812 showed activity in all biological assays mentioned above. In NCI human tumor cell line growth inhibition assay, it was tested against 72 human tumor cell lines and showed anticancer activity for 12 of these cell lines. Of these 12 human tumor cell lines, NCI-H460 human non-small cell lung tumor cell line [34], 786-0 Renal cell line [35], CAKI-1 Renal cell line [36], ACHN Renal cell line [37] express MMP-9. In NCI yeast anticancer drug screen, molecule 303812 was tested against 6 yeast strains and showed growth inhibition activity for 4 of them. In addition, molecule 303812 showed antitumor activity for 6 types of tumor module in NCI in vivo anticancer drug screen. These tumor modules are tumor model B16 Melanoma (intraperitoneal) in B6C3F1 mice; tumor model HT29, CX-1 Human Adenocarcinoma (MER+) (intrarenal inoculation) in NU/NU Swiss (nude) mice; tumor model Colon Carcinoma 38 (subcutaneous) in B6D2F1 (BDF1) mice; tumor model Mammary Adenocarcinoma CD8F1 (subcutaneous) in CD8F1; tumor model

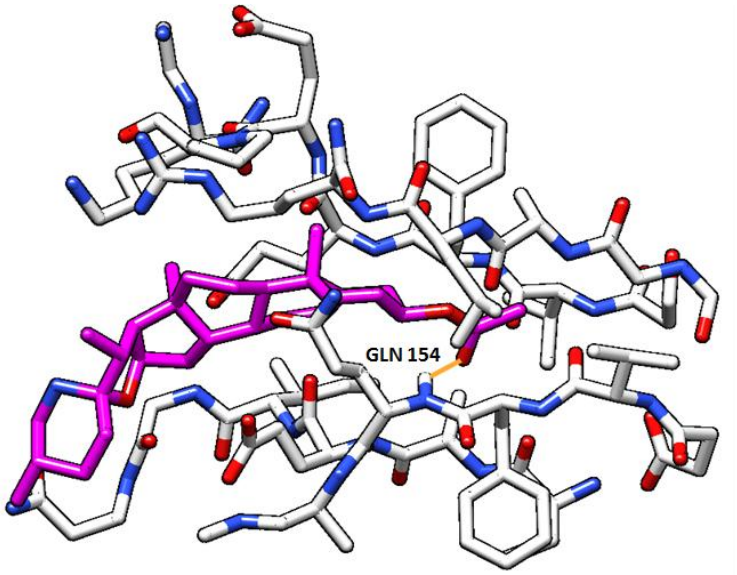
L1210 Leukemia (intraperitoneal) in CD2F1 (CDF1) mice, and tumor model Sarcoma M5076 (intraperitoneal) in B6C3F1 mice.

3.2.1.3. Molecule 96021

Molecule 96021 formed one hydrogen bond with the hemopexin domain of MMP-9 (Figure 35). The hydrogen bond donor is the amino group in the backbone of GLN 154, which interacts with the electron-rich O atom on the ester group of molecule 96021.

The molecule 96021 did not show any activity in biological assays mentioned above.

Figure 37. Interaction between molecule 96021 and the hemopexin domain of MMP-9



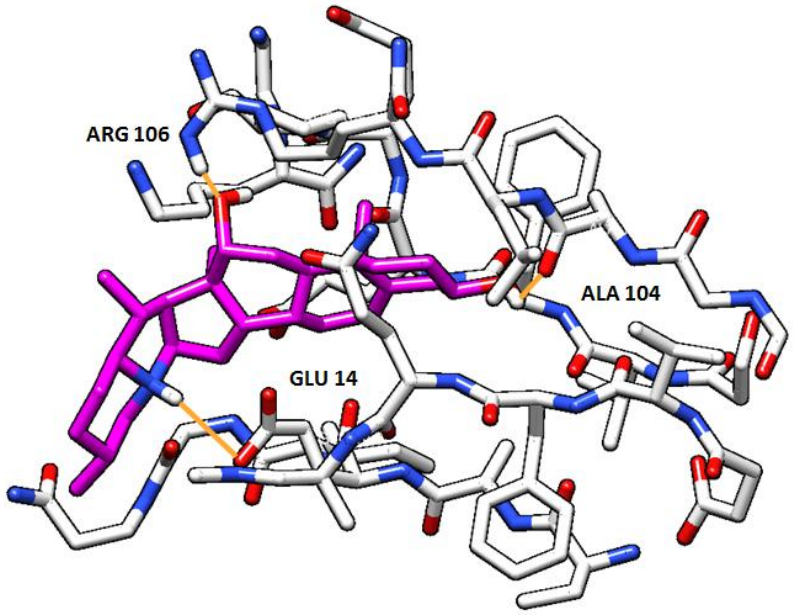
The orange bonds represent hydrogen bonds; and the red and blue region represent oxygen and nitrogen respectively. All hydrogen atoms were hidden except the ones which involved in hydrogen bonds formation.

3.2.1.4. Molecule 76026

Molecule 76026 formed three hydrogen bonds with the hemopexin domain of MMP-9 (Figure 36). Although their 5-ring cores overlaid in the docked conformations, molecule

76026 did not form similar hydrogen bond pattern with the hemopexin domain of MMP-9 as what molecule 96021 did. The molecule 96021 did not show any activity biological assays mentioned above.

Figure 38. Interaction between molecule 76026 and the hemopexin domain of MMP-9



The orange bonds represent hydrogen bonds; and the red and blue region represent oxygen and nitrogen respectively. All hydrogen atoms were hidden except the ones which involved in hydrogen bonds formation.

By comparing the hydrogen bonds formed by four active molecules, we observed that the amino group on the backbone of GLN 154 served as hydrogen bond donor for molecule 96021 and 303812. The electron-rich O atom on the backbone of ALA 104 served as hydrogen bond acceptor for molecule 683770 and 76026.

3.2.2. Binding assay

In binding assay, the hemopexin domain of MMP-9 exists as an equilibrium mixture of dimer and monomer in solution, and the compounds might bind to the dimer or monomer. Therefore, there are two possible types of binding in the assay.

In the first type of binding scheme, the compounds bound to the monomer of the hemopexin domain of MMP-9 (Equation 3). The dissociation constant of dimerization (K_{D1}) is 520 nM, which was calculated by gel filtration analysis. According to the overall K_D's ($K_{Doverall}$) of four compounds measured in binding assay, we calculated the K_{D2}'s of each compound (Table 14).

$$[\ 2PEX\ ] \underset{}{\overset{K_{D1}}{\rightleftharpoons}} 2\ [\ PEX\ ] \underset{K_{D2}}{\overset{+\ 2\ [\ L\ ]}{\rightleftharpoons}} 2\ [\ PEX{\cdot}\ L\ ] \qquad \textit{Equation 3}$$

$$K_{D1} = \frac{[PEX]^2}{[2PEX]} \quad K_{D2} = \frac{[PEX][L]}{[PEX{\cdot}L]} \quad K_{Doverall} = \frac{[L]^2\ [2PEX]}{[PEX{\cdot}L]^2} \quad K_{D2} = \sqrt{K_{Doverall}\ {\cdot}K_{D1}}$$

[2PEX] and [PEX] represent the dimer and monomer of the hemopexin domain of MMP-9 respectively; [PEX·L] represents the ligand-monomer complex.

Table 14. The K_{D2}'s of compounds that bind to the hemopexin domain of MMP-9

compound No.	K_{D2} (μM)
8580066	0.9
683770	0.6
303812	0.8
96021	0.9
76026	0.8

In the second type of binding scheme, the compounds bind to the dimer of the hemopexin domain of MMP-9 (Equation 4). The dissociation constant of dimerization (K_{D1}) is 520 nM. According to the $K'_{D\ overall}$'s of four active compounds measured in binding assay, we calculated the K_{D3} of each compound (Table 15).

$$2\ [\ PEX\ ] \underset{}{\overset{K_{D1}}{\rightleftharpoons}} [\ 2PEX\ ] \underset{K_{D3}}{\overset{+\ 2\ [\ L\ ]}{\rightleftharpoons}} [\ 2PEX{\cdot}2L\ ] \qquad \textit{Equation 4}$$

$$K_{D1} = \frac{[PEX]^2}{[2PEX]} \quad K_{D3} = \frac{[2PEX][L]^2}{[2PEX \cdot 2L]} \quad K'_{Doverall} = \frac{[PEX]^2\,[L]^2}{[2PEX \cdot 2L]} \quad K_{D3} = \frac{K'_{Doverall}}{K_{D1}}$$

[2PEX] and [PEX] represent the dimer and monomer of the hemopexin domain of MMP-9 respectively; [2PEX·2L] represents the ligands-dimer complex.

Table 15. The K_{D3}'s of compounds that bind to the hemopexin domain of MMP-9

compound No.	K_{D3} (μM)
8580066	2.7
683770	1.2
303812	2.1
96021	3.1
76026	2.1

3.3. Conclusion and future work

Our experiments set up a systematic method to screen for small molecules that bind to the hemopexin domain of MMP-9, which combines computational approach and biochemical assay. Its main advantages are low cost and high throughput. The employment of docking makes it possible and convenient to test a big library of small molecules. These small molecules are ranked according to their binding energies for the receptor of interest. This rank list provides a principle basis for choosing potential ligands which have high binding affinity and structural diversity. In addition, testing binding of small molecules to a protein by plate reader is also very efficient and low cost. Monitoring the change of fluorescence by plate reader only requests a small amount of protein and compounds. A lot of samples can be measure at one time.

However, the problem of our method is that only molecules whose fluorescence spectrum does not overlap with the protein's fluorescence spectrum can be tested by our

binding assay. For example, tryptophan in protein is excited at 280 nm, and the emission wavelength is around 330-350. If a compound has similar chromophore as tryptophan, it will also be excited at 280 nm and emit at 330-350 too. Therefore, it is hard to tell the fluorescence quench is caused by binding of compound or by an increased concentration of compound. Because of this drawback, large of amounts of computational hits were not tested.

In order to scale up the screening, it is necessary to find a better method used in binding assay. The new method should be able to monitor the binding of molecules which have similar fluorescence spectrum as the protein, such as fluorescence polarization.

We demonstrated the binding of these 4 molecules to the hemopexin domain of MMP-9. However, in order to be anticancer drug leads, much more biological experiments are required to be done. For example, next determination of whether these compounds can interrupt the biological function of MMP-9 should be made, followed by evaluation of their selectivity for MMP-9 in vivo. Final confirmation that the binding orientation of these compounds is as predicted by docking awaits the determination of liganded crystal structure.

Reference

1. Nagase, H. and J. J. Frederick Woessner, *Matrix metalloproteinases.* J Biol Chem, 1999. **274**(July 30): p. 21491-21494.
2. Kräling, B.M., et al., *The role of matrix metalloproteinase activity in the maturation of human capillary endothelial cells in vitro.* J Cell Sci, 1999. **112**: p. 1599-1609.
3. Lopez-Otin, C. and M. L.M, *Emerging roles of proteases in tumour suppression.* Nat Rev Cancer, 2007. **7**: p. 800-8.
4. Massova, I., *Matrix metalloproteinases: structures, evolution, and diversification.* FASEB J., 1998. **12**.
5. Egeblad, M. and Z. Werb, *New functions for the matrix metalloproteinases in cancer progression.* Nat Rev Cancer, 2002. **2**(3): p. 161-74.
6. Bjorklund, M. and E. Koivunen, *Gelatinase-mediated migration and invasion of cancer cells.* Biochim Biophys Acta, 2005. **1755**(1): p. 37-69.
7. Kessenbrock, K., V. Plaks, and Z. Werb, *Matrix metalloproteinases: regulators of the tumor microenvironment.* Cell, 2010. **141**(1): p. 52-67.
8. Becker, J.W., et al., *Stromelysin-1: three-dimensional structure of the inhibited catalytic domain and of the C-truncated proenzyme.* Protein Sci, 1995. **4**(10): p. 1966-76.
9. Li, J., et al., *Structure of full-length porcine synovial collagenase reveals a C-terminal domain containing a calcium-linked, four-bladed beta-propeller.* Structure, 1995. **3**(6): p. 541-9.
10. Overall, C.M. and O. Kleifeld, *Validating matrix metalloproteinases as drug targets and anti-targets for cancer therapy.* Nat Rev Cancer, 2006. **6**: p. 227-239.

11. Overall, C.M. and C. Lopez-Otin, *Strategies for MMP inhibition in cancer: innovations for the post-trial era.* Nat Rev Cancer, 2002. **2**(9): p. 657-672.

12. Ribatti, D., *Endogenous inhibitors of angiogenesis: a historical review.* Leuk Res, 2009. **33**(5): p. 638-44.

13. Dufour, A., et al., *Role of the hemopexin domain of matrix metalloproteinases in cell migration.* J Cell Physiol, 2008. **217**(3): p. 643-51.

14. Bauvois, B., *New facets of matrix metalloproteinases MMP-2 and MMP-9 as cell surface transducers: Outside-in signaling and relationship to tumor progression.* Biochim Biophys Acta, 2012. **1825**(1): p. 29-36.

15. Philippe, E. and S.V. den, *Biochemistry and molecular biology of gelatinase B or matrox metallopreteinases-9 (MMP-9).* Crit Rev Biochem Mol Biol, 2002. **37**(6): p. 375-536.

16. Turpeenniemi-Hujanen, T., *Gelatinases (MMP-2 and -9) and their natural inhibitors as prognostic indicators in solid cancers.* Biochimie, 2005. **87**(3-4): p. 287-97.

17. Klein, G., et al., *The possible role of matrix metalloproteinase (MMP)-2 and MMP-9 in cancer, e.g. acute leukemia.* Crit Rev Oncol Hematol, 2004. **50**(2): p. 87-100.

18. Hangai, M., et al., *Matrix Metalloproteinase-9-dependent exposure of a cryptic migratory control site in collagen is required before retinal angiogenesis.* Am J Pathol, 2002. **161**(4): p. 1429-1437.

19. Du, R., et al., *HIF1alpha induces the recruitment of bone marrow-derived vascular modulatory cells to regulate tumor angiogenesis and invasion.* Cancer Cell, 2008. **13**(3): p. 206-20.

20. Goldberg, G., et al., *Interaction of 92-kDa type IV collagenase with the tissue inhibitor of metalloproteinases prevents dimerization, complex formation with interstitial collagenase, and activation of the proenzyme with stromelysin.* J Biol Chem, 1992. **267**(3): p. 4583-4591.

21. Hyunju, C., K. Erhard, and H. Robert, *Structural basis of the adaptive molecular recognition by MMP9.* J Mol Biol, 2002. **320**: p. 1065-1079.

22. Lengyel, E., et al., *Expression of latent matrix metalloproteinase 9 (MMP-9) predicts survival in advanced ovarian cancer.* Gynecol Oncol, 2001. **82**(2): p. 291-8.

23. Dufour, A., et al., *Role of matrix metalloproteinase-9 dimers in cell migration: design of inhibitory peptides.* J Biol Chem, 2010. **285**(46): p. 35944-56.

24. Dufour, A., et al., *Small-molecule anticancer compounds selectively target the hemopexin domain of matrix metalloproteinase-9.* Cancer Res, 2011. **71**(14): p. 4977-88.

25. Piccard, H., P.E. Van den Steen, and G. Opdenakker, *Hemopexin domains as multifunctional liganding modules in matrix metalloproteinases and other proteins.* J Leukoc Biol, 2007. **81**(4): p. 870-92.

26. Ewing, T.J.A., et al., *DOCK 4.0: Search strategies for automated molecular docking of flexible molecule databases.* J Comput-Aided Mol Des, 2001. **15**: p. 411-428.

27. Lang, P.T., et al., *DOCK 6: combining techniques to model RNA-small molecule complexes.* RNA, 2009. **15**(6): p. 1219-30.

28. Irwin, J.J. and B.K. Shoichet, *ZINC - a free database of commercially available compounds for virtual screening.* JCIM, 2004. **45**: p. 177-182.

29. Rowsell, S., et al., *Crystal structure of human MMP9 in complex with a reverse hydroxamate inhibitor.* J Mol Biol, 2002. **319**(1): p. 173-81.

30. Natchus, M.G., et al., *Development of new carboxylic acid-based MMP inhibitors derived from functionalized propargylglycines.* J Med Chem, 2001. **44**(7): p. 1060-71.

31. Overall, C.M. and O. Kleifeld, *Towards third generation matrix metalloproteinase inhibitors for cancer therapy.* Br J Cancer, 2006. **94**(7): p. 941-6.

32. Huey, R., et al., *A semiempirical free energy force field with charge-based desolvation.* J Comput Chem, 2007. **28**(6): p. 1145-52.

33. Wang, Y., et al., *An overview of the PubChem BioAssay resource.* Nucleic Acids Res, 2010. **38**(Database issue): p. D255-66.

34. Kozaki, K.-i., et al., *Establishment and characterization of a human lung cancer cell line NCI-H460-LNM35 with consistent lymphogenous metastasis via both subcutaneous and orthotopic propagation.* Cancer Res, 2000. **69**: p. 2535-2540.

35. Waheed, R.M., et al., *Modulation of human renal cell carcinoma 787-0 MMP-2 and MMP-9 activity by inhibitors and inducers in vitro.* JOC, 2006. **23**(2): p. 245-250.

36. Hong, S., et al., *Ascochlorin inhibits matrix metalloproteinase-9 expression by suppressing activator protein-1-mediated gene expression through the ERK1/2 signaling pathway: inhibitory effects of ascochlorin on the invasion of renal carcinoma cells.* J Biol Chem, 2005. **280**(26): p. 25202-9.

37. Gagliano, N., et al., *Malignant phenotype of renal cell carcinoma cells is switched by ukrain administration in vitro.* Anticancer Drugs, 2011. **22**(8): p. 749-62.