T.C.
DOKUZ EYLÜL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

MOLECULAR SPECIFICITY OF TAM FAMILY RECEPTORS TOWARDS DIFFERENT LIGANDS

TÜLAY KARAKULAK

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

İZMİR – 2019

THESIS CODE: DEU.IBG.MSc-2016850034

T.C.
DOKUZ EYLÜL UNIVERSITY
IZMIR INTERNATIONAL BIOMEDICINE AND GENOME INSTITUTE

MOLECULAR SPECIFICITY OF TAM FAMILY RECEPTORS TOWARDS DIFFERENT LIGANDS

DEPARTMENT OF MOLECULAR BIOLOGY AND GENETICS

MASTER OF SCIENCE THESIS

TÜLAY KARAKULAK

Supervising Faculty Member: Asst. Prof. Ezgi KARACA EREK

THESIS CODE: DEU.IBG.MSc-2016850034

YÜKSEK LİSANS TEZ SAVUNMA SINAVI TUTANAĞI

Dokuz Eylül Üniversitesi İzmir Uluslararası Biyotıp ve Genom Enstitüsü Genom Bilimleri ve Moleküler Biyoteknoloji Anabilim Dalı Moleküler Biyoloji ve Genetik Yüksek Lisans Programı 2016850034 numaralı öğrencisi Tülay KARAKULAK **"TAM RESEPTÖR TİROZİN KİNAZ AİLESİNİN FARKLI LİGANDLARA KARŞI SEÇİCİLİĞİ"** konulu Yüksek Lisans tezini 19/04/2019 tarihinde yapılan savunma sınavı sonucunda başarılı olmuştur.

BAŞKAN
Dr. Öğr. Üyesi Ezgi KARACA EREK
Dokuz Eylül Üniversitesi
İzmir Uluslararası Biyotıp ve Genom Enstitüsü

JÜRİ ÜYESİ
Dr. Öğr. Üyesi İsmail Hakkı Akgün
Ege Universitesi
Biyomühendislik

JÜRİ ÜYESİ
Dr. Öğr. Üyesi Zerrin Işık
Dokuz Eylül Üniversitesi
Bilgisayar Mühendisliği

İzmir Uluslararası Biyotıp ve Genom Enstitüsü
Dokuz Eylül Üniversitesi
Sağlık Yerleşkesi No:58/1
Balçova 35340 - İzmir / TÜRKİYE
Tel : (232) 277 63 53
Faks : (232) 412 86 59
e-posta : ibg@deu.edu.tr

TABLE OF CONTENTS…… PAGE NUMBER

INDEX OF TABLES

INDEX OF FIGURES

LIST OF ABBREVIATIONS

Å	Angstrom
AML	Acute Myeloid Leukemia
BSA	Buried Surface Area
EMT	Epithelial to Mesenchymal Transition
EPPIC	Evolutionary Protein-Protein Interface Classifier
FNIII	Fibronectin type III
FATCAT	Flexible structure AlignmenT by Chaining Aligned fragment pairs allowing Twists
Gas6	Growth arrest-specific protein 6
H	Hypothesis
HADDOCK	High Ambiguity Driven biomolecular DOCKing
Ig	Immunoglobulin
$Ig1^{Axl}$	Immunoglobulin-like 2 domain of Axl
$Ig1^{Tyro3}$	Immunoglobulin-like 2 domain of Tyro3
$Ig1^{Mer}$	Immunoglobulin-like 2 domain of Mer
$Ig2^{Axl}$	Immunoglobulin-like 2 domain of Axl
$Ig2^{Tyro3}$	Immunoglobulin-like 2 domain of Tyro3
$Ig2^{Mer}$	Immunoglobulin-like 2 domain of Mer
IQR	Interquartile range
Kd	Dissociation Constant
LG	Laminin G domain
$LG1^{Gas6}$	Laminin G 1domain of Gas6
$LG1^{Pros1}$	Laminin G 1 domain of Pros1
$LG2^{Gas6}$	Laminin G 2domain of Gas6
$LG2^{Pros1}$	Laminin G 2 domain of Pros1
NSCLC	Non-small Cell Lung Cancer
PDB	Protein Data Bank
Pros1	Vitamin K-dependent Protein S
RTK	Receptor Tyrosine Kinase
SDP	Selectivity Determining Positions
vdW	van der Waals

ACKNOWLEDGEMENT

First of all, I am grateful to my thesis advisor, Dr. Ezgi Karaca, for her continued guidance, endless support, patience, her great mentorship as well as shaping my scientific thinking. I will always feel exceedingly lucky to start my academic career in her lab.

I also would like to thank Prof. Mehmet Öztürk and Tuğçe Batur for their valuable discussion and helpful comments made on my thesis.

I also would like to thank TRUBA for making molecular dynamics simulations possible via their super computers.

Many thanks to Nazmiye Arslan, Hamdiye Uzuner and Doğa Eskier for their help in computational biology and programming.

I would like to thank Merve Türker and Tutku Yaraş for their friendship and always being very supportive in every part of my life. Also I would like thank all my friends in Izmir Biomedicine and Genome Center and Izmir International Biomedicine and Genome Institute to make last two years unforgettable.

I would like to thank Erden Eren for his unconditional love and always supporting me to follow my dreams. He always encourage me and always being there for me despite the long distance between us.

Last but not least, I would like to thank the strongest woman in my life, my mother, and two most understanding men in my life, my father and my brother for always supporting me to follow my dreams, their unconditional love and patience. This thesis is dedicated to my family.

MOLECULAR SPECIFICITY OF TAM FAMILY RECEPTORS TOWARDS DIFFERENT LIGANDS

Tülay Karakulak, Izmir International Biomedicine and Genome Institute, Dokuz Eylul University Health Campus, Balcova, 35340, Izmir / Turkey

ABSTRACT

TAM (Tyro3, Axl, Mer) proteins belong to the receptor tyrosine kinase family. One of the prominent TAM activation pathways involves the interaction of TAM's extracellular Ig domains with a ligand. The most commonly studied TAM ligands are two paralogous Vitamin K-dependent proteins, i.e. Gas6 and Pros1. Although Gas6 and Pros1 sequences are significantly similar, Gas6 can bind to all TAM receptors (with the highest affinity towards Axl) while Pros1 can interact only with Tyro3 and Mer (with little or no affinity towards Axl). Expanding on this knowledge, in this thesis, we aim at characterizing the molecular grounds of Axl's ligand selectivity through a structure-based approach. For this, we modeled the complex structures of all possible TAM-ligand interactions, by using the available Axl-Gas6 complex structure as a template. This process was followed by refinement of each complex with HADDOCK. Strikingly, HADDOCK score of each complex, obtained at the end of each refinement session, correlated very well with the observed experimental binding affinities. In order to have a deeper understanding of Axl-ligand selectivity, we later focused on two extremities: the best binder, i.e. Axl-Gas6 and the worst one, i.e. Axl-Pros1. To that end, we run four parallel 200 ns molecular dynamics simulations yielding 800 ns ensembles of Axl-Gas6 and Axl-Pros1. Detailed interaction analyses of these pairs highlighted that electrostatic interactions are dominating Axl's partner selection, which disappear in the case of Axl-Pros1. By combining these results with the conservation and comparative analysis, we identified three specificity determining positions directing Axl's ligand selectivity, i.e. ARG 48^{Axl}, GLU 70^{Axl} and GLU 83^{Axl}.

Keywords: TAM Family, Specificity, Structural Modelling, Molecular Dynamics, HADDOCK

TAM RESEPTÖR TİROZİN KİNAZ AİLESİNİN FARKLI LİGANDLARA KARŞI SEÇİCİLİĞİ

Tülay Karakulak, İzmir Uluslararası Biyotıp ve Genom Enstitüsü, Dokuz Eyül Universitesi, Sağlık Kampüsü, Balçova, 35340, Izmir / Turkey

ÖZET

TAM (Tyro3, Axl, Mer) protein ailesi, reseptör tirozin kinaz ailesinin bir üyesidir. Bu ailenin üyelerinin en önemli aktivasyon mekanizmalarından biri, hücre dışında bulunan domainlerinin ligandlara bağlanması ile yönetilmektedir. TAM ailesinin en çok çalışılan ligandları birbirlerine paralog olan K vitamini bağımlı proteinlerden Gas6 ve Pros1'dir. Gas6 ve Pros1'in sekansları önemli ölçüde benzer olmasına karşın, Gas6 TAM reseptörlerinin tümüne (ve Axl'e karşı çok yüksek afinite ile) bağlanabilirken, Pros1, yalnızca Tyro3 ve Mer ile etkileşime geçebilmektedir. Bu tezde, Axl'in ligand seçiciliğini hesaplamalı yapısal biyoloji tekniklerini kullanarak ve bütün TAM-ligand etkileşimlerini hesaba katarak karakterize etmeyi amaçladık. Bunun için olası tüm TAM-ligand etkileşimlerinin kompleks yapılarını, mevcut Axl-Gas6 kompleksinin üç boyutlu yapısını kalıp olarak kullanarak modelledik. Bu süreci, her bir kompleksin HADDOCK ile yapı iyileştirme prosedürü ile devam ettirdik. Bu işlem sonunda elde edilen her bir kompleksin HADDOCK skorunun, raporlanan deneysel bağlanma afinitileri ile doğrusal ilişki içinde olduğunu ortaya koyduk. Bu bulgunun üzerine, Axl'in seçiciliğinin daha iyi anlaşılması için, iki farklı ligand ile olan etkileşimine odaklandık: en iyi bağlanan Axl-Gas6 ve en kötü bağlanan Axl-Pros1 partnerleri. Axl-Gas6 ve Axl-Pros1 komplekslerinin her biri için toplamda 800 ns olmak üzere dört paralel 200 ns uzunluğunda moleküler dinamik simülasyonluları yürüttük. Bu simülasyonların sonucunda, detaylı etkileşim analizleri, Axl-Gas6 ara yüzündeki elektrostatik kuvvetlerin, Axl-Pros1 ara yüzündekilere oranla oldukça yüksek olduğunu ortaya koyduk. Elde ettiğimiz seçiciliği sağlayan amino asitleri korunmuşluk ve TAM ailesi ile karşılaştırmalı analiz ile birleştirerek, Axl'ın ligandının seçiciliğini yöneten üç pozisyonu belirledik: ARG 48^{Axl}, GLU 70^{Axl} ve GLU 83^{Axl}.

Anahtar Sözcükler: TAM Ailesi, Seçicilik, Yapı Modellemesi, Moleküler Dinamik, HADDOCK

1.INTRODUCTION AND AIM

1.1.Statement and Importance of the Problem

TAM (Tyro3, Axl, Mer) proteins are a member of receptor tyrosine kinase (RTK) family (van der Meer, et al., 2014). They have a critical role in many cell processes including survival, proliferation and cell migration. Thus, its dysregulation causes different types of severe diseases. TAM family is being regulated via different mechanisms including ligand-dependent dimerization, ligand-independent dimerization (heteromeric dimerization of two different TAM receptors and heteromeric dimerization with non-TAM receptors) and trans-cellular binding of TAM extracellular domains. Within these mechanisms, TAM's ligand-dependent dimerization mechanism has been extensively studied (Hafizi and Dahlback, 2006). In this case, TAM family is activated by two paralogous ligands belonging to Vitamin-K dependent family, Gas6 (Growth arrest-specific protein 6) and Pros1 (Vitamin K-dependent protein S). Although Gas6 and Pros1 are significantly similar (40% sequence similarity), under physiological conditions, Gas6 can bind to all TAM receptors with different affinities while Pros1 can interact with Tyro3 and Mer (Hafizi and Dahlback, 2006; Huey, et al., 2016; Yanagihashi, et al., 2017). Among TAM family, Axl's therapeutic potential has earlier been demonstrated, however having only the ligand-induced activation of Axl by Gas6 resolved limits our knowledge in understanding the molecular mechanism of Axl's selectivity. Having this understanding can pave the way for designing new therapeutic molecules, capable of manipulating ligand-induced Axl activation.

1.2.Aim of the Study

Although Axl's therapeutic potential has been demonstrated, the regulatory amino acids directing its ligand selectivity remains elusive. To that end, in this thesis, we aimed at **characterizing Axl's specificity determining residues capable of selectively picking Gas6 over Pros1** by the use of prominent computational structural biology tools.

1.3.Hypothesis of the Study

H1: Missing TAM-ligand complex structures can confidently be modeled, considering that the Axl-Gas6 binding mode is conserved.

H2: Structural investigation of Axl-ligand interactions holds the potential to dissect its specificity differences across Gas6 and Pros1.

H3: Axl's specificity differences can be revealed by investigating the all possible TAM-ligand interactions.

H4: Comparative molecular dynamics simulations of Axl-Gas6 and Axl-Pros1 complexes will help to deduce Axl amino acids capable of mediating its ligand selectivity.

2. GENERAL INFORMATION

2.1. Receptor Tyrosine Kinases

Receptor tyrosine kinases (RTKs) are transmembrane proteins comprising of 58 members classified into 20 families (Lemmon and Schlessinger, 2010). RTKs have diverse roles in essential cellular processes, such as proliferation, survival, growth, motility and adhesion (Du and Lovly, 2018). All RTKs are comprised of an extracellular region followed by a single-pass transmembrane domain and an intracellular tyrosine kinase domain, respectively. Similarity of amino acid sequence in tyrosine kinase domains is used for the classification of RTKs into different families. Figure 1 demonstrates the domain organization of 20 RTK subfamilies of 58 members (Lemmon and Schlessinger, 2010). As depicted in this figure, RTKs can be expressed as a monomer or as an oligomer. For instance, Ins family is found at the cell membrane in the dimer form.

RTKs are generally activated through ligand induced dimerization (Hubbard and Miller, 2007; Lemmon and Schlessinger, 2010). In this case, a receptor-specific ligand binds to the extracellular region of a RTK. Through this binding, another RTK is brought in close vicinity, resulting in a receptor dimerization. This leads to a series of conformational changes impacting the intracellular tyrosine kinase domain, where these domains phosphorylate each other (Du and Lovly, 2018). As a result, downstream signaling proteins are recruited to the phosphorylated tyrosine kinase amino acids such that necessary signaling cascades are initiated. As RTKs have a major role in activation of diverse downstream signaling events, their dysregulation leads to many pathological disorders from developmental syndromes to cancers due to the disruption of cell hemostasis, i.e cell proliferation, apoptosis (Du and Lovly, 2018; McDonell, et al., 2015).

Figure 1: Domain organization of the all 20 receptor tyrosine kinase (RTK) families comprising of 58 members. The legend of 19 different domains found at the RTKs are given at the bottom of the figure. Tyrosine kinase domains are represented in red rectangular box.

2.2. TAM Receptor Tyrosine Kinases

TAM belongs to the RTK family (from the left, 11th family depicted in Figure 1). It has three members: Tyro3 (also called Brt, Dtk, Etk-2, Rek, Rse, Sky, and Tif), Axl (also called Ark, Tyro7, and Ufo) and Mer (also called c-Eyk, Mertk, and Tyro12) (van der Meer, et al., 2014). The conserved sequence, KW(I/L)A(I/L)ES, located within their tyrosine kinase domain discriminates this family from the other 19 RTK subfamilies. TAM receptors are one of the lastly evolved RTKs during evolution (http://pfam.xfam.org/protein/P30530). For instance, TAM could not be found in *Caenorhabditis elegans* and *Drosophila melanogaster*. Instead, they are found in vertebrate and prevertebrate chordates.

TAM proteins have 31-36% identical (52-57% similar) sequence in their extracellular region while their intracellular regions share higher sequence identity, i.e. 54-59% (72-75% similarity). The main isoform of Tyro-3, Axl and Mer are of 97, 98 and 110 kDa, respectively. After posttranslational modifications (such as mono and/or polyubiquitination, and phosphorylation) these weights increase up to 100-140 kDa for Tyro-3 and Axl and 165-205 kDa for Mer (Linger, et al., 2008). TAM proteins can be found in diverse range of cell types, including antigen-presenting cells, monocytes, natural killer cells, osteoclasts, endothelial cells and vascular smooth muscle cells in the vasculature. Though, TAM expression has not been observed in granulocytes or blood lymphocytes (van der Meer, et al., 2014). Unlike the other RTKs, no essential role could be attributed to TAM proteins for embryonic development (Lemke and Rothlin, 2008). Concerning this, it was shown that TAM receptors knockouts (single, double and triple) are still viable, i.e. they had no differences in their phenotypes throughout their first 2-3 weeks after birth (Lemke, 2013) and they could survive more than one year (Lu, et al., 1999). Instead of having role in embryogenesis, TAM receptors are specialized in the homeostasis regulation in the adult tissues. They especially have role in apoptotic clearance via phagocytosis of apoptotic cells. In this mechanism, ligands of TAM family function as a bridge between apoptosis cells and TAM receptors which are expressed on phagocyte (Lemke, 2013). These ligands recognize apoptotic cell's negatively charged flipped phospholipids via their N-terminal, and bind to TAM receptors via their C-terminal region so that apoptosis can be initiated. Thus, absence of TAM family in adult tissues causes diverse phenotypic effects (see Section 2.2.3).

2.2.1. A close up into TAM domain architecture

TAM proteins have two tandem N-terminal immunoglobulin-like domains (Ig1, and Ig2, respectively) followed by two tandem membrane-proximal fibronectin type III-like (FNIII) domains in their extracellular region (Figure 2.A) (Hafizi and Dahlback, 2006; Linger, et al., 2008). After a single transmembrane domain, in their intracellular region, they have a tyrosine kinase domain. In addition to TAM family, Tie (Tie1) and Tek (Tie2) are the only RTKs carry both Ig and FNIII domain in their extracellular domains (from the left, 12th family depicted in Figure 1) (Linger, et al., 2008). Except them, some of the other RTKs can carry one of these domains. For instance, Ig domain is found in the FGF, VEGF and PDGF growth factor receptors

and FNIII domain is found in Ephrin and Insulin families. (Figure 1). Ig domains are generally found in many protein families including antibodies, titin and RTKs. They have reported to play a role in cell-cell recognition, cell-surface receptors interaction and muscle structure (http://www.ebi.ac.uk/interpro/entry/IPR036179). Relatedly, these domains within TAM receptors mediate ligand-dependent TAM dimerization and initiate cell-cell contacts between apoptotic cells and phagocytes (Linger, et al., 2008) (Figure 2.B). The other domain, FNIII, is one of the three fibronectin types (FNI-FNII-FNIII). It is extensively found in animal proteins such as in cell-surface receptors, muscle proteins and extracellular-matrix molecules as well as in enzymes. FNIII repeats has a Arg-Gly-Asp (RGD) sequence motif which has role in interactions with integrin molecule (http://www.ebi.ac.uk/interpro/entry/IPR003961?q=fibronectin). The final domain, tyrosine kinase, is the common domain in all RTKs and function as initiating intracellular signaling cascade by recruiting the molecules to phosphorylated tyrosine residues at this domain (Paul and Mukhopadhyay, 2004).

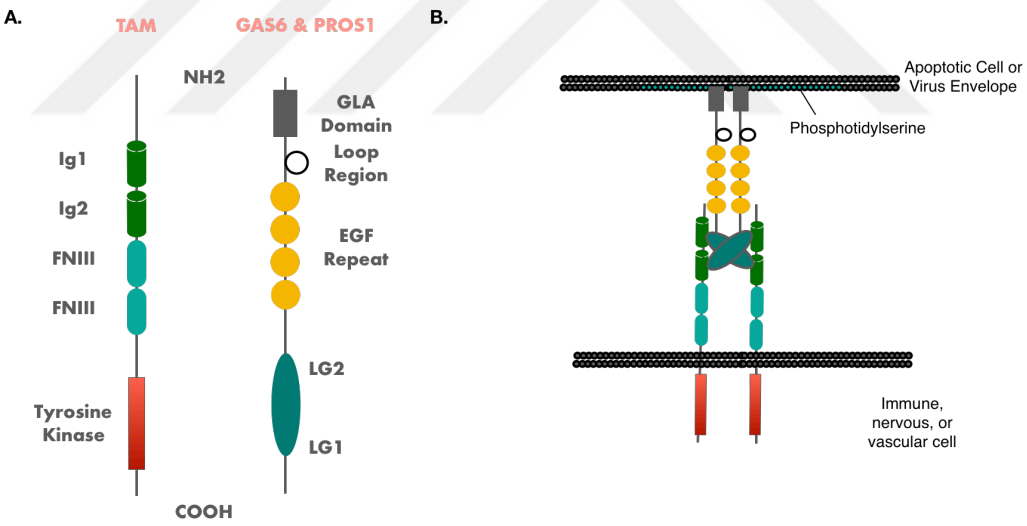


Figure 2: A. Domain organization of TAM family and its ligands, Gas6 and Pros1**.** TAM family consists of Ig1, Ig2, two FNIII and tyrosine kinase domain, respectively. Gas6 and Pros1 consists of Gla domain, Loop region, EGF Repeat, and LG2, LG1 regions, respectively. **B.** Representation of ligand-induced TAM dimerization. LG1:LG2 domain of ligands interacts with the Ig1 domain of receptor and Ig2 domain of the other receptor. Gla domains of ligands unite TAM receptors on phagocyte with the cell membrane of apoptotic cell or virus envelope.

TAM Family receptors are activated through conventional or unconventional mechanisms. The conventional way defines ligand-mediated dimerization as described in section 2.1. Alternatively, TAM receptors are activated by ligand-independent dimerization (Burchert, et al., 1998), heteromeric dimerization among two different TAM members (Bellosta, et al., 1995) and trans-cellular binding of the extracellular domains. As a result of these dimerization processes the alpha helical transmembrane domain is moved in the lateral direction, followed by the induced and conformational rearrangement of the intra-cellular domains. Before ligand binding, receptors are found in inactive form via cis-autoinhibition. Aforementioned conformational changes lead to *trans*-autophosphorylation of tyrosine residues at the tyrosine kinase domains that enables the formation of active conformations. Autophosphorylation of receptors could result in two outcomes: (i) phosphorylation of downstream molecules due to increased catalytic efficiency and (ii) phosphorylated tyrosine residues at the cytoplasmic region form docking sites for the downstream signaling molecules (Linger, et al., 2008).

2.2.2 Ligand-induced activation of TAM

The mostly studied activation mechanism of the TAM family is its ligand-induced activation since TAM-ligand interactions have an important role in many essential cellular processes including the apoptosis processes (Lemke and Rothlin, 2008). Therefore, being able to manipulate TAM-ligand interaction has shown to have a therapeutic potential against cancer (Huey, et al., 2016; Wu, et al., 2018). Thus, understanding the molecular basis of TAM's ligand preference is of great importance.

Reported TAM are Galectin-3, tubby and tubby-like protein 1, Gas6 and Pros1 (Wium, et al., 2018). It is known as Galectin-3 and tubby proteins only interact with Mer while tubby-like protein binds to all three receptors (Caberoy, et al., 2012; Caberoy, et al., 2010). In macrophages and retinal pigment epithelium, these ligands are shown to facilitate phagocytosis (Wium, et al., 2018). Though, interaction mechanisms of Galectin-3, tubby, tubby-like proteins with TAM members have not been characterized yet. The most widely studied ligands of TAM receptors are its two paralogous ligands, i.e. Growth arrest-specific protein 6 (Gas6) and Vitamin K-dependent protein S (Pros1) where Gas6 and Pros1 are both Vitamin-K dependent proteins (Hafizi and Dahlback, 2006). Pros1 is an anticoagulant plasma protein which takes

role as a cofactor for activated protein C in the process of degradation of clotting factors Va and VIIIa. Gas6, on the other hand, have role in cell growth, survival, cell migration and adhesion through its role as a ligand for TAM Family (Studer, et al., 2014). Likewise, TAM RTKs, these proteins are also present only in vertebrates and pre-vertebrates. The weight of Gas6 and Pros1 are ~80 kDa. Moreover, they share 43% amino acid sequence identity (61.3% similarity) with the same domain architecture: Their N-terminal region consists of Vitamin K-dependent carboxylation/gamma-carboxyglutamic (Gla) domain with 11 gamma-carboxyglutamic acid residues and four EGF-like repeats. These regions are followed by C-terminal region containing sex hormone-binding globulin (SHBG)-like structure made of two globular Laminin-G (LG) domains (Hafizi and Dahlback, 2006) (Figure 2.A). *Gla* is the gamma-carboxyglutamate which is formed by post-translational modifications of many glutamate residues by vitamin K-dependent carboxylation and it has high-affinity for calcium ions. These calcium ions are necessary for the folding of Gla properly and binding to apoptotic cells. Gla domain recognizes the "eat-me" signal on cell membranes of apoptotic cells (Figure 2.B). These signals are negatively charged membrane phospholipids (i.e phosphatidylserine) which are normally confined into inner (cytoplasmic) site of plasma membrane in healthy cells. LG domains of Gas6 and Pros1 mediate their binding to TAM members (Lemke, 2013; Linger, et al., 2008).

Although LG domains of Gas6 and Pros1 are significantly similar (58.7% sequence similarity), Gas6 can bind to all TAM receptors with different affinities while Pros1 can interact only with Tyro3 and Mer. Under physiological conditions, Pros1 has little or no affinity to Axl (Hafizi and Dahlback, 2006; Yanagihashi, et al., 2017). A recently published study demonstrated that Pros1 can activate Axl in the glioma cell lines (Sadahiro, et al., 2018). In this study, they have showed that Pros1-Axl signaling cascade has a role in the growth of glioma stem cells. In another study, It has been reported that AXL is over-expressed in some glioblastoma cancers as a result of chromosome 9q13.1 amplification (Margareto, et al., 2009). This can lead to unspecific binding between Axl and Pros1. Also, the binding affinity between Axl and Pros1 was not measured in that study (Sadahiro, et al., 2018).

2.2.3 Dysregulation of TAM Family and Ligands in Cancer and Autoimmune Diseases

TAM receptors and their ligands are involved in platelet aggregation and thrombus formation, endothelial and vascular smooth-muscle homeostasis in normal physiology (Paolino and Penninger, 2016). Moreover, TAM-dependent pathways take part in many essential biological processes, such as spermatogenesis, bone physiology, atherosclerosis, nervous-system, blood-brain barrier's permeability. Concerning these roles in the cell, upregulation of TAM receptors has been associated with many disorders. In recent years, many studies have shown that TAM receptors are involved in activation of pathways associated with cell proliferation, cell survival, epithelial to mesenchymal transition (EMT) and migration. Thus, they act as pro-oncogenes in tumor cells (Paolino and Penninger, 2016). Among TAM receptors, upregulation of Axl has been shown in many cancers such as prostate cancer, ovarian cancer, non-small cell lung cancer (NSCLC), osteosarcoma, acute myeloid leukemia (AML), glioma and thyroid cancer cell lines. Tyro3 has been implicated as upregulated in leukemia, thyroid cancer, metastatic colorectal tumors and melanoma. Lastly, upregulation of Mer has been observed in NSCLC, melanoma and AML. It has been shown that phosphorylated Axl expression has positive correlation with the matrix metalloproteinase (MMP)-9 expression in osteosarcoma patients and breast cancer (Tai, et al., 2008; Wium, et al., 2018). As MMP-9 plays an important role in proteolysis of the extracellular matrix, this correlation is associated with the metastatic role. Additionally, Gas6 has been reported to be overexpressed in melanoma, schwannoma, glioma and pancreatic ductal adenocarcinoma (PDA) cell lines as well as in ovarian cancer and thyroid cancer. Considering the role of TAM Family in apoptosis, it is not surprising that their dysregulation also causes autoimmune diseases as well. In normal conditions, TAM receptors provide phagocytosis of apoptotic cells. When TAM family is dysregulated and apoptotic clearance fails, the own intracellular components of the body accumulate and cause the activation of autoreactive B cells resulting in formation of autoantibodies. The role of TAM family has been demonstrated in autoimmune diseases, such as systemic lupus erythematosus and multiple sclerosis (Rothlin and Lemke, 2010).

2.2.4 Three-dimensional Structures of TAM Family, TAM Ligands and TAM-Ligand Complexes

Among TAM ligands, crystal structure of LG1-LG2 domain of Gas6 was resolved (Protein Data Bank ID: 1H30) (Sasaki, et al., 2002). This structure consists of two V-shaped canonical LG domains. Among TAM members, the crystal structure of Tyro3 has been published (PDB ID: 1RHF) (Heiring, et al., 2004). Regarding TAM-ligand interactions, only Axl-Gas6 structure was elucidated (PDB ID: 2C5D) (Sasaki, et al., 2006). According to this study, two Ig domains of Axl ($Ig1^{Axl}$) are interacting with two Gas6 molecules in a hetero-tetrameric form, without including any receptor-receptor or ligand-ligand interactions (Figure 3). The Ig1 domain of one Axl is bound to the LG1 domain of Gas6 ($LG1^{Gas6}$). This $Ig1^{Axl}$:$LG1^{Gas6}$ 1:1 dimerization is followed by the formation of an Ig1-$Ig2^{Axl}$:LG1-$LG2^{Gas6}$ 2:2 tetramer assembly. Tetrameric Axl-Gas6 crystal structure highlighted two interaction sites (Figure 3.A), i.e. major and minor interfaces. The major interface is established via $Ig1^{Axl}$:$LG1^{Gas6}$ interactions (burying 2366 $Å^2$ solvent accessible surface), while the minor site occurs between $Ig2^{Axl}$:$LG1^{Gas6}$ (burying 765 $Å^2$ solvent accessible surface). Among the two sites, the minor site is highly conserved within other TAM proteins. 2:2 tetramer structure of Axl-Gas6 (Ig1-$Ig2^{Axl}$:LG1-$LG2^{Gas6}$) also pointed out the difference between free and bound form of LG1-$LG2^{Gas6}$ domains: although they are very similar, small reorientation by 9^o happens upon complexation in the interdomain angle. Additionally, two disordered segments (residues between 457-467 and A-B loop, respectively) in free LG1 become ordered (Figure 3.B). When bound form of Ig1-$Ig2^{Axl}$ domains are compared with free form of Ig1-$Ig2^{Tyro3}$, Ig1-$Ig2^{Axl}$ shows linear arrangement while free Ig1-$Ig2^{Tyro3}$ shows bent conformation between Ig1 and Ig2 domains (Sasaki, et al., 2006).

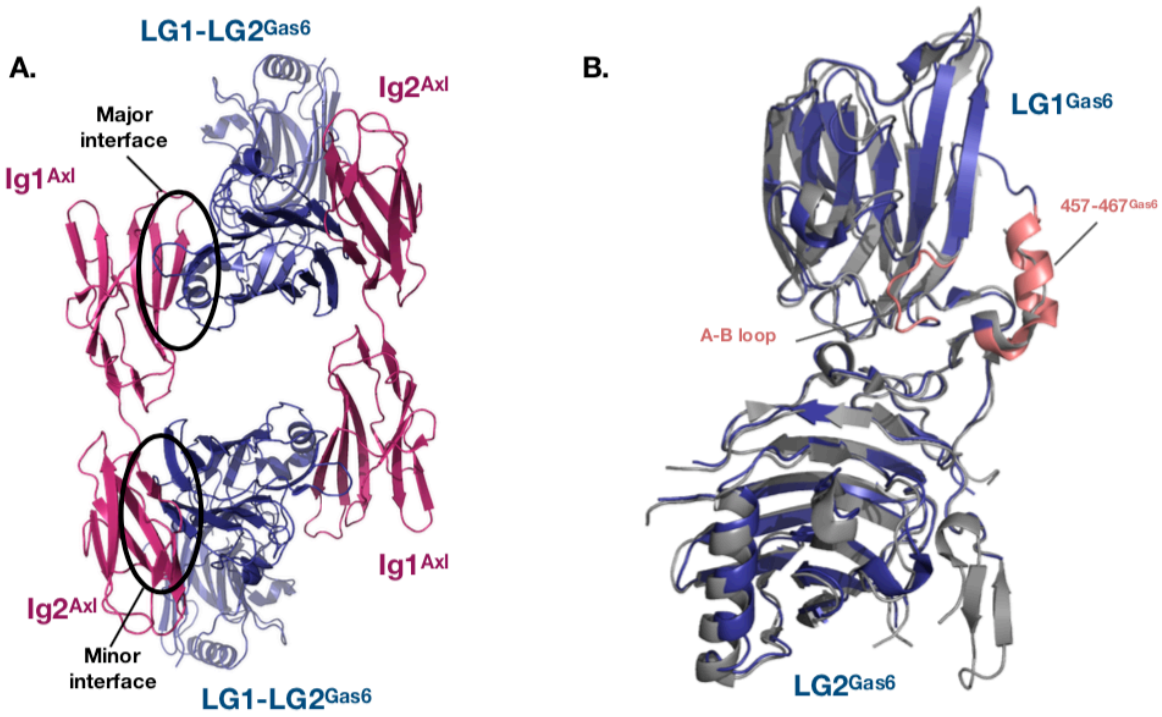


Figure 3: Crystal structure of Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} interaction (PDB ID: 2C5D): 2:2 **A.** Symmetric crystal structure of Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} interaction consists of two contact sites: the major one occurs between Ig1^{Axl} and LG1^{Gas6} (left) and the minor is between Ig2^{Axl} and LG1^{Gas6} (right). **B.** LG1 and LG2 domain of Gas6 in bound (blue) and unbound form (gray). Upon complexation, two disordered regions become ordered. These domains are A-B loop and alpha helix between sequence 457-467. They are depicted as salmon.

The major contact between Ig1^{Axl} and LG1^{Gas6} is formed by side stacking of antiparallel beta sheets interactions (Figure 3.A). Across the major interface, charged and apolar residues are distributed across opposite faces of the stacked beta strands. One face of LG1^{Gas6} is composed of charged residues: ARG308, ARG310 and LYS312. These residues of LG1^{Gas6} are interacting with GLU56 and GLU59 residues on B-C loop of Ig1^{Axl} (Figure 4.A). ASN420 of LG1^{Gas6} carries two N-acetylglucosamine moieties that are packaged against the ARG310 in the crystal structure, however they do not show any interaction with the receptor. On the other hand, the other face of the receptor-ligand interaction is formed by apolar interactions between THR75, THT77, VAL79, ILE90 and VAL92 on Ig1^{Axl} site and ILE307, LEU309, PHE311, THR457, ILE458, THR461 and MET468 on LG1^{Gas6} site (Figure 4.B). The minor contact

(Ig2^{Axl}:LG1^{Gas6}) interaction happens between their G and J strands respectively. Only one salt-bridge formation is seen in this site between LYS204^{Axl} and ASP398^{Gas6}.

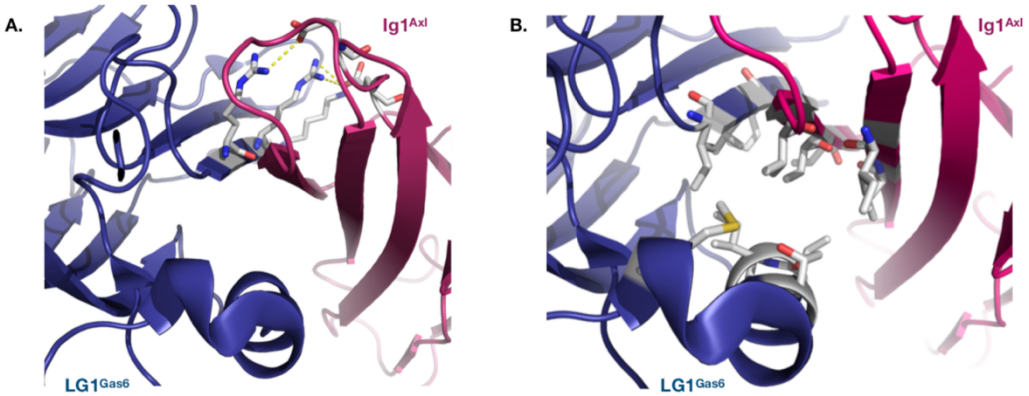


Figure 4: Zoom into major interface between Ig1^{Axl}:LG1^{Gas6} . **A.** One face of reciprocal beta sheets of Ig1^{Axl} and LG1^{Gas6} which is consist of charged residues. **B.** The other face of reciprocal beta sheets of Ig1^{Axl} and LG1^{Gas6} and alpha helix of LG1^{Gas6}. This face is composed of reciprocal nonpolar residues.

Importantly, it has been shown that the predicted glycosylation sites of the Axl are ASN36, ASN150 and ASN191 which are not close to LG domains of Gas6 and no influence on high-affinity Axl-Gas6 binding. When Tyro3 and Mer sequences are mapped onto Axl's structure, it is seen that predicted putative N-linked glycosylation sites of Tyro3 and Mer occur on the solvent-exposed position (Sasaki, et al., 2006). This finding supports that all TAM family have conserved binding mode for the Gas6 and Pros1.

2.2.5 Mutagenesis Experiments on Axl-Gas6

According to mutagenesis experiments, GLU59ARG^{Axl} and THR77ARG^{Axl} double mutations completely abolish the binding of Axl towards Gas6 (Kariolis, et al., 2014). Surprisingly, even though GLU56 makes electrostatic interaction with Gas6, GLU56ARG^{Axl} mutation did not alter the binding affinity in a considerable manner. Moreover ARG310GLU^{Gas6} and LYS312GLU ^{Gas6} point mutations highly reduced affinity toward Axl

(Sasaki, et al., 2006). Additional minor site mutation, ALA404VAL on Gas6 did not affect the affinity in considerable manner however reduced the activation of Axl which indicates the importance of both major and minor interaction for Axl activation (Sasaki, et al., 2006).

2.2.6. Axl and Gas6 targeted therapies

As TAM Family are over-expressed in cancer, many small molecules were developed to inhibit their over-activation. Even though they strongly inhibit the activity of TAM, they are not selective against one of the TAM proteins. For instance, one of the Axl inhibitor, BGB324, inhibits Mer and Tyro3 (Lee-Sherick, et al., 2015). Some studies investigating the impact of targeted-Axl or Gas6 inhibition in tumor cells have showed strong anti-tumor efficiency with suppression of tumor progression. The knockdown of Axl in breast cancer prevents the invasive phenotype. Axl-selective inhibitor, R428, impedes pathways activated by Axl containing Akt phosphorylation and breast cancer cell invasion (Holland, et al., 2010). Moreover, it has been shown that inhibition of Axl hinders angiogenesis by blocking the endothelial tube formation so that resulting in prevention of tumor growth. In addition to Axl-targeted inhibition, Kariolis *et al.* designed two Axl variants that binds Gas6 with high affinity than the wild-type Axl. In this manner, Gas6 was bound and blocked outside of the cell so that Gas6/Axl signaling axis is inhibited (Kariolis, et al., 2017; Kariolis, et al., 2014). These variants confer mutations away from the interface and they function indirectly by changing the conformation of protein.

Even though these therapies have been developed centering Axl and other TAM family members, they are not targeting directly the Axl-ligand interactions. Also, they have never ensured that these therapeutics are specific to Axl only. However, available crystal structure of Axl-Gas6 could pave the way for new targeted therapies for Axl by re-designing the interface of Axl or ligands. For instance, Axl could be modified by computational interface re-design strategies to make it a better binder against Gas6. Alternatively, Pros1 and Axl could be designed to bind each other. Therefore, it is of utmost importance to characterize the specificity determining amino acids of Axl interface which selects Gas6 over Pros1 ligands. All these analyses can assist the interface design of TAM-ligand interactions and developing Axl-based therapeutics in the near future.

2.3. Investigation of Specificity with Computational Methods

Binding patterns between paralogous family are characterized mostly by binding affinity and specificity (Vishwanath, et al., 2017). Binding affinity defines the strength of the interactions between pairs and generally reported by the equilibrium dissociation constant (Kd). In this case, a smaller Kd defines stronger binding (Kastritis and Bonvin, 2013). Binding strength is affected by intermolecular interactions, such as salt bridges, hydrogen bonds or hydrophobic interactions. Salt bridges form between two ionized states of residues by transferring of a proton from a carboxylic acid group of the residue to a primary amine or guanidine group of the other residue. Hydrogen bond (h-bond) is another strong bond formed between hydrogen atom, bound to a h-bond donor, and a h-bond acceptor which has ion pair of electrons for interactions with hydrogen. Contribution of these non-covalent interactions defines the Kd of molecules. The biomolecule-pair with 1 mM can be considered as low affinity but nanomolar (nM) and picomolar (pM) affinities can be considered as high affinity. Understanding the affinity of molecules pave the way for the new design of molecules because it enables to understand also mutations' effect on protein affinity. The computational methods trying the predict binding affinity as *in silico* can based on the calculation of energies of intermolecular interactions by using the three dimensional structure of molecules, or machine learning methods. The SKEMPI v2.0 is an valuable benchmark that contains PDB ID of the interacting partners and changes in the protein-protein binding affinity information upon mutation (Jankauskaite, et al., 2019). Thus, it is used to train and test validity of binding affinity prediction methods (Geng, et al., 2019).

Specificity, on the other hand, defines the selection of a specific partner among others (Kastritis and Bonvin, 2013). Specificity determining positions (SDP) are important residues that directs the specificity of proteins. The highly specific interactions must meet following three criteria: complementarity of ions, complementarity of hydrogen bonds and steric complementarity (Kastritis and Bonvin, 2013). Complementarity of ions defines formation of salt bridges between all charged groups across the interface of biomolecules. If this criterion is not met, charged interface residues tend to make ionic bonds with the solvent, thus destabilizing the interaction. Hydrogen bond interactions across the biomolecular interface also provides

with additional stability to the interaction. Lastly, steric complementarity defines van der Waals forces. Even though they are weak forces, their contribution to binding selectivity is large due to large number of atoms at the interface.

When structures of complexes of paralogous families are available, specificity related information can be derived *in silico*. For this, one could calculate the interface size, buried surface area (BSA) of the complex, electrostatics and van der Waals energies of the interface or characteristics of the interface residues such as their physicochemical properties, conservation and coevolution (Ivanov, et al., 2017). These interactions' site can also be classified into two regions, core and rim. The former consists of the residues which are completely buried in the interfaces, whereas the latter one consists of the residues which are partially exposed to solvent. Importantly, core region is enriched with hydrophobic and highly conserved residues while rim region consists of charged and less conserved residues.

Even with the presence of the above-mentioned parameters, the discrimination of specificity among protein-protein interactions (PPIs) of very similar surfaces is very challenging. To that end, M. Ivanov *et al.* has showed that charge-charge interactions can be used as better discriminants for specificity among paralogous PPIs (Aiello and Caffrey, 2012). It has been reported that SDPs are found at the rim region of proteins. Generally they are highly conserved in a proteins having the same functional specificity. However, they are substituted to different specificity groups in paralogous.

However, for many cases, analyzing a snapshot of a PPI complex won't provide with essential insights into the depiction SDPs. For this, running atomistic molecular dynamics (MD) simulations would be essential to analyze dynamic behavior of biomolecules within a given time course. By use of classical mechanics, MD provides an opportunity to closely investigate folding of molecules, interface changes, residue fluctuations, as well as molecule compactness (Appendix D1). It also generates a platform to trace the change in the interfacial hydrophobic interactions, hydrogen-bonds and salt-bridges with respect to time (Likhachev, et al., 2016). This kind of information can be used to discriminate stable versus nonstable interactions over the time, which can help to pinpoint SDPs (Glazer, et al., 2009).

3. MATERIALS AND METHOD

3.1 Type of Study

In this thesis, we performed a theoretical study. For this, we used web-servers, simulation tools, and statistical analyses.

3.2 Time and Location of Study

This research has been carried out at the Izmir Biomedicine and Genome Center and at the Izmir International Biomedicine and Genome Institute between February 2018 and February 2019.

3.3 Population and Sample of Study

In this thesis, we have investigated the ligand binding preference of a RTK family, specially TAM family.

3.4 Materials of Study

In this thesis, we have used the GROMACS 5.1.4 for molecular dynamics simulations. Our molecular dynamics simulations were carried out in TRUBA (TÜBİTAK's national supercomputer), in its levrekv2-cuda and akya-cuda machines. 24 cores and 1 GPU were used in levrekv2-cuda and 40 cores and 1 GPU were used in akya-cuda.

3.5 Variables of Study

TAM family interactions with its ligands is the dependent variable of this study (as they will all be built by using Axl-Gas6 as a template). All other calculated parameters are the independent variables of the study.

3.6 Tools for Data Collection

3.6.1 Web-servers & resources used for data collection

We used the following tools and simulation packages during the time course of the thesis.

UNIPROT: UNIPROT (https://www.uniprot.org) is a free resource to annotate protein sequences and functions. It was used to acquire the sequence and domain information of TAM Family, Gas6 and Pros1 (UniProt, 2019).

NEEDLE: Needle (https://www.ebi.ac.uk/Tools/psa/emboss_needle/) is a webserver for pairwise sequence alignment of protein (Li, et al., 2015). It was used to predict percentage of similarity between Tyro3 and Axl, Mer and Axl as well as Pros1 and Gas6.

RCSB: RCSB (https://www.rcsb.org/) is the protein data bank. This resource is an archive for 3D coordinates of proteins, nucleic acids and complex assemblies (Burley, et al., 2019). It was used to retrieve the crystal structure of Axl-Gas6.

I-TASSER: I-TASSER (https://zhanglab.ccmb.med.umich.edu/I-TASSER/) program predicts the three dimensional protein structure from its sequence (Zhang, 2008). It was used to model unavailable TAM proteins and its ligands.

FATCAT: Flexible structure alignment by chaining aligned fragment pairs allowing twists (FATCAT) (http://fatcat.sanfordburnham.org) is a webserver for structural alignment of proteins onto a reference structure (Ye and Godzik, 2003). This server was used to model TAM-ligand tetramer interactions.

HADDOCK: We used HADDOCK 2.2 webserver to refine our complex models (https://milou.science.uu.nl/services/HADDOCK2.2/haddock.php) (van Zundert, et al., 2016). HADDOCK 2.2's refinement module calculates intermolecular energies (van der Waals, electrostatics etc.), which were shown to be descriptors of binding affinity for a variety of cases.

PROCHECK: PROCHECK (https://servicesn.mbi.ucla.edu/PROCHECK/) is a webserver to check stereochemical quality of three dimensional structures of proteins (Morris, et al., 1992). We used it to check quality of our refined structures.

GROMACS: This program is a molecular dynamics package that simulates inter-atomic atoms in a time-dependent manner (Van Der Spoel, et al., 2005).

CONSURF: The web server ConSurf (http://consurf.tau.ac.il) calculates evolutionary conservation of amino acid and nucleic acid sequences (Ashkenazy, et al., 2016; Landau, et al., 2005). This information was used to dissect the functionally important regions in proteins.

INTERFACEA: Python package to calculate the intermolecular interactions: Hydrophobic interactions, hydrogen bonds and ionic bonds (https://github.com/JoaoRodrigues/interfacea/). It was used to calculate the inter-molecular interactions in static structures and over the snapshots extracted from MD.

TRUBA: "TÜBİTAK ULAKBİM High Performance and Grid Computing Center" is a national center providing high performance computing and data storage. We used TRUBA to run the molecular dynamics simulations (http://www.truba.gov.tr/index.php/en/main-page/).

PyMOL: PyMOL is a molecular visualization tool (https://pymol.org/2/). This tool was used to visualize the static and dynamic protein structures (Schrödinger, 2010).

Throughout this thesis, statistical and interaction analyses were carried out with the relevant **R packages (R Development Core Team, 2017).**

3.6.2 Methodology

3.6.2.1 Search for the Experimentally Determined Structures of TAM Family, Gas6 and Pros1

The structures of TAM members (Ig1-Ig2^{Receptor}), Gas6 and Pros1 (LG1-LG2^{Ligand})'s were searched at protein data bank (RCSB). Among TAM proteins and its ligands, Axl and Gas6 has a crystal structure in a bound 2:2 tetramer form (PDB ID: 2C5D) (Ig1-Ig2^{Axl}:LG1-LG2^{Gas6}) with 3.3 Å resolution. Gas6 (LG1-LG2^{Gas6}) has also experimentally determined crystal structure in its free form (PDB ID: 1H30) with 2.2 Å resolution . Moreover, Tyro3 (Ig1-Ig2^{Tyro3}) has a crystal structure in the unbound form (PDB ID: 1RHF) with 1.96 Å resolution. Experimentally determined structure of Mer (Ig1-Ig2^{Mer}) and Pros1(LG1-LG2^{Pros1}) have not been determined.

3.6.2.2. Pairwise Sequence Alignment

Sequence alignment is used to identify regions of similarity among proteins or nucleic acids. It is important to reveal functionally or structurally important regions or evolutionary conserved parts in sequences. Sequence similarity among biomolecules is also essential to model three dimensional structure of proteins. The similarity should be high enough (about >30%) to obtain reliable template-based model of proteins. We calculated the pairwise sequence identity and similarity between LG1:LG2^{Gas6} and LG1:LG2^{Pros1} , Ig1:Ig2^{Axl} and Ig1:Ig2^{Tyro3} as well as Ig1:Ig2^{Axl} and Ig1:Ig2^{Mer} by pairwise alignment tool EMBOSS Needle in its default mode (Li, et al., 2015). This tool aims at finding the optimum alignment of two sequences by using the Needleman-Wunsch alignment algorithm (Needleman and Wunsch, 1970) that is global alignment technique. As the sequence lengths are similar in each compared case, global alignment is more suitable for use in comparison.

3.6.2.3 Homology Modelling

Homology Modelling is a computational method to build high quality three dimensional protein structure from its sequence. This technique is based on the idea that protein structures are conserved more than their sequences in evolutionary processes. We used iterative implementation of the Threading ASSEmbly Refinement (I-TASSER) webserver (Zhang,

2008) to model Tyro3, Mer and Pros1. The reason is that the best 3D structure of protein models were modelled by I-TASSER webserver at CASP7, CASP8, CASP9, CASP10, CASP11, CASP12 and CASP13 which are community-wide blind experiment. The quality of the models in I-TASSER is encoded in Confidence score (C-Score) and TM-Score. C-Score ranges between -5 and 2 where the higher C-Score, the more confident the model is (Zhang, 2008). TM-score measures the global accuracy of the structures where a TM-Score greater than 0.5 indicates a correct topology. It was shown that I-TASSER generated 92 models with TM-score > 0.5 among 110 targets that makes it reliable to use for modelling the structures. LG1:LG2^{Gas6} structure (2C5D: chain A) was used as a template to model LG1:LG2^{Pros1} while Ig1:Ig2^{Axl} structure (2C5D: chain C) was given as a template to model Ig1:Ig2^{Tyro3} and Ig1:Ig2^{Mer} proteins. Even though crystal structure of free form of Ig1:Ig2^{Tyro3} is known, we modelled it due to obtain its bound form. The structure qualities were evaluated according to their TM-Score and C-Score. The ones with highest quality were selected for further steps. The C-Score of the highest quality models of Tyro3, Mer and Pros1 were given as 0.95, 0.10 and 1.27, respectively. TM-Score of the highest quality models of Tyro3, Mer and Pros1 were given as 0.84+-0.08, 0.73+-0.11, 0.89+-0.07, respectively.

3.6.2.4 Template-Based Modelling

FATCAT (Ye and Godzik, 2003) was used for template-based modelling, where we used Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} binding mode as a template. Ig1-Ig2^{Mer} and Ig1-Ig2^{Tyro3} receptors were superimposed onto the Ig1-Ig2^{Axl}. Likewise, LG1-LG2^{Pros1} was superimposed onto LG1-LG2^{Gas6} via FATCAT. Then each TAM-ligand tetramer structure was isolated in PyMOL and they were saved as separate pdb files. At the end, five receptor-ligand pair models were obtained: Ig1-Ig2^{Tyro3}: LG1-LG2^{Gas6}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, Ig1-Ig2^{Axl}:LG1-LG2^{Pros1}, Ig1-Ig2^{Tyro3}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}.

3.6.2.5 HADDOCK Refinement

3.6.2.5.1 The General HADDOCK Protocol

HADDOCK is a most cited molecular docking program that predicts the three dimensional structure of biomolecular complexes (van Zundert, et al., 2016). HADDOCK performs docking with following three stages:

1. *it0*, Randomization of orientations and rigid body minimization: The interactions (e.g bonds lengths, bond angles and dihedral angles) between partners that will be docked are frozen and treated as rigid bodied. Then, partners are randomly rotated around their center of mass. Rigid body minimization is applied after this step in order to optimize the interactions.
2. *it1*: Semi-flexible simulated annealing in torsion angle space: In this stage, the interfaces modeled at the end of *it0* are refined in three-steps: (i) the orientations of partners are optimized . (ii) interfacial side chain (residues within 5 Å) cut-off) conformations are optimized. (iii) interfacial backbone and side-chain conformations are optimized.
3. *water*: Refinement with explicit solvent: In this final stage of HADDOCK, interacting partners are submerged in a solvent shell (either TIP3P model water or DMSO environment). In this solvent environment, very short molecular dynamics simulation (in femtosecond timescale) at 300 K is applied to complex to improve the energetics of interaction.

In our study, we used only the refinement module (step 3) of HADDOCK 2.2 webserver to calculate energetics of generated five complexes obtained in previous step (Ig1-Ig2^{Tyro3}: LG1-LG2^{Gas6}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, Ig1-Ig2^{Axl}:LG1-LG2^{Pros1}, Ig1-Ig2^{Tyro3}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}) and of the template crystal structure (Ig1-Ig2^{Axl}: LG1-LG2^{Gas6}). HADDOCK uses the following Equation 1 scoring function to rank the poses at the end of water refinement:

HADDOCK-water = 1.0 van der Waals + 0.2(electrostatics) + 1.0 desolvation (1)*

3.6.2.5.2 Applied Water Refinement Protocol

In order to refine TAM-ligand tetramers, the multi-body interface of HADDOCK 2.2 was used with the following parameters:

- In the "Distance Restraints" section:

1. Define center of mass restraints to enforce contact between the molecules → True

- In the "Sampling Parameters" section:
- Number of structures for rigid body docking (it0) → 200
- In the "Advanced Sampling Parameters" section:

1. Randomize starting orientations → False
2. Perform initial rigid body minimisation → False
3. Allow translation in rigid body minimisation → False
4. Number of MD steps for rigid body high temperature TAD → 0
5. Number of MD steps during first rigid body cooling stage → 0
6. Number of MD steps during second cooling stage with flexible side-chains at interface →0
7. Number of MD steps during third cooling stage with fully flexible interface → 0

3.6.2.5.3 HADDOCK Analysis

HADDOCK runs of each complexes were downloaded. HADDOCK reports vdW energy, electrostatics energy, desolvation, total HADDOCK Score, as well as BSA of each solution. HADDOCK also calculates residue-based energy scores of each complex (expressed in electrostatics, van der Waals (vdW) and electrostatics+vdW) in *ene-residue.disp* file (can be found under the HADDOCK output folder: structures/it1/water/analysis).

In order to generate HADDOCK-related plots, "lattice" library and levelplot function of R were used (Sarkar, 2008). Moreover electrostatics, vdW energies and BSA were obtained via calculating the average of the four best ranked structures for $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$.

3.6.2.7 Quality Control of Tetramer Models

Stereochemical quality of tetramers were checked with the PROCHECK webserver after HADDOCK refinement. PROCHECK generates the Ramachandran Plot. Ramachandran plot is diagram to visualize the quality of protein structures according to torsional angles psi (Φ) and phi (ψ) between amino acids: Φ is the angle between N and Cα, ψ is the angle between Cα and C atoms (Ho and Brasseur, 2005). In the Ramachandran plot, Φ values of residues are plotted on x-axis while ψ values on y-axis. Combination of these angles defines which combination of these angles are possible (allowed regions) in the protein structure and which combinations are not possible due to steric hindrance among residues (disallowed regions). We generated Ramachandran plots of crystal structure of Ig1:Ig2^{Axl}:LG1:LG2^{Gas6} (2C5D), refined version of crystal structure of Ig1:Ig2^{Axl}:LG1:LG2^{Gas6}, and the other five refined models in previous steps.

3.6.2.8 Electrostatics Potential Energy Maps

Electrostatic potential energy maps are used to visualize distribution of charges on a biomolecular surface. The Adaptive Poisson-Boltzmann Solver (APBS), is a method to calculate electrostatic potentials of proteins (Jurrus, et al., 2018). It has an APBS Tools 2.1 PyMOL plugin that makes possible to calculate electrostatics within PyMOL and visualize protein's electrostatics in color-coded surface. APBS Tool requires PQR file for calculation. PQR file differs from the pdb file with 2 columns: the occupancy and B-factor columns in the pdb file are replaced by charge and radius of atoms. PQR file of each protein were generated via PDB2PQR webserver (Dolinsky, et al., 2004). With this tool, major interfaces of Ig1:Ig2^{Axl}, LG1:LG2^{Gas6} and LG1:LG2^{Pros1} were investigated.

3.6.2.9 Interface Analysis

The interface of Ig1^{Axl}:LG1^{Gas6} and Ig1^{Axl}:LG1^{Pros1} were investigated in two ways: (i) analyzing the interaction types, (ii) analyzing residue-residue pairwise energies at the dimer interface.

3.6.2.9.1 Interaction Type Analysis

Interfacea is the Python library (https://github.com/JoaoRodrigues/interfacea) to calculate atomic interactions at interface of biomolecules: hydrophobic interactions, hydrogen bonds and ionic bonds. The interactions are calculated according to following criteria:

- If any pair of non-polar atoms in the reciprocal groups are closer than 4.4 Å, there is a hydrophobic interaction between them.
- If hydrogen donor and hydrogen acceptor are closer than 2.5 Å, there is a hydrogen bond between them (N/O/F/S atoms bound to hydrogen are defined as donors. N/O/F/S atoms within the distance of 2.5 Å of that hydrogen are defined as acceptors).
- If the oppositely charged groups (nitrogen or oxygen atoms) are closer than 4.0 Å, there is an ionic bond between them.

3.6.2.9.2 Pairwise Interaction Energy Analysis

Score_Residue_pairs.py is a python script to calculate the interaction energies of interfacial residues across a PPI (Joao Rodrigues, personal communication). This script first minimize the potential energy of the system. Then it labels the interface residues, and uses AMBER force field to calculate the inter-residue energies. It calculates vdW and electrostatics energy according to Leonard-Jones potential and Coulomb's Law, respectively.

3.6.2.8 Molecular Dynamics

GROMACS 5.1.4 software was used for molecular dynamics simulations and for the quality analysis of simulations (e.g temperature, RMSD, Rg analysis) (Van Der Spoel, et al., 2005). All simulations were carried out at TRUBA (TÜBİTAK ULAKBİM High Performance and Grid Computing Center) by using levrek-cuda and akya-cuda machines. The slurm files that were used to run simulations can be found in Appendix A.

3.6.2.8.1 Molecular Dynamics

Molecular dynamics (MD) is the computer simulation method that enables the movement of atoms and molecules in a fixed period of time so that dynamics evolution of system can be interpreted. In MD simulation, trajectory of atoms and molecules are produced. This trajectory are determined by the classical laws of motion, the Newton's equation. It defines the positions and velocities of the molecules in the system. The formula (Eq 2) describes that force F_{xi} is applied to a particle having mass *m* so that it motions throughout the x_i coordinate resulting in the formation of trajectory.

$$\frac{\delta^2 x_i}{\delta t^2} = \frac{F_{x_i}}{m_i} \qquad (2)$$

The forces applied to particles are calculated via molecular mechanics force fields that is the set of parameters which approximate the potential energy of the system. The force fields can be classified into two categories: non-bonded and bonded terms. The formula (Eq 3) describes these bonded terms: bond stretching, angle bending and torsion angles, and non-bonded terms: van der Waals and electrostatics.

$$E_{total} = \overbrace{\sum_{bonds} K_r(r - r_{eq})^2 + \sum_{angles} K_\theta(\theta - \theta_{eq})^2 + \sum_{dihedrals} \frac{V_n}{2}[1 + \cos(n\phi - \gamma)]}^{\text{Bonded}} + \overbrace{\sum_{i<j} \left[\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^{6}} + \frac{q_i q_j}{\epsilon R_{ij}} \right]}^{\text{Non-bonded}} \qquad (3)$$

For the MD simulations, AMBER99SB-ILDN force field (Lindorff-Larsen, et al., 2010) and TIP3P water model (Jorgensen, et al., 1983) were used. They were placed in rhombic dodecahedron unit cell. Periodic Boundary Condition (PBC) was set to 1.4 nm. The mdp files are generated following https://github.com/haddocking/molmod-data. Accordingly, each complex was first minimized in a vacuum by using steepest descent algorithm. After, they were solvated in the TIP3P water model, ions were added to the simulation box (51 NA+ and 48 CL- ions were added to neutralize $Ig1^{Axl}$:$LG1^{Gas6}$ systems while 58 NA+ and 49 CL- ions were added to neutralize $Ig1^{Axl}$:$LG1^{Pros1}$ systems). All topology files were edited manually according

to add the number of NA+ and CL- ions. Second energy minimization was performed on the solvated systems. Solvent and hydrogen atoms were relaxed by a 20 ps molecular dynamics under constant temperature (NVT) ensemble conditions (random seed value at NVT step was changed in each parallel simulation). This step was followed by 20 ps NPT (isothermal-isobaric) equilibrations at 300 K. As a last step before production run, the position restraints were released progressively by decreasing strength of the force constant. After system is equilibrated at desired temperature and pressure, we run 200 ns molecular dynamics simulation for each parallel simulation.

3.6.3.8.2 Analysis of Molecular Dynamics Simulations

In order to analyze the simulations PBC effect was corrected and the system was stripped off the solvent and ion atom. After isolating only complex coordinates at each simulations, Root Mean Square Deviations (RMSD) , Radius of gyrations (Rg) and Root mean square Fluctuation (RMSF) of proteins were calculated by GROMACS tools. Amount of hydrophobic interactions, hydrogen bond and ionic bonds were calculated with the interfacea python library. Energetics of all pairwise interactions were calculated via Score_Residue_pairs.py python code. The order of analysis is given below:

- After 200 ns simulations were completed, backbone RMSD values of the each simulations were calculated with respect to initial states with *rms* function of GROMACS. RMSD is calculated with the equation 4: where $r_i(t)$ is the position of atom *i* at time *t* and m_i is equal to mass of atom i.

$$\mathrm{RMSD}(t) = \left[\frac{1}{M} \sum_{i=1}^{N} m_i |\mathbf{r}_i(t) - \mathbf{r}_i^{\mathrm{ref}}|^2 \right]^{1/2} \quad (4)$$

  The average RMSD values were calculated with the *MDPlot* package of R (Margreitter and Oostenbrink, 2017). The RMSD values were loaded with *load_rmsd* function and average values were calculated with *rmsd_average* function.

- Backbone Rg of each trajectory (25-200 ns time interval for $Ig1^{Axl}$:$LG1^{Gas6}$ and 50-200 ns time interval for $Ig1^{Axl}$:$LG1^{Pros1}$) was calculated with respect to initial structure with the *gyrate* function of GROMACS. Rg is calculated with the equation 5: where where $r_i(t)$ is the position of atom *i* with respect to the center of mass at time *t* and m_i is equal to mass of atom i.

$$R_g = \left(\frac{\sum_i |\mathbf{r}_i|^2 m_i}{\sum_i m_i} \right)^2 \qquad (5)$$

The plots were created with the R *plot* function.

- Residue-based RMSF of each trajectories (25-200 ns time interval for $Ig1^{Axl}$:$LG1^{Gas6}$ and 50-200 ns time interval for $Ig1^{Axl}$:$LG1^{Pros1}$) was calculated with respect to initial structures with the *rmsf* function of GROMACS. RMSF is calculated with the equation 6: where T is the overall time course, $r_i(t)$ is the position of atom *i* at time j, with respect to reference position r_i^{ref}.

$$\mathrm{RMSF}_i = \left[\frac{1}{T} \sum_{t_j=1}^{T} |\mathbf{r}_i(t_j) - \mathbf{r}_i^{\mathrm{ref}}|^2 \right]^{1/2} \qquad (6)$$

The distribution of RMSFs of each trajectory were plotted with the box and whisker analysis. The representative structures were extracted from one of the $Ig1^{Axl}$:$LG1^{Gas6}$ simulations and one of the $Ig1^{Axl}$:$LG1^{Pros1}$ simulations by *rmsf* function with combination with *-oq* parameters. The structures were visualized with PyMOL (The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC).

- After deciding the equilibration time for each simulation, the snapshots from the simulations were extracted as following:
    - The 350 frames from each $Ig1^{Axl}$:$LG1^{Gas6}$ simulations were extracted in each 500 ps between 25-200 ns trajectory.
    - The 300 frames from each $Ig1^{Axl}$:$LG1^{Pros1}$ simulations were extracted in each 500 ps between 50-200 ns trajectory.

- Hydrophobic interactions, hydrogen bounds and ionic bonds were calculated for each snapshot for each parallel simulation: Interfacea python library (described in section 3.6.2.9.1) was used to calculate those kind of interactions. Distribution of each kind of interaction was calculated with the box and whisker analysis of plotly R package (*boxplot* function) where the whisker length of 1.5IQRs.
- The salt-bridges that are seen across $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ interfaces were kept if they were seen at least in 25% within the time frame of a given trajectory. These residues were named after consistent salt-bridges.
- Pairwise energetics of whole residue-residue interactions across 350 snapshots and four parallel simulations of $Ig1^{Axl}$:$LG1^{Gas6}$ (total 1400 snapshots) were calculated via Score_Residue_pairs.py python code.
- Distribution of only energetically favorable interactions (vdW+electrostatic < 0 energies) of all four parallel $Ig1^{Axl}$:$LG1^{Gas6}$ simulations were plotted after transformed into logarithmic scale.
- The average pairwise energies of consistent salt-bridges across 350 snapshots from Axl-Gas6-Parallel1 were marked on the plot.

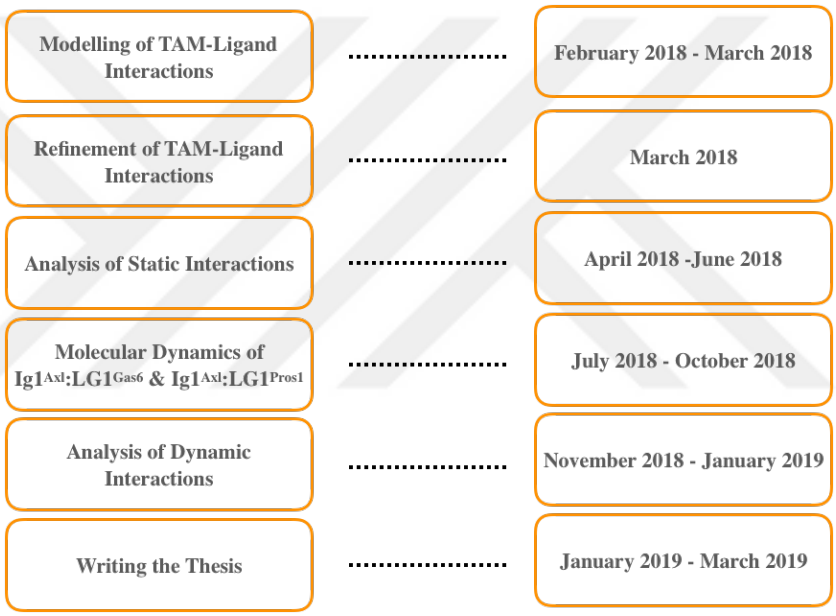
3.6.2.9 Deciding for Selectivity Determining Positions

- The same $Ig1^{Axl}$ residue making consistent salt-bridges in all complexes was eliminated to be SDPs. The rest (uncommon) $Ig1^{Axl}$ residues which are making salt-bridges with $LG1^{Gas6}$ and not with $LG1^{Pros1}$ were classified as possible SDPs. Then, positions in $Ig1^{Tyro3}$, $Ig1^{Mer}$ and $LG1^{Gas6}$, $LG1^{Pros1}$ corresponding to these possible SDPs were compared structurally and at the sequence level (by PyMOL and Pairwise Alignment

respectively). At the side, evolutionary conservation scores of these residues were obtained ConSurf web-server.

- As a final step, EPPIC webserver was to classify the positioning of possible SDPs: whether they are at the rim or at the core the PPI. This analysis was performed by using the crystal structure of Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} (2C5D). Core residues are the ones that are buried at least 95% in protein structure. The rest of interface residues are counted as rim residues.
-

3.7. Study Plan and Calendar



3.8. Data Evaluation

We have performed the following quality controls:

- Crystal structure of Axl-Gas6 was evaluated by its resolution (in Å)
- Homology models of LG1:LG2^{Pros1}, Ig1:Ig2^{Tyro3} and Ig1:Ig2^{Mer} were evaluated with the TM-Score and C-Score of I-TASSER webserver.
- Complex structures were evaluated with Ramachandran plots provided by the PROCHECK webserver.
- Molecular Dynamics Simulations were evaluated by RMSD (root mean square deviation) from the initial structure and convergence of temperature analysis.

3.9. Limitations of Study

As running molecular dynamics is a time-consuming study, we carried out molecular dynamics of selected complexes (maximum two complexes with four parallel from each).

3.9. Ethics Committee Approval

No ethics committee approval was required or requested due to the purely computational scope of the study.

4. RESULTS

4.1 The Structural Modelling of TAM Family and its Ligands

As discussed in the Introduction 2.2.3, Axl-Gas6 interaction is mediated through Ig1 and Ig2 domains of Axl (Ig1-Ig2^{Axl}) and LG1 and LG2 domains of Gas6 (LG1-LG2^{Gas6}) (PDB id: 2C5D, (Sasaki, et al., 2006). By taking this information as a basis, we aimed at modeling the TAM-ligand interaction pairs that has not been resolved, i.e. Tyro-Gas6, Mer-Gas6, Axl-Pros, Tyro-Pros, Mer-Pros. For this, we initially focused on modeling of Ig1-Ig2 of Tyro (Ig1-Ig2^{Tyro}) and Mer (Ig1-Ig2^{Mer}) and LG1-LG2 of Pros1 (LG1-LG2^{Pros1}) structures in their bound states, by taking Ig1-Ig2^{Axl} and LG1-LG2^{Gas6} as a template, respectively.

According to the pairwise sequence alignment EMBOSS Needle web tool (https://www.ebi.ac.uk/Tools/psa/emboss_needle/ , see Methods 3.6.2.2): (i) LG1-LG2^{Gas6} and LG1-LG2^{Pros1} share 39.5% sequence identity and 58.7% sequence similarity (Figure 5), (ii) Ig1-Ig2^{Axl} and Ig1-Ig2^{Tyro} share 37.8% sequence identity and 57.7% sequence similarity (Figure 7), and Ig1-Ig2^{Axl} and Ig1-Ig2^{Mer} share 34.7% sequence identity with 51.0% sequence similarity (Figure 6). Based on the >30% criteria described in section 3.6.2.2, Ig1-Ig2^{Axl} can be used as a template to homology model the bound states of Ig1-Ig2^{Tyro} and Ig1-Ig2^{Mer}, where LG1-LG2^{Gas6} can be used as a template to model LG1-LG2^{Pros1}.

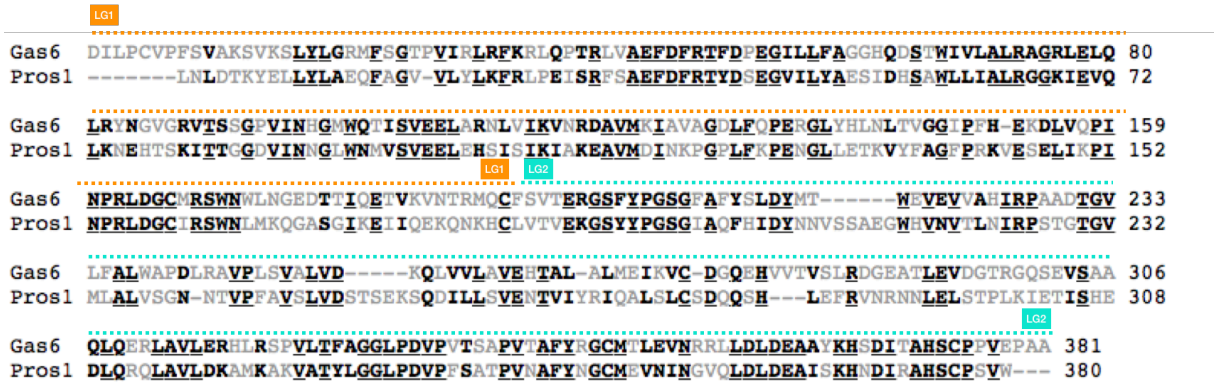


Figure 5: Pairwise alignment of LG1-LG2^{Gas6} and LG1-LG2^{Pros1}: Residues with the same physicochemical properties are marked in black while non-conserved residues are colored in gray. LG1 and LG2 domains are marked with orange and dashed blue lines (above the sequence), respectively.

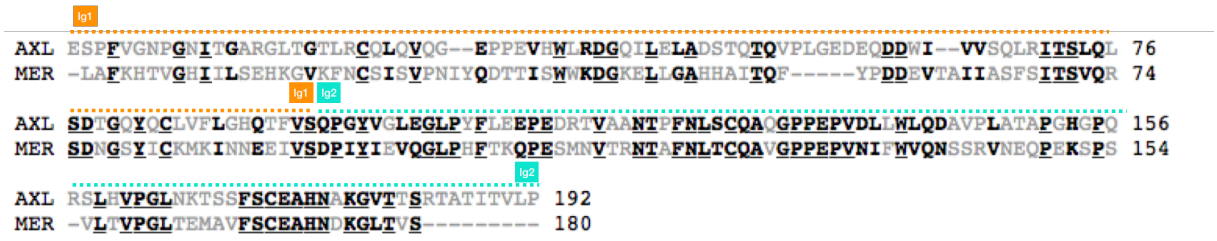


Figure 6: Pairwise alignment of Ig1-Ig2^{Axl} and Ig1-Ig2^{Mer}: Residues having the same physicochemical properties are colored as black while non-conserved residues are colored as gray. Ig1 and Ig2 domains are marked with orange and dashed blue lines (above the sequence), respectively.

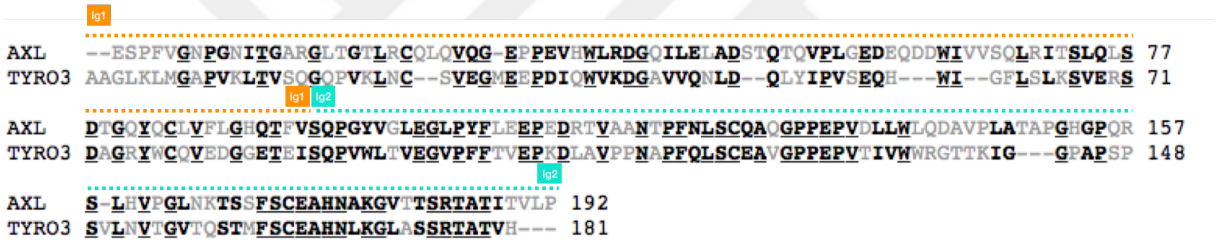


Figure 7: Pairwise alignment of Ig1-Ig2^{Axl} and Ig1-Ig2^{Tyro3}: Residues having the same physicochemical properties are colored in black while non-conserved residues are colored in gray. Ig1 and Ig2 domains are marked with orange and dashed blue lines (above the se the sequence), respectively.

The homology modeling of Ig1-Ig2^{Tyro} and Ig1-Ig2^{Mer} was performed with the I-TASSER web service (https://zhanglab.ccmb.med.umich.edu/I-TASSER/ , see Methods 3.6.2.3) by taking the bound state of Ig1-Ig2^{Axl} (taken from 2C5D) as a template. The same procedure was carried out to homology model LG1-LG2^{Pros1} by taking LG1-LG2^{Gas6} (taken from 2C5D) as a template. In order to assess the quality of the models, I-TASSER criteria was followed, which is encoded in Confidence score (C-Score) and TM-Score. The corresponding C-scores and TM-scores of Ig1-Ig2^{Tyro3}, Ig1-Ig2^{Mer}, LG1-LG2^{Pros1} homology models are given

in Table 1. As explained in the section 3.6.2.3, the C-score ranges between -5 and 2 where the higher C-Score reflects the more confident the model is and TM-score measures the global accuracy of the structures where a TM-Score greater than 0.5 indicates a correct topology. These values in Table 1 reflect that the homology modeling performed by I-TASSER resulted in high quality models.

Table 1: I-TASSER quality measure encoded as C-score, TM-Score and RMSD of models: Pros1, Tyro3 and Mer

	C-score	**TM Score**	**RMSD**
Pros1	1.27	0.89 +- 0.07	4.1 +- 2.7
Tyro3	0.95	0.84+-0.08	3.3 +- 2.3
Mer	0.10	0.73+-0.11	5.1 +- 3.3

4.2 The Structural Modelling of TAM-Ligand Interactions

4.2.1 Construction of the TAM-Ligand Interactions by Structural Alignment

After obtaining Ig1-Ig2^{Tyro3}, Ig1-Ig2^{Mer}, LG1-LG2^{Pros1} with homology modeling, we proceeded with the modeling of the TAM ligand pairs that have not been resolved (Ig1-Ig2^{Tyro3}:LG1-LG2^{Gas6}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, Ig1-Ig2^{Axl}:LG1-LG2^{Pros1}, Ig1-Ig2^{Tyro3}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}). For this, as indicated earlier, we took Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} interactions, as observed in 2C5D as a basis. Initially, the homology models of Ig1-Ig2^{Tyro3} and Ig1-Ig2^{Mer} were aligned on the bound coordinates of Ig1-Ig2^{Axl} (2C5D) with FATCAT structural alignment web tool (http://fatcat.sanfordburnham.org , see Methods 3.6.2.4). The same procedure was carried out to match the coordinates of LG1-LG2^{Pros1} and LG1-LG2^{Gas6} (2C5D). For each case, the same protocol was repeated to match the symmetric copies of the relevant Ig1-Ig2's and LG1-LG2's, since Ig1-Ig2 and LG1-LG2 bind in a 2:2 fashion in a cyclic symmetry (Figure 8.A). The matched coordinates were visualized in PyMOL molecular visualization program (Figure 8). Each relevant complex coordinates were isolated in PyMOL as a different object and saved as a separate coordinate file (in pdb format). As a result, six

complex coordinates (pdb files) were obtained: Ig1-Ig2^{Tyro3}: LG1-LG2^{Gas6}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, Ig1-Ig2^{Axl}:LG1-LG2^{Pros1}, Ig1-Ig2^{Tyro3}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1} (Figure 8.B).

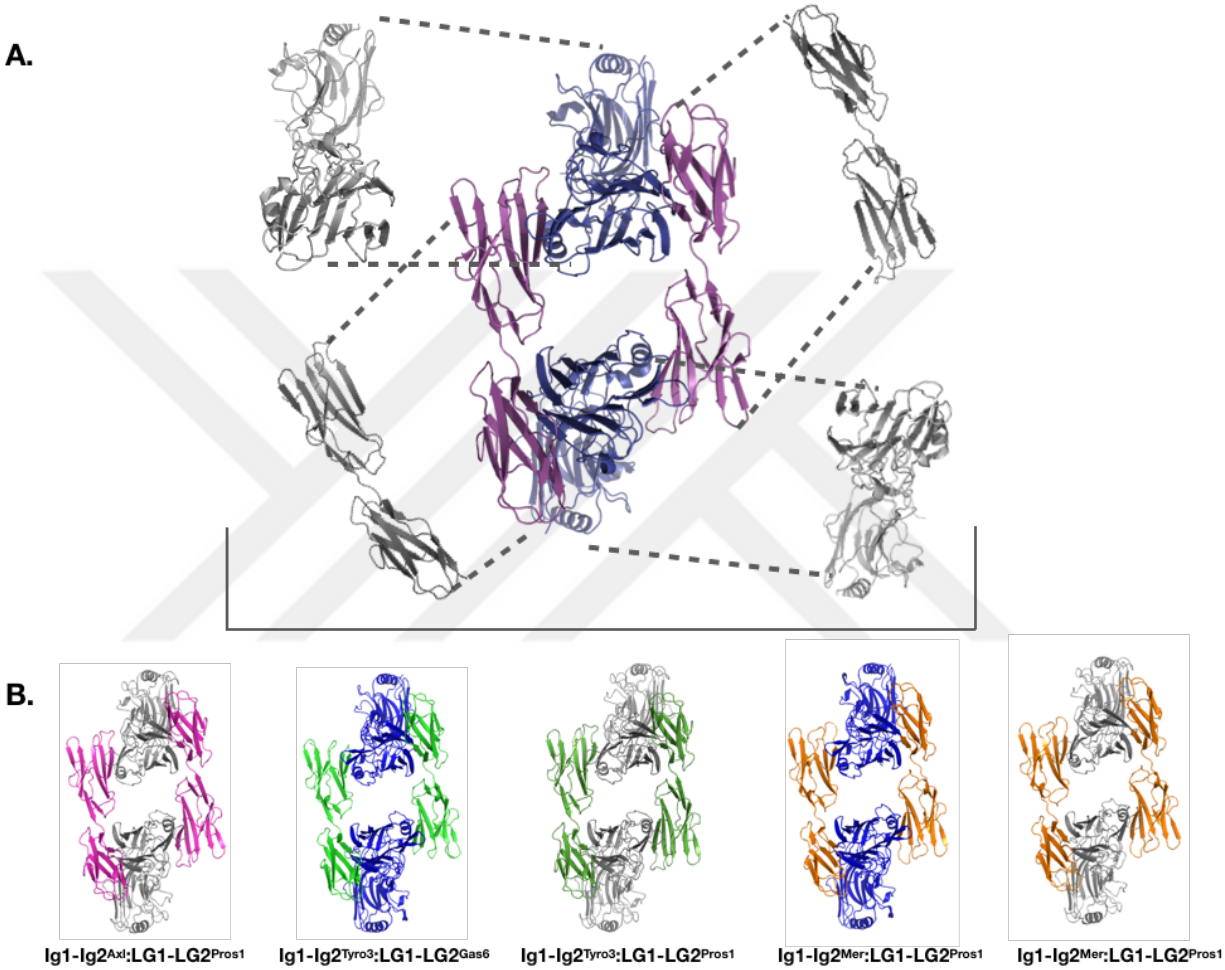


Figure 8: Representation of structural modelling of TAM-Ligand interactions: **A.** Structural alignment of Ig1-Ig2^{Receptors} domains (gray cartoon) on to the bound state of Ig1-Ig2^{Axl} (dark pink cartoon) and LG1-LG2^{Ligand} (gray cartoon) onto bound state of LG1-LG2^{Gas6} (purple cartoon). **B.** Five TAM-ligand interactions built by structural alignment (as depicted in panel A of this figure): Non-binder Ig1-Ig2^{Axl}:LG1-LG2^{Pros1} and binder Ig1-Ig2^{Tyro3}:LG1-LG2^{Gas6}, Ig1-Ig2^{Tyro3}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, respectively.

4.2.2 Refinement of the constructed TAM-Ligand complexes

After completing the crude modeling, as explained above, we moved to the refinement of each complex model. This is a necessity in order to obtain correct geometry and physicochemical complementarity across different interfaces. To that end, we used the interface refinement option of HADDOCK molecular docking platform (https://milou.science.uu.nl/services/HADDOCK2.2/haddockserver-refinement.html , see Methods, (van Zundert, et al., 2016). The standard refinement protocol of HADDOCK performs a short molecular dynamics simulation to generate 20 models. These models slightly differ from each other as each is refined with molecular dynamics protocol starting from a different initial velocity. In the end, the generated models are ranked with the HADDOCK score, which is a sum of *electrostatics* (elec), *van der Waals* (vdW) *and desolvation* terms (desolv) (Eq 1). The top ranking four models, i.e. the four models with the lowest HADDOCK scores, are offered as the final refined complex states.

To refine our complexes with the optimum parameters, we first increased the number models that are sampled during HADDOCK refinement (see Methods 3.6.2.5). This optimization step was carried out by using the crystal structure of Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} (2C5D). Here, instead of 20, we generated 50, 100, 150 and 200 refined structures. This allowed us to see how sampling affects the HADDOCK score distribution (Figure 9). As a result of this analysis, we observed that the energy distribution stops changing after 150 structures. In order to be on the safe side, we generated 200 refined structures for each TAM-ligand complex and isolated the top ranking model as the final solution.

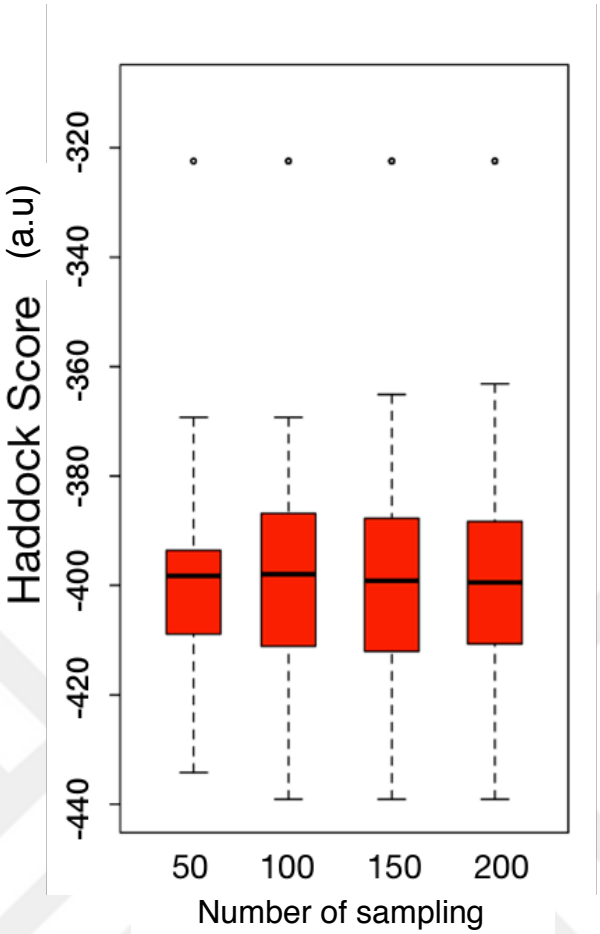


Figure 9: HADDOCK Score distribution of Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} (2C5D) obtained from refinement for different numbers of sampling used, i.e. 50, 100, 150, 200 structures. Box and whiskers were calculated with the whisker lengths equal to 2.0 IQRs (IQR: interquartile range).

4.2.3 Quality Control of the TAM-ligand Complex Models

For ensuring the viability of the top scoring tetramer models, we performed their structure analysis with the PROCHECK webserver (https://servicesn.mbi.ucla.edu/PROCHECK/). Table 2 reports the Ramachandran plot statistics of each model (see Methods 3.6.2.7). All of our models, including refined Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} structure yields better torsion angle than the crystal structure of Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} : first of all, percentages of residues in most favored region are higher in all tetramers than crystal Axl-Gas6 structure while percentage of residues in additional and generously allowed regions are lower than crystal Axl-Gas6 structure. This indicates that our models are safe to use for further hypothesis generation and testing.

Table 2: Ramachandran plot results of tetramer complexes: Residues in most favored regions (%), residues in additional allowed region (%), residues in generously allowed regions (%) and residues in disallowed regions (%) for the crystal structure of Ig1-Ig2^{Axl}: :LG1-LG2^{Gas6} and models: Ig1-Ig2^{Axl}:LG1-LG2^{Pros1} and binder Ig1-Ig2^{Tyro3}:LG1-LG2^{Gas6}, Ig1-Ig2^{Tyro3}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}, Ig1-Ig2^{Mer}:LG1-LG2^{Pros1}.

	Residues in most favored region (%)	Residues in additional allowed region (%)	Residues in generously allowed region (%)	Residues in disallowed region (%)
Axl-Gas6 crystal structure (2C5D)	74.1	23.6	2.3	0.0
Axl-Gas6 Refined-crystal structure	84.8	14.2	0.2	0.8
Axl-Pros model	78.4	17.6	2.2	1.8
Tyro-Gas6 xx	83.1	15.3	0.8	0.7
Tyro-Pros	76.5	19.2	2.1	2.2
Mer-Gas6	83.6	15.2	0.0	1.2
Mer-Pros	77.9	18.1	1.8	2.2

4.3 Interface analysis of the TAM-ligand complex structures

4.3.1 Haddock Scores of TAM-Ligand Interactions

In order to perform the interface analysis of the TAM-ligand pairs, we first investigated the HADDOCK scores of the tetramers. Here, our assumption was that these scores could provide us with a clue regarding the interface complementarity of each TAM-ligand pair. In Figure 10.B, the average HADDOCK scores of the best four ranked TAM-ligand pairs are given. The provided scale highlights that the more negative (red) the score, the more interface complementarity is predicted for the given complex. Intriguingly, HADDOCK score reflects the experimentally observed binding pattern for TAM-Gas6 and TAM-Pros1 complexes: In

TAM-Gas6 complexes, Ig1-Ig2^{Axl}:LG1-LG2^{Gas6} shows the highest complementarity which is followed by Ig1-Ig2^{Tyro3}:LG1-LG2^{Gas6} and Ig1-Ig2^{Mer}:LG1-LG2^{Gas6}. In the TAM-Pros1 complexes, the most complementing pair is found as Ig1-Ig2^{Tyro}:LG1-LG2^{Pros1} which is followed by Ig1-Ig2^{Mer}:LG1-LG2^{Pros1} and Ig1-Ig2^{Axl}:LG1-LG2^{Pros1}. Most importantly, overall energy results indicate that the non-binder complex, Ig1-Ig2^{Axl}:LG1-LG2^{Pros1} ranks the last among all complexes. So, in line with our initial assumption, HADDOCK Scores of the tetramer structures reflect the binding pattern among TAM-Ligand interactions.

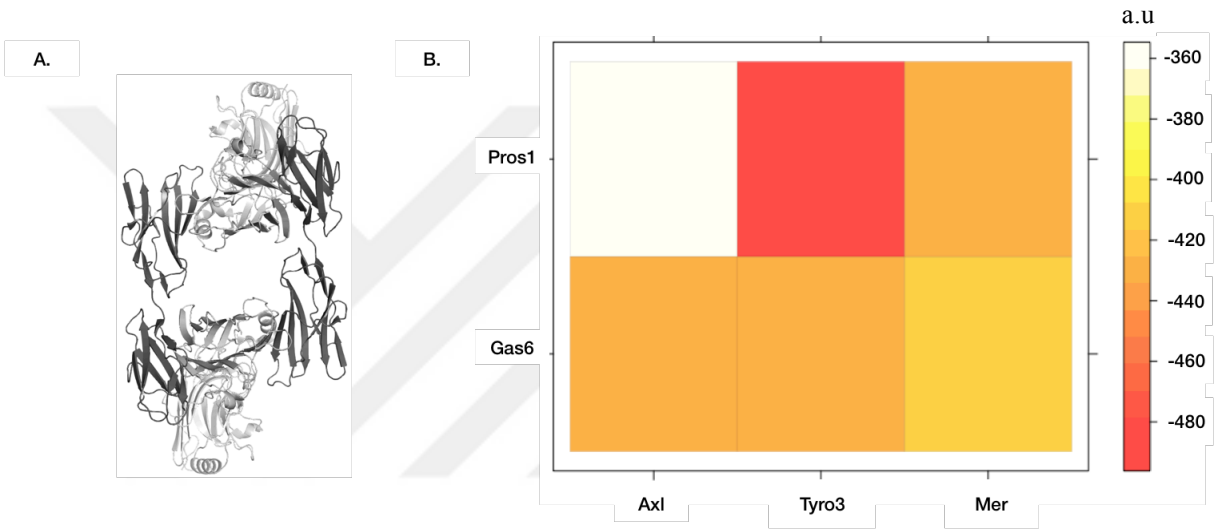


Figure 10: HADDOCK Refinement of Tetramers: **A.** A representative the tetramer structure of which HADDOCK score is calculated. **B.** Average HADDOCK scores of the top four models of TAM-Gas6 (second row) and Pros-TAM (first row) complexes: Ig1-Ig2^{Axl}: LG1-LG2^{Gas6} (-432.9+/-5.03 a.u) ranks the first which is followed by Ig1-Ig2^{Tyro3}: LG1-LG2^{Gas6} (-430.7+-3.7 a.u) and Ig1-Ig2^{Mer}: LG1-LG2^{Gas6} (-416.2+- 7.5 a.u) . In Pros1 complexes, Ig1-Ig2^{Tyro3}: LG1-LG2^{Pros1} (-487.5+-3.5 a.u) ranks the first which is followed by the Ig1-Ig2^{Mer}: LG1-LG2^{Pros1} (-430.1+-2.3 a.u) and Ig1-Ig2^{Axl}: LG1-LG2^{Pros1} (-363.3067+- 5.7 a.u). The color bar ranges between -480 and -360 a.u. HADDOCK score is expressed in arbitrary unit (a.u).

4.3.2 HADDOCK Scores of the Major Interface

As given in the Introduction (Section 2.2.3), the amino acids taking part in the minor interface formation are well conserved across the TAM family (Sasaki, et al., 2006). Taking this information as a basis, we isolated the following major interfaces from the modeled tetramers. $Ig1^{Axl}$:$LG1^{Gas6}$, $Ig1^{Axl}$-$LG1^{Pros1}$, $Ig1^{Tyro3}$-$LG1^{Gas6}$, $Ig1^{Tyro3}$-$LG1^{Pros1}$, $Ig1^{Mer}$-$LG1^{Gas6}$, $Ig1^{Mer}$-$LG1^{Pros1}$. Figure 11.A shows the average HADDOCK score of four best ranked TAM-Ligand major interface. As in the HADDOCK Score of tetramer structures (Figure 10.B), HADDOCK score reflects the experimentally observed binding pattern for TAM-Gas6 and TAM-Pros1 complexes. This time, $Ig1^{Axl}$:$LG1^{Gas6}$ shows the highest complementarity among all TAM-Ligand dimers which is followed by $Ig1^{Tyro3}$:$LG1^{Gas6}$, $Ig1^{Mer}$:$LG1^{Gas6}$, $Ig1^{Tyro3}$:$LG1^{Pros1}$ and $Ig1^{Mer}$:$LG1^{Pros1}$ (Figure 11.B). $Ig1^{Axl}$:$LG1^{Pros1}$, again, ranks the last among all complexes. These results indicate that the simple energetics terms of HADDOCK can clearly explain the binding profile differences of the strongest complex Axl-Gas6 and weakest complex Axl-Pros. This result also reflects that binding signal is encoded in the major interface.

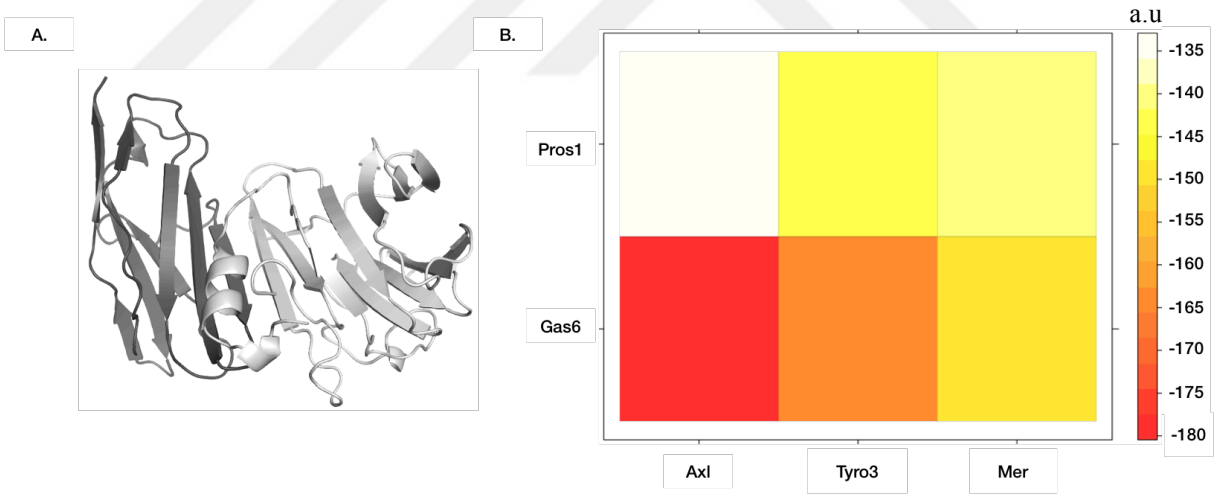


Figure 11: A. A representation of the dimer structure where HADDOCK Score of their interface were calculated. **B**. Average major interface HADDOCK Scores of 4 best ranked structures of Gas6-TAM (bottom) and Pros1-TAM (above) complexes. In the figure red s score shows the more energetically favorable complexes. In overall, $Ig1^{Axl}$:$LG1^{Pros1}$ ranks the last in all complexes. $Ig1^{Axl}$:$LG1^{Gas6}$ (-177.54+-4.3 a.u.) ranks the first which is followed by $Ig1^{Tyro3}$:$LG1^{Gas6}$ (-164.5+-1.7 a.u.) and $Ig1^{Mer}$:$LG1^{Gas6}$ (-148.41+-1.8 a.u.) . In Pros1 complexes,

$Ig1^{Tyro3}$:$LG1^{Pros1}$ (-143.88+-2.9 a.u.) ranks the first which is followed by the $Ig1^{Mer}$:$LG1^{Pros1}$ (-139.56+-1.2 a.u.) and $Ig1^{Axl}$:$LG1^{Pros1}$ (-135.83+- 2.6 a.u.). HADDOCK score is expressed in a.u.

4.3.3 Energetics of Major Interface across Axl-Gas6 and Axl-Pros

As the total HADDOCK Score discriminate, the strongest binder Axl-Gas6 and weakest binder Axl-Pros1, we decided to investigate these two extreme cases. Moreover, as simple energetic terms could discriminate between these two complex, we hypothesized that selectivity signals of Axl's selectivity are encoded in these energy terms. Thus, we searched for the dominant force in Axl's selectivity across two ligands. To do that, we analyzed electrostatics, van der Waals (vdW) and Buried Surface Area (BSA) of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ dimers, respectively. We found that vdW and BSA is highly similar in both complexes. On the contrary, electrostatics of $Ig1^{Axl}$:$LG1^{Gas6}$ interface is 3.6 times higher than $Ig1^{Axl}$:$LG1^{Pros1}$ interface (Figure 12). These results uncover that Axl's selectivity might be controlled mostly by electrostatic interactions rather than hydrophobic forces.

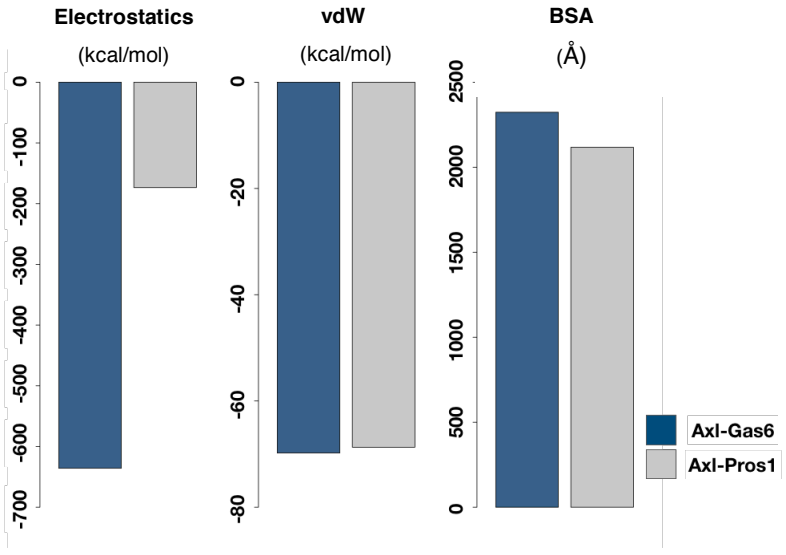


Figure 12: Average individual HADDOCK Score components of first four ranked Axl-Gas6 and Axl-Pros1 structures. Electrostatics (blue bar) of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ (gray bar) are -635.85+-29.3 kcal/mol and -173.51+- 21.4 kcal/mol, vdW (middle) scores are -69.

79+- 1.24 kcal/mol and -68.90 +- 2.2 kcal/mol, and BSA (right) scores are found as 2323.12+-20.9 Å and 2117.28 +-28.7 Å, respectively.

4.3.4 Electrostatic Potential Maps of the major interfaces across the $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ complexes

To understand how residues are distributed across the Axl's Ig1 domain and LG of Gas6 and Pros1, as well as electrostatic complementarity of major interfaces ($Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$), the electrostatics potential map of Axl, Gas6, and Pros were calculated by the APBS plugin of PyMOL (see Methods 3.2.6.8). As the first four ranked structures reflect similar energies, here and thereafter, we analyzed only the top ranking refined structure. In figure 13, red indicates a surface of negative potential while blue of positive potential (ranging between charges -5 and +5). The patches taking part in the major interface formation are encircled in the figure. This analysis showed that Axl's major binding site is mainly dominated by negatively charged potential which might be complemented well with a ligand of a positively charged (blue) surface. When the major binding sites of Gas6 and Pros1 are compared, it was seen that only Gas6 would complement Axl with its positively charged patch.

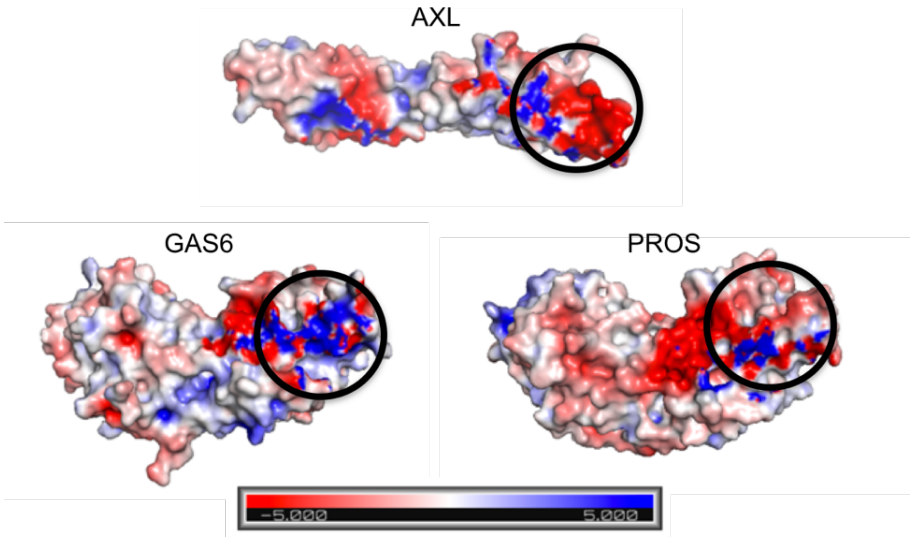


Figure 13: Electrostatics potential map of Ig1-Ig2 of Axl and LG1-LG2 of Gas6 and Pros1. The region taking part in the major interface formation is encircled. The more red (toward -5) means the more negative surface potential while the more blue (toward +5) means the more positive surface potential.

4.3.5. Analysis of the electrostatics-driven interactions across the major interface of Axl-Gas6 and Axl-Pros1

4.3.5.1. The Distribution of the Hydrogen Bonds and Salt-Bridges

As we found out that the electrostatics is the major driving force for the Axl receptor binding, we decided to focus on the electrostatics-driven interactions: hydrogen bonds and salt-bridges. Hydrogen bonds and salt bridges are the major non-covalent contributors of electrostatics interactions. As a result of our analysis, the number of h-bonds in $Ig1^{Axl}$:$LG1^{Gas6}$ interface dropped by 56.75% in $Ig1^{Axl}$:$LG1^{Pros1}$ (Figure 14). Likewise hydrogen bond number, number of salt-bridges in $Ig1^{Axl}$:$LG1^{Gas6}$ interface dropped by 85.71% in $Ig1^{Axl}$:$LG1^{Pros1}$ interface. Reduction in number of electrostatic-driven interactions in $Ig1^{Axl}$:$LG1^{Pros1}$ interface points out Axl and Gas6 have more residues having capacity to make hydrogen bonds and salt-bridges across their major interface.

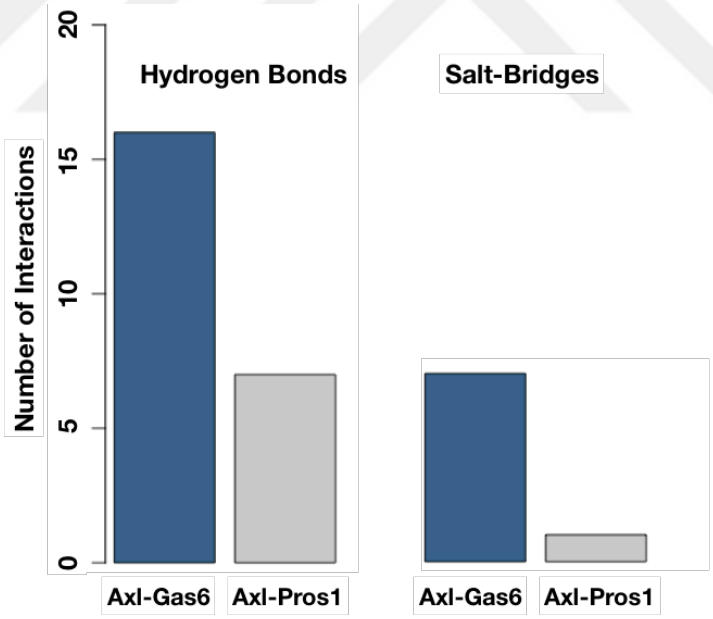


Figure 14: Number of H-bond and salt-bridges at the major interface of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ dimers. Number of interactions at $Ig1^{Axl}$:$LG1^{Gas6}$ interface are colored in blue while $Ig1^{Axl}$:$LG1^{Pros1}$ in gray.

4.3.5.2. Depiction of the most energetically favorable Axl interface Residues Facing Gas6 and Pros1

To decide the residues contributing the binding most compared the other residues, we investigated the energetics of the $Ig1^{Axl}$'s interface amino acids that interacts with $LG1^{Gas6}$ & $LG1^{Pros1}$. Per-amino acid energetics are calculated as a result of the HADDOCK refinement and outputted in a file called, *ene-residue.disp* file (see Methods 3.6.2.5.3). This file describes electrostatics, vdW and electrostatics+vdW contributions of each interface amino acid to the formation of the inter-molecular interaction. Here, we considered both electrostatics+vdW and only electrostatics scores of $Ig1^{Axl}$ interface amino acids, as they might be critical residues in binding. There are 31 $Ig1^{Axl}$:$LG1^{Gas6}$ (on Axl) interface amino acids, whereas 27 $Ig1^{Axl}$:$LG1^{Pros1}$ (on Axl) interface amino acids. We plotted the distribution of each case via box and whisker plot (length of the whiskers: 2xIQR). We see that Axl residues in $Ig1^{Axl}$:$LG1^{Gas6}$ complex shows much lower energies and lower mean value than in $Ig1^{Axl}$:$LG1^{Pros1}$ complex. ARG 48^{Axl}, GLU 56^{Axl}, GLU 59^{Axl}, GLU 70^{Axl}, GLU 73^{Axl}, GLU 83^{Axl} were found as negative outliers in $Ig1^{Axl}$:$LG1^{Gas6}$ complex, however only ARG 48^{Axl} and GLU 59^{Axl} were found as negative outliers in the $Ig1^{Axl}$:$LG1^{Pros1}$ complex (Figure 15). These outlier residues have the lowest energy among all residues at major interface meaning that they contribute to binding energy more than other residues. Also, all are charged residues having potential to make salt-bridges and/or hydrogen bonds with another residue across interface. The common residues in both complexes are residue ARG 48^{Axl} and GLU 59^{Axl}. These residues of $Ig1^{Axl}$ have the potential to make salt-bridges or hydrogen bonds in both complexes. These identified residues have potential to be SDPs of Axl.

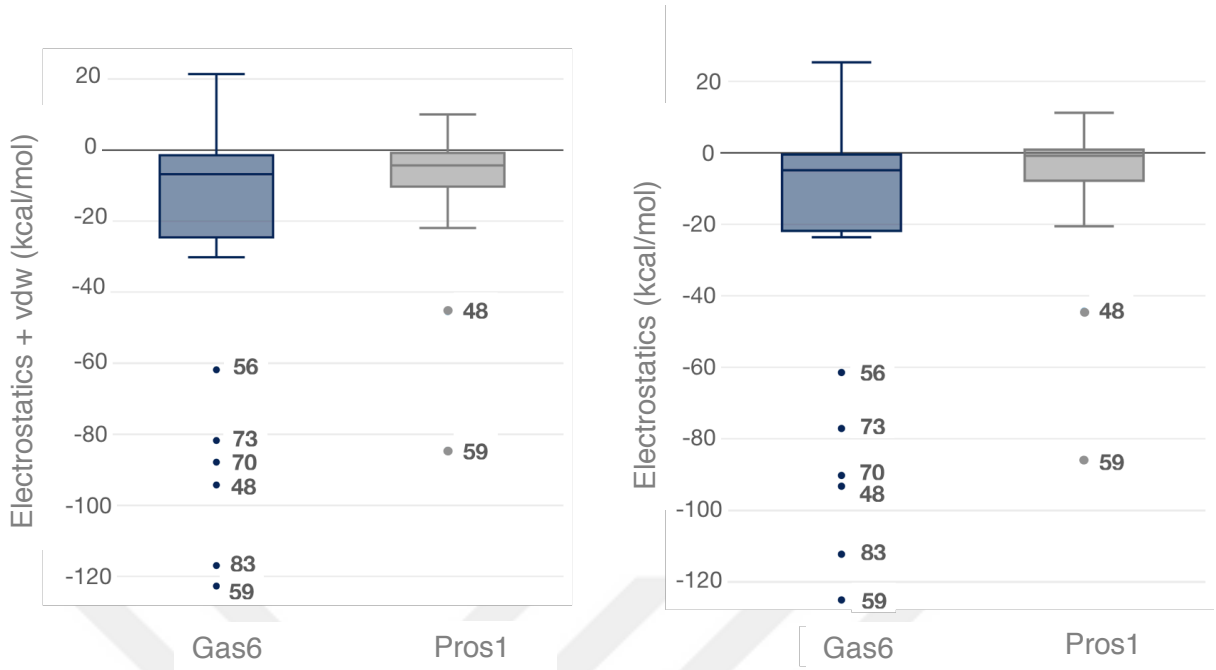


Figure 15: Electrostatics+vdW (left) and electrostatics (right) distribution of Axl interface residues facing Gas6 and Pros1, respectively. The residues across Axl-Gas6 or Axl-Pros1 interface in both analyzes are the same with little energy differences. 48^{Axl}, 56^{Axl}, 59^{Axl}, 70^{Axl}, 73^{Axl}, 83^{Axl} were found as negative outliers in Axl-Gas6 complex, however only 48^{Axl} and 59^{Axl} were found as negative outliers in the Axl-Pros complex. Box and whiskers were calculated where whisker is equal to 2. Energy distribution of residues at $Ig1^{Axl}$:$LG1^{Gas6}$ interface are colored in blue while $Ig1^{Axl}$:$LG1^{Pros1}$ in gray.

4.4 Tracing the dynamic interaction profile of Axl-Gas6's and Axl-Pros' Major Interface by Molecular Dynamics

The analysis described under section 4.3 provided us with possible SDPs across Axl's major interface on $Ig1^{Axl}$. In order to probe the viability of the predicted SDPs, performed MD simulations of $Ig1^{Axl}$:LG^{Gas6} and $Ig1^{Axl}$:$LG1^{Pros1}$ complexes were performed. Here, the aim was to trace the dynamic profile of the predicted SDPs across Axl's major interface, such that the molecular grounds of Axl's binding profile could be dissected. For this, $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros}$ were simulated for 200 ns by using the GROMACS 5.1.4 package (Van Der Spoel, et al., 2005) (see Method 3.6.2.8). For each complex, three extra parallel simulations

were run, which resulted in the generation of 1.6 microsecond simulation in total. The initial seeds, number of water atoms, number of ion elements and protein atoms in each simulation, can be found in Table 3.

Table 3: The 200 ns parallel simulations performed with the GROMACS 5.1.4 package. Their corresponding timescales, initial velocities (random seed), number of solvents (water), ions and protein atoms described in the simulation set up are listed in different columns.

Simulation Code	Random Seed	Number of water atoms	Number of Ion Elements	Number of Protein atoms
Axl-Gas6-Parallel1	500	47367	51 NA, 48 CL	4449
Axl-Gas6-Parallel2	1992	47373	51 NA, 48 CL	4449
Axl-Gas6-Parallel3	142	47376	51 NA, 48 CL	4449
Axl-Gas6-Parallel4	225	47376	51 NA, 48 CL	4449
Axl-Pros1-Parallel1	500	47814	58 NA, 49 CL	4462
Axl-Pros1-Parallel2	1992	47799	58 NA, 49 CL	4462
Axl-Pros1-Parallel3	142	47814	58 NA, 49 CL	4462
Axl-Pros1-Parallel4	225	47817	58 NA, 49 CL	4462

4.5.2 Quality Control of the Molecular Dynamics Simulations

The quality of each simulation was controlled by analyzing the convergence of its thermodynamical parameters and calculating the minimum distance between periodic images. The results of these analyses reflected that the performed simulations are technically sound (Appendix B). The structural quality control tests were performed through evaluating RMSDs (from the initial structure) and the Rg of the protein complexes with respect to the time.

4.5.3 Root Mean Square Deviations

RMSD is defined as the average displacement of the atoms compared to a reference state (see Method 3.6.3.8.2). Thus, it is used as a measure of similarity among biomolecules where smaller RMSD indicates similar structures. RMSD calculations over the time course of the trajectory also informs us on the structural convergence towards an equilibrium state. A MD simulation is considered to reach equilibrium, when its RMSD values start to oscillate around a fixed mean value. Following this criterion, $Ig1^{Axl}$:LG^{Gas6} determined to reach an equilibrium in 25 ns whereas $Ig1^{Axl}$:LG^{Pros1} in 50 ns (Appendix C). The fact that $Ig1^{Axl}$:LG^{Pros1} reaches equilibrium in a longer time than $Ig1^{Axl}$:LG^{Gas6} could be due to two facts: (i) $Ig1^{Axl}$:LG^{Pros1} is a model and thus needs more time to reach an equilibrium state, (ii) $Ig1^{Axl}$:LG^{Pros1} is more unstable than $Ig1^{Axl}$:LG^{Gas6}. After trimming the first 25 ns from the $Ig1^{Axl}$:LG^{Gas6} and 50 ns from the $Ig1^{Axl}$:LG^{Pros1} simulations, 175 ns of production run for $Ig1^{Axl}$:$LG1^{Gas6}$ and 150 ns for $Ig1^{Axl}$:$LG1^{Pros1}$ were obtained. Further analyses were carried out on the production runs of each complex.

As the first structural analysis, the average backbone RMSDs of all the simulations run for a complex was calculated by taking the first structures as the reference (Figure 16, see Methods 3.6.3.8.2). As a result, $Ig1^{Axl}$:LG^{Gas6} simulations reflect 40% smaller average RMSD (~0.15 nm) than $Ig1^{Axl}$:LG^{Pros1} simulations (~0.25 nm). Also, $Ig1^{Axl}$:LG^{Pros1} fluctuates between higher RMSD values (minimum: 0.15 nm and maximum: 0.36 nm) than $Ig1^{Axl}$:LG^{Gas6} (minimum: 0.13 nm and maximum: 0.22 nm).

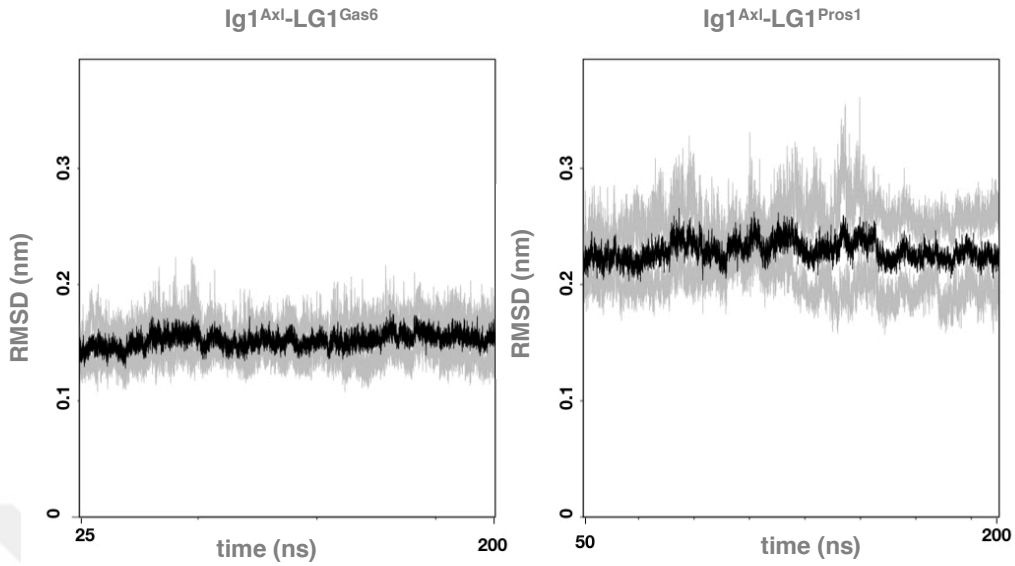


Figure 16: Backbone atoms' RMSD of $Ig1^{Axl}$:LG^{Gas6} (left) and $Ig1^{Axl}$:LG^{Pros1} (right) simulations over 175 and 150 ns trajectory with respect to their initial state. Black lines indicate the mean RMSD of the four simulations while gray lines indicate minimum and max maximum RMSDs observed for that given time points.

4.5.3 Radius of gyration

R_g reflects the globular shape and the compactness of a biomolecule (see Methods 3.6.3.8.2). So, in the case of a protein-protein complex, a smaller R_g might denote a compact structure, thus a tighter complex. To understand which protein complex reflects a more compact structure throughout the simulation, we analyzed the R_g of $Ig1^{Axl}$:LG^{Gas6} and $Ig1^{Axl}$:$LG1^{Pros1}$ for all of their production runs. As a result, $Ig1^{Axl}$:LG^{Gas6} consistently exerted a R_g around ~1.92 nm over the course of the simulation (Figure 16). On the other hand, R_g of the $Ig1^{Axl}$:$LG1^{Pros1}$ simulations fluctuated between ~1.92 nm and ~2.05 nm, indicating a certain degree on dependency on the initial velocity of the simulation (Figure 17). It also indicates that $Ig1^{Axl}$:$LG1^{Gas6}$ behaves more compact than the $Ig1^{Axl}$:$LG1^{Pros1}$ during the time course of trajectory.

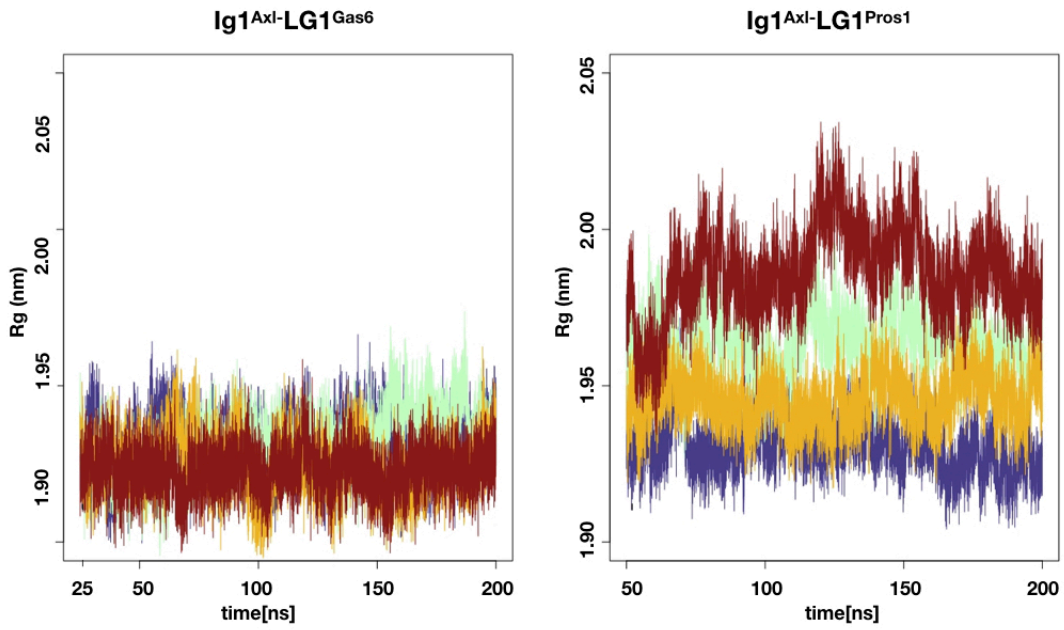


Figure 17: Backbone Rg fluctuations of $Ig1^{Axl}$:LG^{Gas6} (left) and $Ig1^{Axl}$:LG^{Pros1} (right) simulations over production run (175 ns and 150 ns, respectively). Radius of gyration is expressed in nm. Lines are colored according to their parallel simulation id (purple: p parallel 1, green: parallel 2, orange: parallel 3, red: parallel 4).

4.5.4 Root Mean Square Fluctuations

RMSFs measure a particular atom's or group of atoms' displacement (i.e. flexibility) within the course of time. Here, $Ig1^{Axl}$'s residue-based RMSFs were depicted with box and whisker plots. This analysis aims to compare the range of Axl amino acid flexibilities when bound to $LG1^{Gas6}$ or $LG1^{Pros1}$ for different simulations (Figure 18). In Figure 18, it is seen that all Axl-Pros1-Parallel and Axl-Gas6-Parallel simulations show similar RMSF distributions.
We further investigated whether we could depict an Axl residue which is stabilized across $LG1^{Gas6}$ and gets non-stabilized across $LG1^{Pros1}$ for all simulations. GLU 85^{Axl} turned out to be such a residue. It has around 0.26 nm RMSF for $Ig1^{Axl}$:$LG1^{Pros1}$simulations and smaller than 0.25 nm RMSF for $Ig1^{Axl}$:$LG1^{Gas6}$ simulations. This residue also found to be at the major interface (Figure 19).

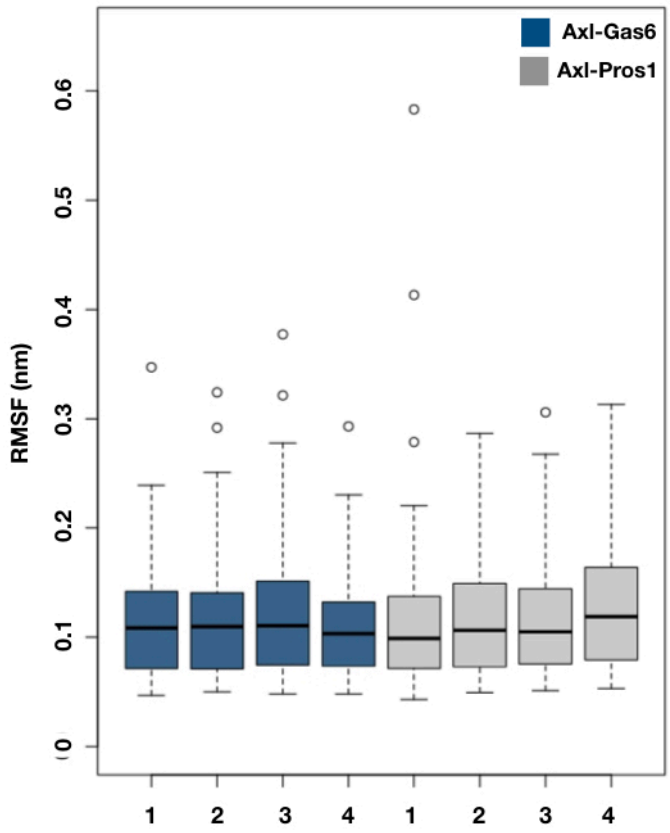


Figure 18: Distribution of $Ig1^{Axl}$ residue-based RMSFs of the (from left to right) Axl-Gas6-Parallel1, Axl-Gas6-Parallel2, Axl-Gas6-Parallel3, Axl-Gas6-Parallel4, Axl-Pros1-Parallel1, Axl-Pros1-Parallel2, Axl-Pros1-Parallel3, Axl-Pros1-Parallel4 the production runs, respectively. $Ig1^{Axl}$:$LG1^{Gas6}$ simulations are colored in blue and $Ig1^{Axl}$:$LG1^{Pros1}$ are colored in gray.

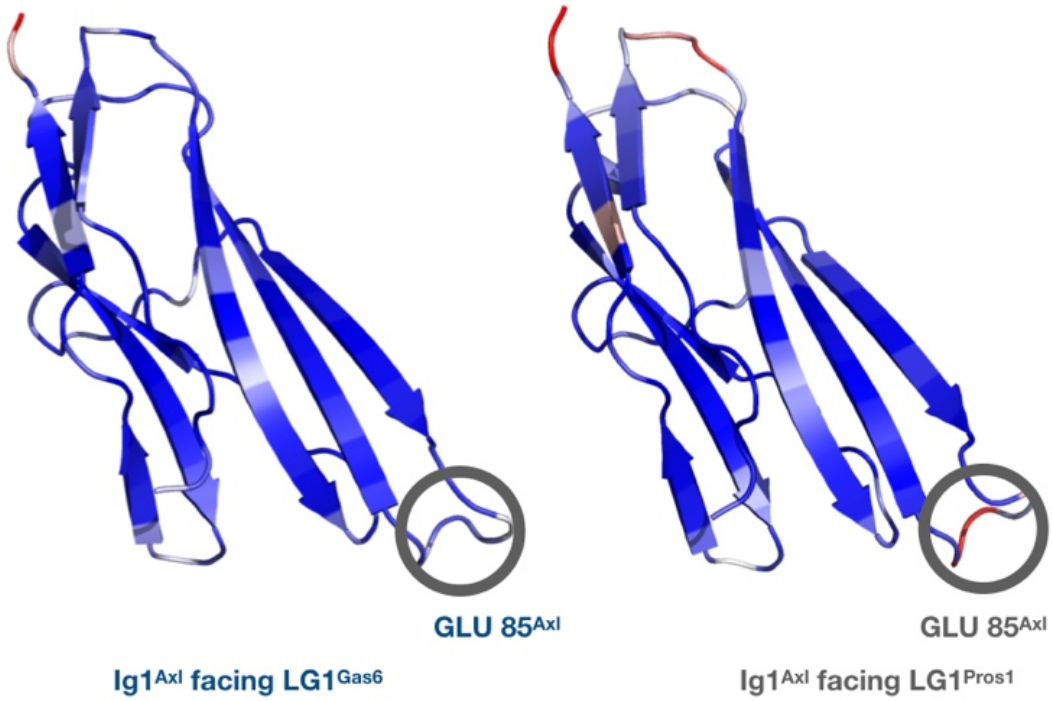


Figure 19: Flexibility of $Ig1^{Axl}$ residues when complexed with $LG1^{Gas6}$ (left) or $LG1^{Pros1}$ (right). The encircled $GLU85^{Axl}$ found to be highly flexible in $Ig1^{Axl}$:$LG1^{Pros1}$ simulations whereas it gets stabilized in $Ig1^{Axl}$:$LG1^{Gas6}$ simulations. The RMSF values are reflected in the color from red (more flexible) through the blue (more stable). These RMSF values were extracted from Axl-Gas6-Parallel4 and Axl-Pros1-Parallel4 simulation.

4.5.5 Interaction Analysis of Axl-Gas6 and Axl-Pros1 Major Interfaces

To examine the interaction profiles of each trajectory, we calculated the number and pair of hydrophobic interactions, hydrogen bonds and ionic interactions of each complex over the time course of each parallel simulation (see Methods 3.6.2.9.1). The interaction profiles are calculated for the coordinates which are extracted in each 500 ps of the simulation (corresponding to 350 snapshots for $Ig1^{Axl}$:$LG1^{Gas6}$ and 300 snapshots for $Ig1^{Axl}$:$LG1^{Pros1}$).

4.5.5.1 Comparison of the Hydrophobic Interaction Profiles

To examine the how hydrophobic interactions across $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ change over time, we calculated the number of pairwise hydrophobic interactions across both interfaces. The mean number of hydrophobic contacts in $Ig1^{Axl}$:$LG1^{Gas6}$ simulations fluctuates around 45-50. This number reduces to 35-45 in the case of $Ig1^{Axl}$:$LG1^{Pros1}$ simulations, except the first $Ig1^{Axl}$:$LG1^{Pros1}$ simulation (Figure 20). As such, as Rg of the simulation indicated, $Ig1^{Axl}$:$LG1^{Gas6}$ makes a tighter complex than $Ig1^{Axl}$:$LG1^{Pros1}$ which also results in the formation of more hydrophobic contacts.

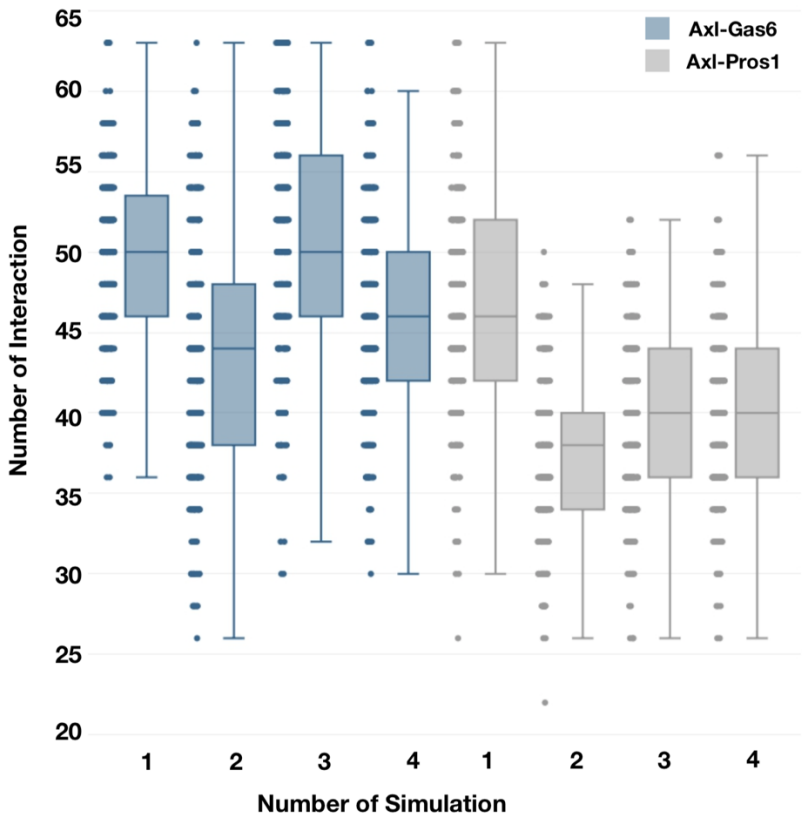


Figure 20: Distribution of the number of pairwise hydrophobic contacts over four $Ig1^{Axl}$:$LG1^{Gas6}$ parallel simulations and $Ig1^{Axl}$:$LG1^{Pros1}$ parallel simulations. $Ig1^{Axl}$:$LG1^{Gas6}$ simulations were colored in blue and $Ig1^{Axl}$:$LG1^{Pros1}$ simulations were colored in gray. Box and whiskers were calculated with the whisker length of 1.5 IQRs.

4.5.5.2 Comparison of the Hydrogen Bonding Profiles

As depicted in the section 4.3.2, Axl's selectivity might be directed by complementary electrostatic interactions. Expanding on this, we analyzed the number of hydrogen bonds formed across of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ interfaces over time. In Figure 21, it is seen that an average complex makes 12-14 hydrogen bonds in the case of $Ig1^{Axl}$:$LG1^{Gas6}$ simulations, while these values shift to 6-12 for $Ig1^{Axl}$:$LG1^{Pros1}$. As in the distribution of hydrophobic contacts (Figure 20), first $Ig1^{Axl}$:$LG1^{Pros1}$ behaves more like the $Ig1^{Axl}$:$LG1^{Gas6}$ simulations. As an interesting point here, our static analyses given in section 4.3.5.1, indicated that there are 16 hydrogen bonds formed across $Ig1^{Axl}$:$LG1^{Gas6}$ and 7 across $Ig1^{Axl}$:$LG1^{Pros1}$ interfaces. However, as reflected in Figure 21, these numbers highly fluctuate in the case of dynamics simulations.

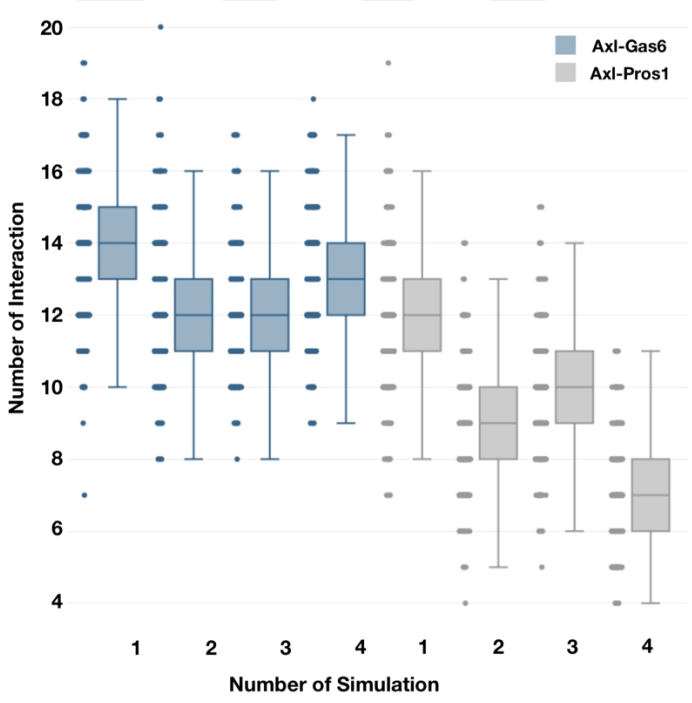


Figure 21: Distribution of number of hydrogen bonds over four $Ig1^{Axl}$:$LG1^{Gas6}$ parallel simulations and $Ig1^{Axl}$:$LG1^{Pros1}$ parallel simulations. $Ig1^{Axl}$:$LG1^{Gas6}$ simulations were colored in blue and $Ig1^{Axl}$:$LG1^{Pros1}$ simulations were colored in gray. Box and whiskers were calculated with the whisker length of 1.5 IQRs.

4.5.5.3 Comparison of the Salt Bridge Profiles

As charged residues are a potential subfamily of SDPs (see Introduction 2.3), we calculated how charged-charged interactions are distributed over each trajectory. Figure 21 reported that the mean values of $Ig1^{Axl}$:$LG1^{Gas6}$ varies between 4 and 5 interactions. However, the number of salt bridges formed during $Ig1^{Axl}$:$LG1^{Gas6}$ simulations fluctuate between 1 to 10 over time. On the other hand, in the case of $Ig1^{Axl}$:$LG1^{Pros1}$ there is even not enough number of salt-bridges to construct box-and-whisker plots, indicating that $Ig1^{Axl}$ is losing its salt-bridge making capacity when faced to $LG1^{Pros1}$. This is also in line with our previous observation where we initially analyzed our static structures (section 4.3.5.1): The most extreme difference was seen in the salt-bridge making capacity of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ complexes.

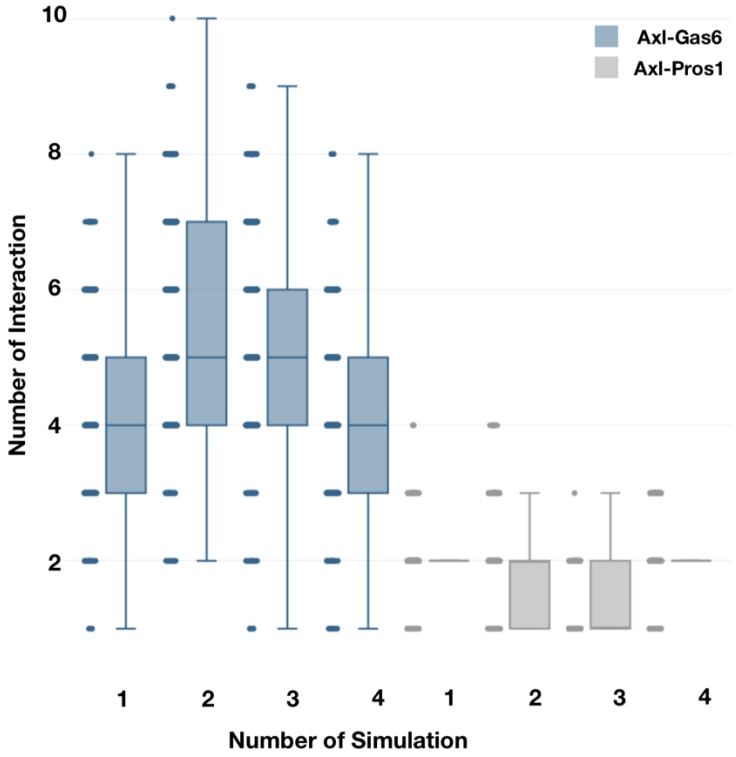


Figure 22: Distribution of number of salt-bridges over four $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ simulations (colored as blue and gray, respectively). Box and whiskers were calculated with the whisker length of 1.5 IQRs.

4.5.5.4 Consistent Salt-Bridges formed across Axl-Gas6 and Axl-Pros1 Major Interface

To deduce the salt-bridges that are mostly seen across $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ interfaces, we kept the salt bridges, which were seen at least in 25% within the time frame of a given trajectory. Afterwards, we decided to concentrate only on the salt bridges that are consistently seen in all parallel simulations. This analysis is performed since we hypotheses that possible selectivity determining positions (SDPs) of $Ig1^{Axl}$ are charged residues making strong interactions with residues on $LG1^{Gas6}$. Table 4 shows salt-bridges meeting the above-given criteria for Axl-Gas6-Parallel1, Axl-Gas6-Parallel2, Axl-Gas6-Parallel3 and Axl-Gas6-Parallel4 simulations, respectively. For these simulations, the consistently observed salt-bridges are: GLU 70^{Axl}:ARG 414^{Gas6}, ASP 73^{Axl} :ARG 313^{Gas6}, GLU 83^{Axl} :ARG 467^{Gas6}, GLU 59^{Axl} :ARG 310^{Gas6} and ARG 48^{Axl} :ASP 455^{Gas6} (Table 4 – orange pairs). The mostly found salt bridge in each simulation is shown as bold. The structural positioning of these salt-bridges are illustrated in the Figure 23.

Table 4: Consistent Salt-Bridges in Axl-Gas6-Parallel1, Parallel2, Parallel3 and Parallel4 simulations, respectively. The residue pairs of $Ig1^{Axl}$ and $LG1^{Gas6}$ where salt bridges form and the percentage of the interaction occurrence across simulations are listed in different columns.

$Ig1^{Axl}$	$LG1^{Gas6}$	Axl-Gas6-Parallel1 (%)	Axl-Gas6-Parallel2 (%)	Axl-Gas6-Parallel3 (%)	Axl-Gas6-Parallel4 (%)
70	**414**	**81.14**	34.57	55.41	70.57
73	**313**	72.86	**81.71**	69.43	**76.86**
83	**467**	60.29	48	68.57	71.14
59	**310**	59.43	32.86	**83.14**	56.86
48	**455**	43.71	44	53.43	30.57
70	312	30.57	40	---	---
88	467	28.29	28	26	---
83	299	---	33.42	60.29	34.86

85	308	---	---	36.86	---
96	454	---	44.71	27.71	---
56	308	---	---	---	25.14
48	460	---	56.86	---	---
59	414	---	56.29	---	---

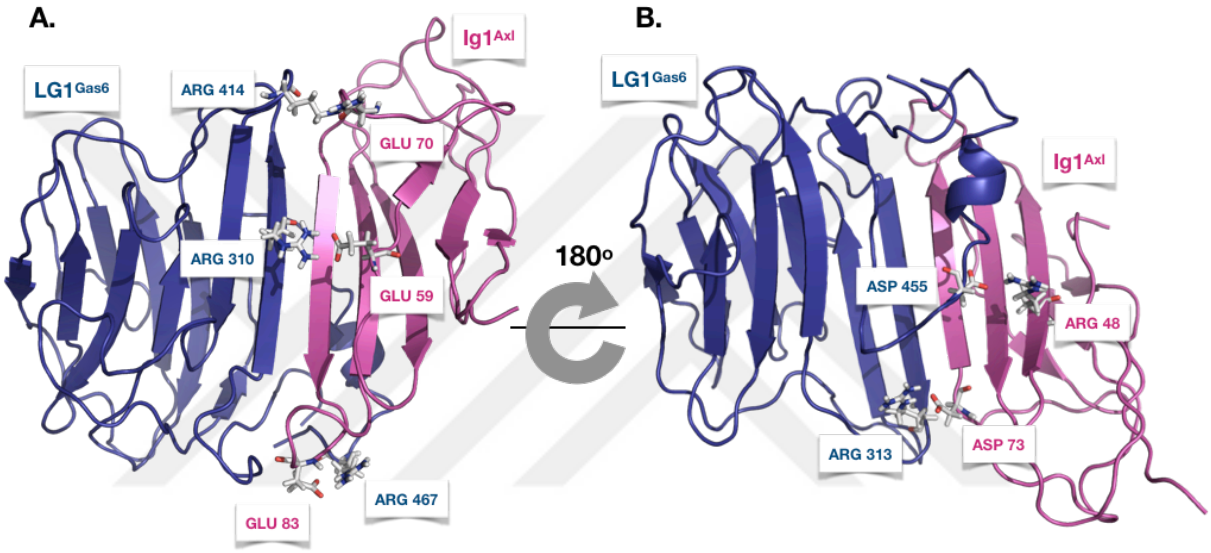


Figure 23: The structural depiction of the consistent salt-bridges formed across $Ig1^{Axl}$:$LG1^{Gas6}$. $LG1^{Gas6}$ is colored in blue and $Ig1^{Axl}$ is colored in magenta. **A.** The depicted salt bridges are GLU 70^{Axl}:ARG 414^{Gas6}, GLU 83^{Axl} :ARG 467^{Gas6}, GLU 59^{Axl} :ARG 310^{Gas6} **B.** 180º rotation of structure around x axis: The depicted salt bridges are ARG 48^{Axl} :ASP 455^{Gas6}, ASP 73^{Axl} :ARG 313^{Gas6} .

The same analysis was performed for the parallel $Ig1^{Axl}$:$LG1^{Pros1}$ simulations. Their corresponding analyses are given in Table 5. It is seen that among frequently observed three salt bridges, only one is consistently observed in all of the parallel simulations. This salt-bridge is formed between GLU 59^{Axl} :LYS 314^{Pros1}. The other four $Ig1^{Axl}$ residues (ARG 48^{Axl}, GLU 70^{Axl}, ASP 73^{Axl} and GLU 83^{Axl}), which should make salt-bridges with its ligand, cannot find a partner on $LG1^{Pros1}$.

So, these results yield that GLU 59^{Axl} can make strong interaction in both complexes. However, the others (ARG 48^{Axl} , GLU 70^{Axl}, ASP 73^{Axl} and GLU 83^{Axl}) cannot find a residue on LG1^{Pros1} to make electrostatic interaction. As these residues loses their salt-bridge making capacity, they probably start to interact with water. Consequently, this residue-water interaction induces destabilization of Ig1^{Axl} . Thus, those salt-bridge making residues except GLU 59^{Axl} on the face of Ig1^{Axl} might be the possible SDPs. These four SDPs were also predicted as possible SDPs with the static analysis of structures (Figure 15).

Table 5: Consistent Salt-Bridges in Axl-Pros1-Parallel1, Parallel2, Parallel3 and Parallel4 simulations, respectively. The residue pairs of Ig1^{Axl} and LG1^{Pros1} where salt bridges form and the percentage of the interaction occurrence across simulations are listed in different columns.

Ig1^{Axl}	LG1^{Pros1}	Axl-Pros1-Parallel1	Axl-Pros1-Parallel2	Axl-Pros1-Parallel3	Axl-Pros1-Parallel4
59	**314**	**97.67**	**97.67**	**94.33**	**94.33**
59	316	--	--	--	66.66
48	465	64.67	51.67	--	33

4.5.5.4 Analysis of the Energetics of Consistent Axl-Gas6 and Axl-Pros Salt-bridges

After finding out of critical salt-bridges and possible SDPs, we calculated their mean energies (see Method 3.6.2.9.2, Table 6). As a result, the most energetically favorable interaction was formed between GLU 70^{Axl}:ARG 414^{Gas6} which is followed by ASP 73^{Axl} :ARG 313^{Gas6}, GLU 83^{Axl} :ARG 467^{Gas6}, GLU 59^{Axl} :ARG 310^{Gas6} and ARG 48 ^{Axl} :ASP 455^{Gas6}, respectively. The energy of only consistent salt-bridge in Ig1^{Axl}:LG1^{Pros1} , ARG 59^{Axl} :LYS 314^{Pros1} was found to be -101 kcal/mol (Table 7).

Table 6: The pairwise salt bridge (vdW+electrostatics) energies across $Ig1^{Axl}$:$LG1^{Gas6}$

Axl	Gas6	Average Energies (kcal/mol)
70	414	-82.4 +- 24.1
73	313	-72.1 +- 8.1
83	467	-70.6 +- 26.4
59	310	-67.6 +- 25.3
48	455	-59.6 +- 12.3

Table 7: The pairwise salt bridge (vdW+electrostatics) energies across $Ig1^{Axl}$:$LG1^{Pros1}$.

Axl	Pros	Average Energies (kcal/mol)
59	314	-101.1 + 9.3

As we determined the energies of the critical salt-bridges, we investigated where these energies are located across the whole pairwise interaction energies of $Ig1^{Axl}$:$LG1^{Gas6}$. Therefore, we calculated the whole pairwise energies across 350 snapshots from four parallel simulations of $Ig1^{Axl}$:$LG1^{Gas6}$ (total 1400 snapshots). Then we only considered the energetically favorable ones: negative energies (vdW+electrostatic). Figure 24 shows the distribution of these energies transformed into log scale. The histogram points out the identified salt-bridges, indicated as red dashed lines, are at the extremities of the histogram. Thus, these residues are among the residue pairs having lowest pairwise energies across the interaction landscape of $Ig1^{Axl}$:$LG1^{Gas6}$ simulations.

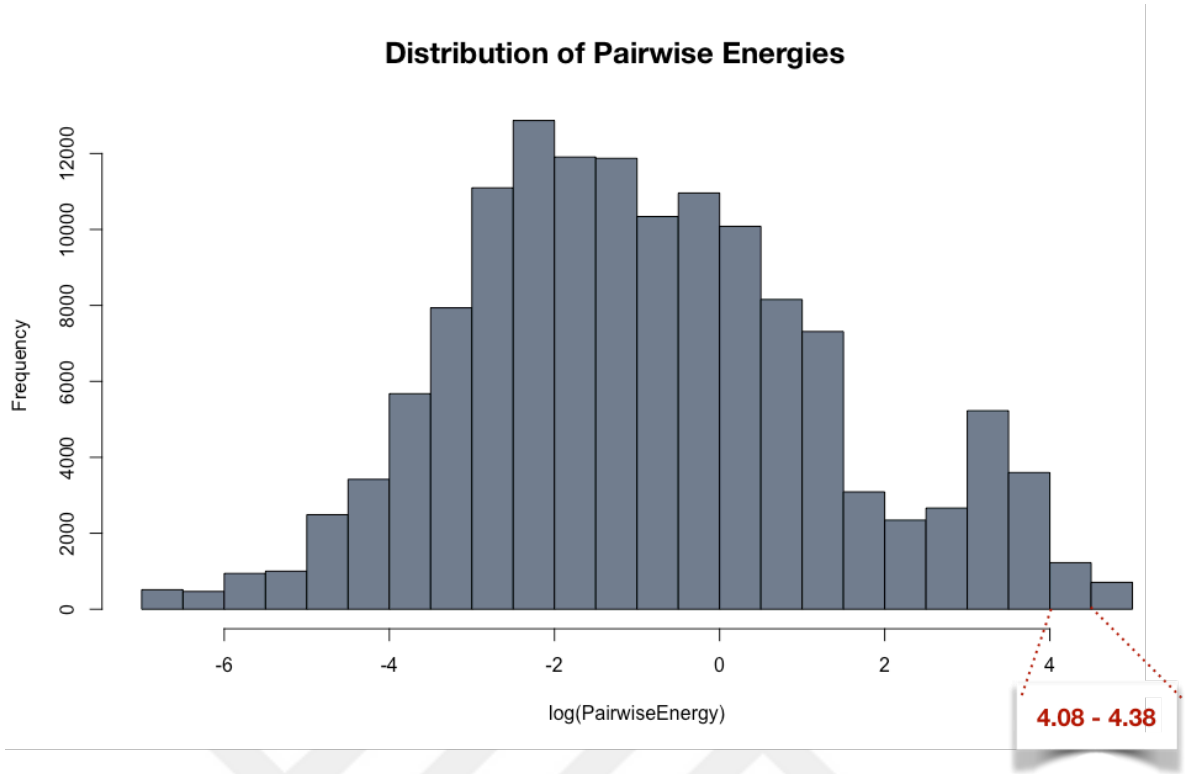


Figure 24: Distribution of negative pairwise energies in log scale for 2. The indicated red dashed lines denotes the consistent salt-bridge energy ranges for $Ig1^{Axl}$:$LG1^{Gas6}$ complex simulations.

4.5.7 Comparison of SDPs and their interactions among TAM Family, Gas6 and Pros1

If ARG 48^{Axl}, GLU 70^{Axl}, ASP 73^{Axl} and GLU 83^{Axl} are SDPs, they should be different at Tyro3 and Mer. As Tyro3 and Mer can bind the Pros1 with higher affinity, we expected the residues determining Axl's selectivity should be different on these positions in Tyro3 and Mer. Thus, we compared the tertiary structures of proteins (see Method 3.6.2.9) and mapped the sequences (see Method 3.6.2.2). Table 8 shows the corresponding $LG1^{Gas6}$ residues in $LG1^{Pros1}$ (left) and corresponding $Ig1^{Axl}$ residues in $Ig1^{Tyro3}$ and $Ig1^{Mer}$ (right). We see that all charged $LG1^{Gas6}$ residues are replaced by uncharged-polar or nonpolar residues in $LG1^{Pros1}$. As expected, one of the possible SDP, GLU 70^{Axl} is replaced by uncharged polar GLN at $Ig1^{Tyro3}$ and nonpolar LEU at $Ig1^{Mer}$. The other possible SDP, 48 ARG ^{Axl}, is replaced by uncharged-polar ASN both in $Ig1^{Tyro3}$ and $Ig1^{Mer}$. Residue GLU 83^{Axl}, found at the long loop region, has

no correspondence in short loop of Tyro3. However negative-charged of this position is conserved in $Ig1^{Mer}$. Lastly, ASP 73^{Axl} is conserved in $Ig1^{Tyro3}$ and it is replaced by positively charged HIS in $Ig1^{Mer}$. As a result, position 70 and position 48 are substituted to uncharged residues on Tyro3 and Mer. This result strengthens their possibility of being SDPs. However, for position 73 and 83, further analysis are required. Additionally, all identified four positions are completely conserved at the mouse Axl protein.

Table 8: Analogous $LG1^{Pros1}$ SDP-interacting residues to $LG1^{Gas6}$ and Analogous $Ig1^{Axl}$ SDP to $Ig1^{Tyro3}$ and $Ig1^{Mer}$, respectively.

PROS1	GAS6	AXL	TYRO3	MER
418 - ASN	414 - ARG	70 - GLU	85 - GLN	138 - LEU
317 - LEU	313 - ARG	73 - ASP	87 - ASP	141 - HIS
472 - ASN	467 - ARG	83 - GLU	--	151 - ASP
460 - ALA	455 - ASP	48 - ARG	63 - ASN	114 - ASN

As less conserved regions are most likely have role in subfamily specificity, we checked the conservation scores of the these residues via ConSurf (see Method 3.6.2.9). Table 9 shows Consurf's color code of possible SDPs and the residue variety at those positions. The color code ranges between 1 and 9 where 1 means less conserved and 9 means highly conserved residue. Here, the number of sequences used by ConSurf for multiple sequence alignment was 150. We found that residue 70^{Axl} and 48^{Axl} are less conserved than the other residues 73^{Axl} and 83^{Axl}.

Table 9: Conservation Scores of Possible Specificity Determining Positions

Residue	Score	Color	Residue Variety
70	0.897	2	R,H,P,V,Y,N,L,D,A,E,Q,S
73	-0.136	5	E,Q,I,D,T,S,M,H,N,V,G
83	-0.252	6	V,T,S,E,I,D,A
48	0.486	3	S,K,T,D,E,Q,G,H,V,N,L,F,R

4.5.9 Rim and Core Region Determination (2c5d)

As discussed in Introduction 2.3, the subfamily SDPs occur mostly at the rim region of protein. Therefore, we checked the core/rim annotation of possible SDPs (Table 10). To do that we used the crystal structure of Axl-Gas6 (PDB 2C5D) (see Method 3.6.2.9). We found that residue GLU 70^{Axl}, GLU 83^{Axl} and ARG 48^{Axl} are found in the rim region while ASP 73^{Axl} was found in the core region. In line with the previous conservation results, residue ASP 73^{Axl} also show high conservation (Table 9). This mean that it should have a role as a hot spot instead of being a SDP.

Table 10: Core/Rim annotation of Possible Specificity Determining Positions on Interface of Axl-Gas6 crystal structure (2c5d) calculated with EPPIC web-server , red: rim residues , black: core residues

Core/Rim	GAS6	AXL	Core/Rim
Rim	**414 - ARG**	**70 - GLU**	**Rim**
Rim	**313 - ARG**	**73 - ASP**	**Core**
Rim	**467 - ARG**	**83 - GLU**	**Rim**
Rim	**455 - ASP**	**48 - ARG**	**Rim**

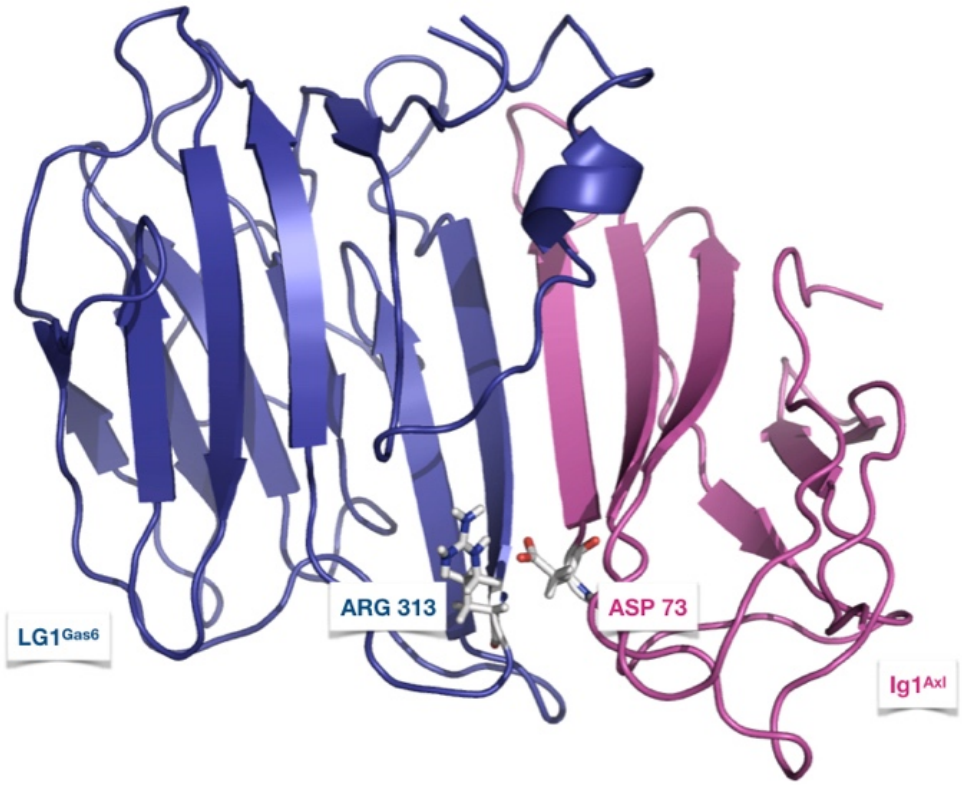


Figure 25: Location of core residue ASP 73^{Axl} and its partner ARG 313^{Gas6} on the $Ig1^{Axl}$:$LG1^{Gas6}$ interface.

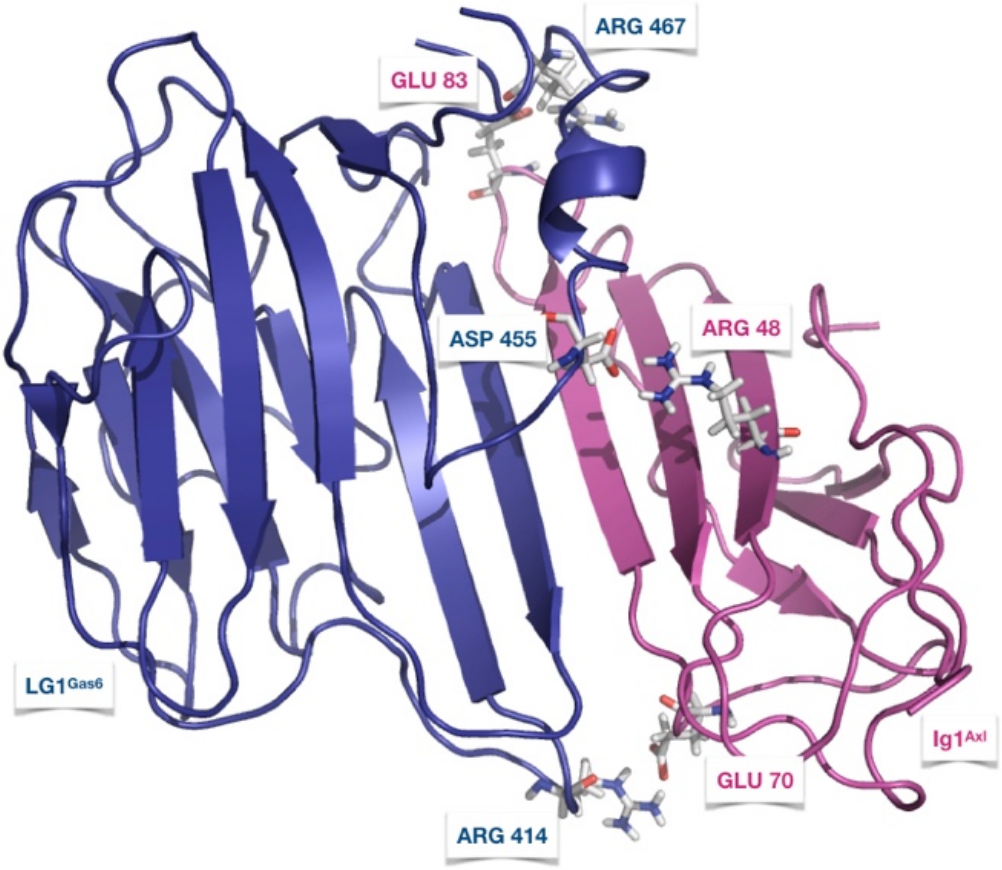


Figure 26: Location of rim residues ARG 48^{Axl}, GLU 70 ^{Axl} and GLU 83 ^{Axl} and residues across Ig1^{Axl}: ASP 455^{Gas6}, ARG 414^{Gas6}, ARG 467^{Gas6}.

5.DISCUSSION

TAM Family has a critical role from control of hemostasis via phagocytosis of apoptotic cells to the regulation of cell survival, proliferation and migration. Among TAM family, Axl is found at the center of many studies on targeted therapy against cancer. As Axl is essential for the cell's homeostasis and Ig^{Axl} has been previously designed as a soluble receptor for therapeutic purposes, deducing the molecular principals behind its ligand-induced activation can pay the way open for new rational design strategies. Expanding on this, in this thesis, we investigated Axl-ligand selectivity across two paralogous ligands Gas6 and Pros1 by using structure modeling approaches and understanding the context of all possible TAM-ligand interactions.

In line with this aim, we modelled the all possible 2:2 ($Ig1\text{-}Ig2^{Receptor}$:$LG1\text{-}LG2^{Ligand}$) and 1:1 ($Ig1^{Receptor}$:$LG1^{Ligand}$) TAM-ligand interactions. These modelling approaches indicated that the binding affinity profile of TAM family is encoded at its major receptor-ligand interface. This is in line with the literature where the major interface has been proposed to play a role in partner selectivity. Through the analysis of the constructed partners, we further discovered that Axl-Gas6 is electrostatically very compatible while Axl-Pros is not. Following this result, we focused on the Axl selectivity mechanism across the Gas6 (the best binder) and Pros1 (the worst binder.

Through the energetics analysis of our model structures, we identified critical residues on $Ig1^{Axl}$ via HADDOCK, which are ARG 48^{Axl}, GLU 56^{Axl}, GLU 59^{Axl}, GLU 70^{Axl}, ASP 73^{Axl} and GLU 83^{Axl}. Among those SDPs, it has been previously shown that $E59R^{Axl}$ mutation highly reduces the binding affinity across Gas6. Our analysis also indicates this residue's significance for Axl's binding to Gas6: it is the most energetically favorable residue among all residues of $Ig1^{Axl}$. To validate our findings and for deeper understanding of $Ig1^{Axl}$ selectivity across $LG1^{Gas6}$ and $LG1^{Pros1}$, we run molecular dynamics on $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$. MD analysis has showed that all four parallel $Ig1^{Axl}$:$LG1^{Gas6}$ simulations reach equilibration state very fast while none of parallel $Ig1^{Axl}$:$LG1^{Pros1}$ simulations come into equilibration. This result can be caused by two reasons. First, $Ig1^{Axl}$:$LG1^{Pros1}$ is the model structure that can require longer time to reach equilibration. The other reason can be based on the unstable characteristics of

$Ig1^{Axl}$:$LG1^{Pros1}$. This is also in line with Rg analysis because all parallel $Ig1^{Axl}$:$LG1^{Gas6}$ simulations show very compact and tight structure. On the other hand, only one of the parallel simulations of $Ig1^{Axl}$:$LG1^{Pros1}$ (Axl-Pros1-Parallel1) is very stable and behaves like $Ig1^{Axl}$:$LG1^{Gas6}$. The other three parallel simulations are very different from each other and fluctuate throughout the simulation. This randomness among parallel simulations is based on the dependency on the initial MD velocities.

Detailed analysis of interaction profile across trajectory of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ have showed that ¾ parallel $Ig1^{Axl}$:$LG1^{Pros1}$ simulations show lower hydrophobic interaction and hydrogen bond than $Ig1^{Axl}$:$LG1^{Gas6}$. Only the Axl-Pros1-Parallel1 shows very similar distribution with Axl-Gas6-Parallel simulations in terms of those kind of interactions. This result is also in line with the Rg results of the simulations: only Axl-Pros1-Parallel1 exhibit the similar (very stable) Rg value to $Ig1^{Axl}$:$LG1^{Gas6}$ simulations. The similarity between Axl-Pros1-Parallel1 and Axl-Gas6-Parallel simulations suggests that under certain conditions, like observed in in glioblastoma cells, Axl might interact nonspecifically with Pros1 through hydrophobic contacts and hydrogen bonds.

Investigation of another electrostatic-driven interaction type, salt-bridge, showed very different result than the hydrophobic and hydrogen bond analysis: the number of salt bridges distributes significantly differently in the case of $Ig1^{Axl}$:$LG1^{Gas6}$ simulations. Each $Ig1^{Axl}$:$LG1^{Gas6}$ simulation contains on average 4-5 fold more salt bridges than in $Ig1^{Axl}$:$LG1^{Pros1}$ simulations. This means that $Ig1^{Axl}$'s interface replete with charge residues and these residues cannot find a partner on $LG1^{Pros1}$. In this case, instead, they might make interactions with water molecule. This protein-water interactions might lead to destabilization of the $Ig1^{Axl}$'s structure.

Through the comparison of salt-bridges in these two complexes, we have found out that GLU 59^{Axl} is capable of making strong salt-bridges within two complexes. As mentioned before, the one mutation ($E59R^{Axl}$) results in decrease in binding affinity between Axl-Gas6. This might be based on the dissociation of the most strong salt-bridge between Axl and Gas6. More importantly, we have revealed four probable SDP residues on $Ig1^{Axl}$ which capable of making salt-bridges with $LG1^{Gas6}$ and not $LG1^{Pros1}$: GLU 70^{Axl}, GLU 83^{Axl}, ARG 48^{Axl} and GLU 73^{Axl} . Among these residues, GLU 70^{Axl}, GLU 83^{Axl} and GLU 73^{Axl} were also found as

probable SDPs with the use of HADDOCK-based analysis. Further comparative analysis with TAM Family, conservation analysis and positioning annotation analysis have yielded three critical SDP on $Ig1^{Axl}$ across $LG1^{Gas6}$: $GLU\ 70^{Axl}$, $ARG\ 48^{Axl}$ and $GLU\ 83^{Axl}$. They are the charged-rim residues and they completely lose their interactions with $LG1^{Pros1}$. Moreover, these SDPs are found to be substituted to nonpolar or uncharged residues on $Ig1^{Tyro3}$ and $Ig1^{Mer}$. This can be the also reason of how Tyro3 and Mer is able to bind Pros1. We have also identified an essential core residue which is $ASP\ 73^{Axl}$.

All in all, by using prominent computational structural biology techniques, we identified potential Axl SDPs that can select Gas6 over Pros1. We anticipate that these findings will pave the way for new therapeutic molecule design targeting Axl and Gas6 interface.

6. CONCLUSION AND SUGGESTIONS

- In this study, all possible TAM-Ligand interactions were modelled and their interface energetics were calculated. According to HADDOCK Score, major interface reflected the experimental binding affinity pattern of TAM-Ligand interactions. Essentially, it could discriminate the best ($Ig1^{Axl}$:$LG1^{Gas6}$) and worst ($Ig1^{Axl}$:$LG1^{Pros1}$) binder among TAM-Ligand complexes.
- According to HADDOCK individual terms, electrostatics were found to be the major factor that determines the $Ig1^{Axl}$ selectivity across $LG1^{Gas6}$ and $LG1^{Pros1}$.
- Comprehensive interaction analysis of $Ig1^{Axl}$:$LG1^{Gas6}$ and $Ig1^{Axl}$:$LG1^{Pros1}$ interface has showed that $Ig1^{Axl}$'s selectivity is driven by salt-bridges.
- Detailed investigation of consistent salt-bridges in each trajectory coupled with comparative and conservation analysis and position annotation of residues have provided us with three $Ig1^{Axl}$ SDPs: ARG 48^{Axl}, GLU 70^{Axl} and GLU 83^{Axl}.
- These charged residues are found on rim region of $Ig1^{Axl}$:$LG1^{Gas6}$ interface which fits with the literature claiming that sub-family specificity is directed by charged residues at rim region.
- These residues were also among the residues predicted as possible SDPs with HADDOCK energy analysis.
- We suggest that the simulation of all possible TAM Family-Ligand interactions should be run for further investigation and detailed analysis. This can provide deep understanding of Pros1 binding mechanism towards Tyro3 and Mer instead of Axl.
- The methodology developed in this thesis is given in Appendix D.2. It is general protocol to apply to other paralogous receptor-ligand pairs to deduce their SDPs.
- As we claim that identified SDPs are important for $Ig1^{Axl}$ selectivity, mutational experimental analysis can be applied to validate our findings: $Ig1^{Axl}$: $LG1^{Pros1}$ can be designed to be better binder by mutating three $LG1^{Pros1}$ residues to $LG1^{Gas6}$ counterparts which make salt-bridge with SDPs.

7. REFERENCES

1. Aiello, D. and Caffrey, D.R. Evolution of specific protein-protein interaction sites following gene duplication. *J Mol Biol* 2012;423(2):257-272.
2. Ashkenazy, H.*, et al.* ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. *Nucleic Acids Res* 2016;44(W1):W344-350.
3. Bellosta, P.*, et al.* The receptor tyrosine kinase ARK mediates cell aggregation by homophilic binding. *Mol Cell Biol* 1995;15(2):614-625.
4. Burchert, A.*, et al.* Determinants for transformation induced by the Axl receptor tyrosine kinase. *Oncogene* 1998;16(24):3177-3187.
5. Burley, S.K.*, et al.* RCSB Protein Data Bank: biological macromolecular structures enabling research and education in fundamental biology, biomedicine, biotechnology and energy. *Nucleic Acids Res* 2019;47(D1):D464-D474.
6. Caberoy, N.B.*, et al.* Galectin-3 is a new MerTK-specific eat-me signal. *J Cell Physiol* 2012;227(2):401-407.
7. Caberoy, N.B., Zhou, Y. and Li, W. Tubby and tubby-like protein 1 are new MerTK ligands for phagocytosis. *EMBO J* 2010;29(23):3898-3910.
8. Dolinsky, T.J.*, et al.* PDB2PQR: an automated pipeline for the setup of Poisson-Boltzmann electrostatics calculations. *Nucleic Acids Res* 2004;32(Web Server issue):W665-667.
9. Du, Z. and Lovly, C.M. Mechanisms of receptor tyrosine kinase activation in cancer. *Mol Cancer* 2018;17(1):58.
10. Geng, C.*, et al.* Finding the ΔΔG spot: Are predictors of binding affinity changes upon mutations in protein–protein interactions ready for it? *WIREs Comput Mol Sci.* 2019(e1410).
11. Glazer, D.S., Radmer, R.J. and Altman, R.B. Improving structure-based function prediction using molecular dynamics. *Structure* 2009;17(7):919-929.
12. Hafizi, S. and Dahlback, B. Gas6 and protein S. Vitamin K-dependent ligands for the Axl receptor tyrosine kinase subfamily. *FEBS J* 2006;273(23):5231-5244.
13. Hafizi, S. and Dahlback, B. Signalling and functional diversity within the Axl subfamily of receptor tyrosine kinases. *Cytokine Growth Factor Rev* 2006;17(4):295-304.
14. Heiring, C., Dahlback, B. and Muller, Y.A. Ligand recognition and homophilic interactions in Tyro3: structural insights into the Axl/Tyro3 receptor tyrosine kinase family. *J Biol Chem* 2004;279(8):6952-6958.
15. Ho, B.K. and Brasseur, R. The Ramachandran plots of glycine and pre-proline. *BMC Struct Biol* 2005;5:14.
16. Holland, S.J.*, et al.* R428, a selective small molecule inhibitor of Axl kinase, blocks tumor spread and prolongs survival in models of metastatic breast cancer. *Cancer Res* 2010;70(4):1544-1554.
17. Hubbard, S.R. and Miller, W.T. Receptor tyrosine kinases: mechanisms of activation and signaling. *Curr Opin Cell Biol* 2007;19(2):117-123.
18. Huey, M.G.*, et al.* Targeting the TAM Receptors in Leukemia. *Cancers (Basel)* 2016;8(11).
19. Ivanov, S.M.*, et al.* Protein-protein interactions in paralogues: Electrostatics modulates specificity on a conserved steric scaffold. *PLoS One* 2017;12(10):e0185928.

20. Jankauskaite, J.*, et al.* SKEMPI 2.0: an updated benchmark of changes in protein-protein binding energy, kinetics and thermodynamics upon mutation. *Bioinformatics* 2019;35(3):462-469.
21. Jorgensen, W.L.*, et al.* Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 1983;79(2):926-935.
22. Jurrus, E.*, et al.* Improvements to the APBS biomolecular solvation software suite. *Protein Sci* 2018;27(1):112-128.
23. Kariolis, M.S.*, et al.* Inhibition of the GAS6/AXL pathway augments the efficacy of chemotherapies. *J Clin Invest* 2017;127(1):183-198.
24. Kariolis, M.S.*, et al.* An engineered Axl 'decoy receptor' effectively silences the Gas6-Axl signaling axis. *Nat Chem Biol* 2014;10(11):977-983.
25. Kastritis, P.L. and Bonvin, A.M. On the binding affinity of macromolecular interactions: daring to ask why proteins interact. *J R Soc Interface* 2013;10(79):20120835.
26. Landau, M.*, et al.* ConSurf 2005: the projection of evolutionary conservation scores of residues on protein structures. *Nucleic Acids Res* 2005;33(Web Server issue):W299-302.
27. Lee-Sherick, A.B.*, et al.* Efficacy of a Mer and Flt3 tyrosine kinase small molecule inhibitor, UNC1666, in acute myeloid leukemia. *Oncotarget* 2015;6(9):6722-6736.
28. Lemke, G. Biology of the TAM receptors. *Cold Spring Harb Perspect Biol* 2013;5(11):a009076.
29. Lemke, G. and Rothlin, C.V. Immunobiology of the TAM receptors. *Nat Rev Immunol* 2008;8(5):327-336.
30. Lemmon, M.A. and Schlessinger, J. Cell signaling by receptor tyrosine kinases. *Cell* 2010;141(7):1117-1134.
31. Li, W.*, et al.* The EMBL-EBI bioinformatics web and programmatic tools framework. *Nucleic Acids Res* 2015;43(W1):W580-584.
32. Likhachev, I.V., Balabaev, N.K. and Galzitskaya, O.V. Available Instruments for Analyzing Molecular Dynamics Trajectories. *Open Biochem J* 2016;10:1-11.
33. Lindorff-Larsen, K.*, et al.* Improved side-chain torsion potentials for the Amber ff99SB protein force field. *Proteins* 2010;78(8):1950-1958.
34. Linger, R.M.*, et al.* TAM receptor tyrosine kinases: biologic functions, signaling, and potential therapeutic targeting in human cancer. *Adv Cancer Res* 2008;100:35-83.
35. Lu, Q.*, et al.* Tyro-3 family receptors are essential regulators of mammalian spermatogenesis. *Nature* 1999;398(6729):723-728.
36. Margareto, J.*, et al.* DNA copy number variation and gene expression analyses reveal the implication of specific oncogenes and genes in GBM. *Cancer Invest* 2009;27(5):541-548.
37. Margreitter, C. and Oostenbrink, C. MDplot: Visualise Molecular Dynamics. *R J* 2017;9(1):164-186.
38. McDonell, L.M.*, et al.* Receptor tyrosine kinase mutations in developmental syndromes and cancer: two sides of the same coin. *Hum Mol Genet* 2015;24(R1):R60-66.
39. Morris, A.L.*, et al.* Stereochemical quality of protein structure coordinates. *Proteins* 1992;12(4):345-364.
40. Needleman, S.B. and Wunsch, C.D. A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J Mol Biol* 1970;48(3):443-453.
41. Paolino, M. and Penninger, J.M. The Role of TAM Family Receptors in Immune Cell Function: Implications for Cancer Therapy. *Cancers (Basel)* 2016;8(10).

42. Paul, M.K. and Mukhopadhyay, A.K. Tyrosine kinase - Role and significance in Cancer. *Int J Med Sci* 2004;1(2):101-115.
43. R Development Core Team. 2017. R: A Language and Environment for Statistical Computing. https://www.R-project.org/
44. Rothlin, C.V. and Lemke, G. TAM receptor signaling and autoimmune disease. *Curr Opin Immunol* 2010;22(6):740-746.
45. Sadahiro, H., *et al.* Activation of the Receptor Tyrosine Kinase AXL Regulates the Immune Microenvironment in Glioblastoma. *Cancer Res* 2018;78(11):3002-3013.
46. Sarkar, D. Lattice: Multivariate Data Visualization with R. Springer; 2008.
47. Sasaki, T., *et al.* Crystal structure of a C-terminal fragment of growth arrest-specific protein Gas6. Receptor tyrosine kinase activation by laminin G-like domains. *J Biol Chem* 2002;277(46):44164-44170.
48. Sasaki, T., *et al.* Structural basis for Gas6-Axl signalling. *EMBO J* 2006;25(1):80-87.
49. Schrödinger. 2010. The PyMOL Molecular Graphics System, Version 2.0 Schrödinger, LLC
50. Studer, R.A., *et al.* Understanding the functional difference between growth arrest-specific protein 6 and protein S: an evolutionary approach. *Open Biol* 2014;4(10).
51. Tai, K.Y., *et al.* Axl promotes cell invasion by inducing MMP-9 activity through activation of NF-kappaB and Brg-1. *Oncogene* 2008;27(29):4044-4055.
52. UniProt, C. UniProt: a worldwide hub of protein knowledge. *Nucleic Acids Res* 2019;47(D1):D506-D515.
53. van der Meer, J.H., van der Poll, T. and van 't Veer, C. TAM receptors, Gas6, and protein S: roles in inflammation and hemostasis. *Blood* 2014;123(16):2460-2469.
54. Van Der Spoel, D., *et al.* GROMACS: fast, flexible, and free. *J Comput Chem* 2005;26(16):1701-1718.
55. van Zundert, G.C.P., *et al.* The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes. *J Mol Biol* 2016;428(4):720-725.
56. Vishwanath, S., *et al.* Specificity and stability of transient protein-protein interactions. *Curr Opin Struct Biol* 2017;44:77-86.
57. Wium, M., Paccez, J.D. and Zerbini, L.F. The Dual Role of TAM Receptors in Autoimmune Diseases and Cancer: An Overview. *Cells* 2018;7(10).
58. Wu, G., *et al.* Targeting Gas6/TAM in cancer cells and tumor microenvironment. *Mol Cancer* 2018;17(1):20.
59. Yanagihashi, Y., *et al.* Mouse macrophages show different requirements for phosphatidylserine receptor Tim4 in efferocytosis. *Proc Natl Acad Sci U S A* 2017;114(33):8800-8805.
60. Ye, Y. and Godzik, A. Flexible structure alignment by chaining aligned fragment pairs allowing twists. *Bioinformatics* 2003;19 Suppl 2:ii246-255.
61. Zhang, Y. I-TASSER server for protein 3D structure prediction. *BMC Bioinformatics* 2008;9:40.
62. Zwier MC and Chond LT. Reaching biological timescales with all-atom molecular dynamics simulations. *Curr Opin Pharmacol,* 2010;10(6):745-52.

APPENDIX A.

The following codes were used to run molecular dynamics simulation via GROMACS 5.1.4 at TÜBİTAK ULAKBİM High Performance and Grid Computing Center (TRUBA).

1st slurm codes:

```
#!/bin/bash
#SBATCH -p akya-cuda
#SBATCH -A tkarakulak
#SBATCH -J akyacuda-axlgas6
#SBATCH -N 1
#SBATCH -n 40
#SBATCH --gres=gpu:1
#SBATCH --time=1:00:00
#SBATCH --mail-type=ALL
#SBATCH --mail-user=tulay.karakulak@ibg.edu.tr
#SBATCH --output=slurm-%j.out
#SBATCH --error=slurm-%j.err

source /etc/profile.d/modules.sh
module load centos7.3/app/gromacs/5.1.4-openmpi-1.8.8-gcc-4.8.5-E5V3-cuda
module load  centos7.3/lib/cuda/7.5

module load centos7.3/lib/openmpi/1.8.8-gcc-4.8.5
module load centos7.3/lib/acml/6.1.0-gfortan64

export OMP_NUM_THREADS=20
echo "SLURM_NODELIST $SLURM_NODELIST"
echo "NUMBER OF CORES $SLURM_NTASKS"

$GROMACS_DIR/bin/gmx_mpi pdb2gmx -f peptide.pdb -o peptide.gro -p peptide.top -ignh -ter <<EOD
6
1
EOD
$GROMACS_DIR/bin/gmx_mpi editconf -f peptide.gro -o peptide.pdb
$GROMACS_DIR/bin/gmx_mpi editconf -f peptide.gro -o peptide-PBC.gro -bt dodecahedron -c -d 1.4
$GROMACS_DIR/bin/gmx_mpi grompp -v -f 01_em_vac_PME.mdp -c peptide-PBC.gro -p peptide.top -o peptide-EM-vacuum.tpr
date
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-EM-vacuum -pin on
date
$GROMACS_DIR/bin/gmx_mpi solvate -cp peptide-EM-vacuum.gro -cs spc216.gro -o peptide-water.gro -p peptide.top
$GROMACS_DIR/bin/gmx_mpi grompp -v -f 02_em_sol_PME.mdp -c peptide-water.gro -p peptide.top -o peptide-water.tpr
$GROMACS_DIR/bin/gmx_mpi genion -s peptide-water.tpr -o peptide-solvated.gro -conc 0.15 -neutral -pname NA+ -nname CL- <<EOD
13
EOD

exit
```

2nd Slurm Codes to use after topology editing:

```
#!/bin/bash
#SBATCH -p akya-cuda
#SBATCH -A tkarakulak
#SBATCH -J akya-cuda-axlgas6
#SBATCH -N 1
#SBATCH -n 40
#SBATCH --gres=gpu:1
#SBATCH --time=1:00:00
#SBATCH --mail-type=ALL
#SBATCH --mail-user=tulay.karakulak@ibg.edu.tr
#SBATCH --output=slurm-%j.out
#SBATCH --error=slurm-%j.err

source /etc/profile.d/modules.sh
module load centos7.3/app/gromacs/5.1.4-openmpi-1.8.8-gcc-4.8.5-E5V3-cuda
module load  centos7.3/lib/cuda/7.5

module load centos7.3/lib/openmpi/1.8.8-gcc-4.8.5
module load centos7.3/lib/acml/6.1.0-gfortan64

export OMP_NUM_THREADS=20
echo "SLURM_NODELIST $SLURM_NODELIST"
echo "NUMBER OF CORES $SLURM_NTASKS"

$GROMACS_DIR/bin/gmx_mpi grompp -v -f 02_em_sol_PME.mdp -c peptide-solvated.gro -p peptide.top -o peptide-EM-solvated.tpr
date "echo 02 start"
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-EM-solvated -pin on
date "echo 02 end"
$GROMACS_DIR/bin/gmx_mpi grompp -v -f 03_nvt_pr1000_PME.mdp -c peptide-EM-solvated.gro -p peptide.top -o peptide-NVT-PR1000.tpr
date "echo 03 start"
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-NVT-PR1000 -pin on
date "echo 03 end"
$GROMACS_DIR/bin/gmx_mpi grompp -v -f 04_npt_pr_PME.mdp -c peptide-NVT-PR1000.gro -p peptide.top -o peptide-NPT-PR1000.tpr
date "echo 04 start"
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-NPT-PR1000 -pin on
date "echo 04 end"

sed -i -e 's/1000  1000  1000/ 100   100   100/g' posre_Protein_chain_A.itp
sed -i -e 's/1000  1000  1000/ 100   100   100/g' posre_Protein_chain_B.itp
sed -i -e 's/1000  1000  1000/ 100   100   100/g' posre_Protein_chain_C.itp
sed -i -e 's/1000  1000  1000/ 100   100   100/g' posre_Protein_chain_D.itp

$GROMACS_DIR/bin/gmx_mpi grompp -v -f 04_npt_pr_PME.mdp -c peptide-NPT-PR1000.gro -p peptide.top -o peptide-NPT-PR100.tpr
date "echo 04-2 start"
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-NPT-PR100 -pin on
date "echo 04-2 end"

sed -i -e 's/100   100   100/ 10    10    10/g' posre_Protein_chain_A.itp
sed -i -e 's/100   100   100/ 10    10    10/g' posre_Protein_chain_B.itp
sed -i -e 's/100   100   100/ 10    10    10/g' posre_Protein_chain_C.itp
```

```
sed -i -e 's/100   100   100/ 10    10    10/g' posre_Protein_chain_D.itp

$GROMACS_DIR/bin/gmx_mpi grompp -v -f 04_npt_pr_PME.mdp -c peptide-NPT-PR100.gro -p peptide.top -o peptide-NPT-PR10.tpr
date "echo 04-3 start"
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-NPT-PR10 -pin on
date "echo 04-3 end"

$GROMACS_DIR/bin/gmx_mpi grompp -v -f 05_npt_NOpr_PME.mdp -c peptide-NPT-PR10.gro -p peptide.top -o peptide-NPT-noPR.tpr
date "echo 05 start"
$GROMACS_DIR/bin/gmx_mpi mdrun -v -deffnm peptide-NPT-noPR -pin on
date "echo 05 end"

exit
```

3rd Slurm File for Production Run:

```
#!/bin/bash
#SBATCH -p akya-cuda
#SBATCH -A tkarakulak
#SBATCH -J axlgas6_akya_parallel2
#SBATCH -N 1
#SBATCH -n 40
#SBATCH --gres=gpu:1
#SBATCH --time=10-00:00:00
#SBATCH --mail-type=ALL
#SBATCH --mail-user=tulay.karakulak@ibg.edu.tr
#SBATCH --output=slurm-%j.out
#SBATCH --error=slurm-%j.err

source /etc/profile.d/modules.sh
module load centos7.3/app/gromacs/5.1.4-openmpi-1.8.8-gcc-4.8.5-E5V3-cuda
module load  centos7.3/lib/cuda/7.5

module load centos7.3/lib/openmpi/1.8.8-gcc-4.8.5
module load centos7.3/lib/acml/6.1.0-gfortan64

export OMP_NUM_THREADS=20
echo "SLURM_NODELIST $SLURM_NODELIST"
echo "NUMBER OF CORES $SLURM_NTASKS"

$GROMACS_DIR/bin/gmx_mpi grompp -v -f 06_md_PME.mdp -c peptide-NPT-noPR.gro -p peptide.top -o md_0_20.tpr
$GROMACS_DIR/bin/mdrun_mpi -v -deffnm md_0_20 -pin on

exit
```

APPENDIX B.

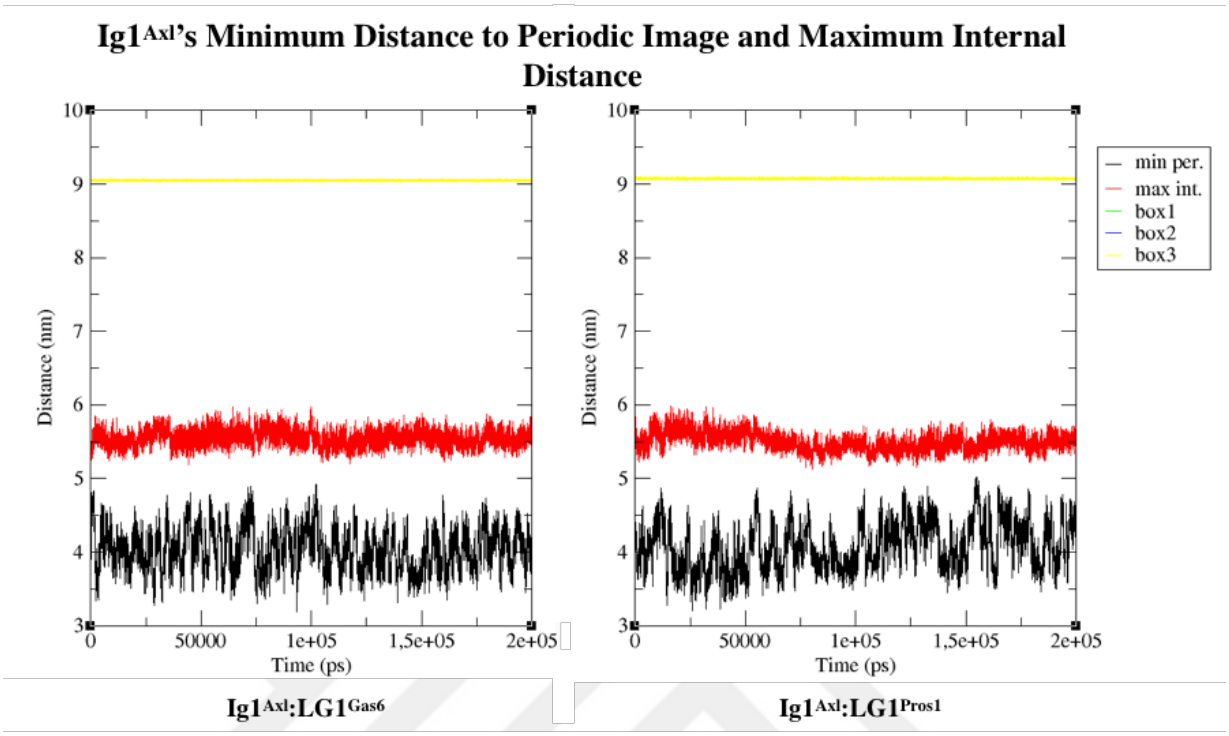


Figure B.1 $Ig1^{Axl}$'s minimum distance to periodic image and maximum internal distance in $Ig1^{Axl}$:$LG1^{Gas6}$ (left) and $Ig1^{Axl}$:$LG1^{Pros1}$ (right)complexes. The values belongs to Axl-Gas6-Parallel1 and Axl-Pros1-Parallel1 simulations. Plot was generated with xmgrace command.

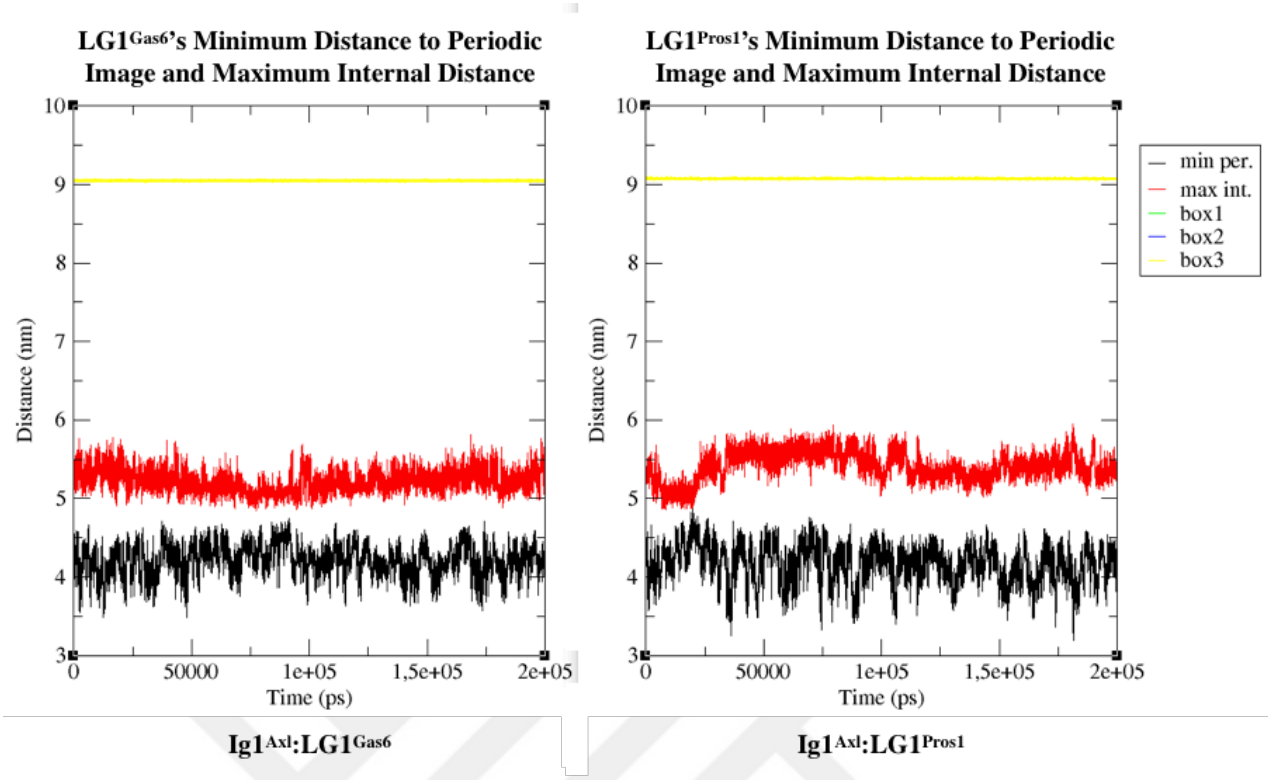


Figure B.2 LG1^{Gas6}'s and LG1^{Pros1}'s minimum distance to periodic image and maximum internal distance in Ig1^{Axl}:LG1^{Gas6} (left) and Ig1^{Axl}:LG1^{Pros1} (right) simulations. The values belongs to Axl-Gas6-Parallel1 and Axl-Pros1-Parallel1 simulations. Plot was generated with *xmgrace* command.

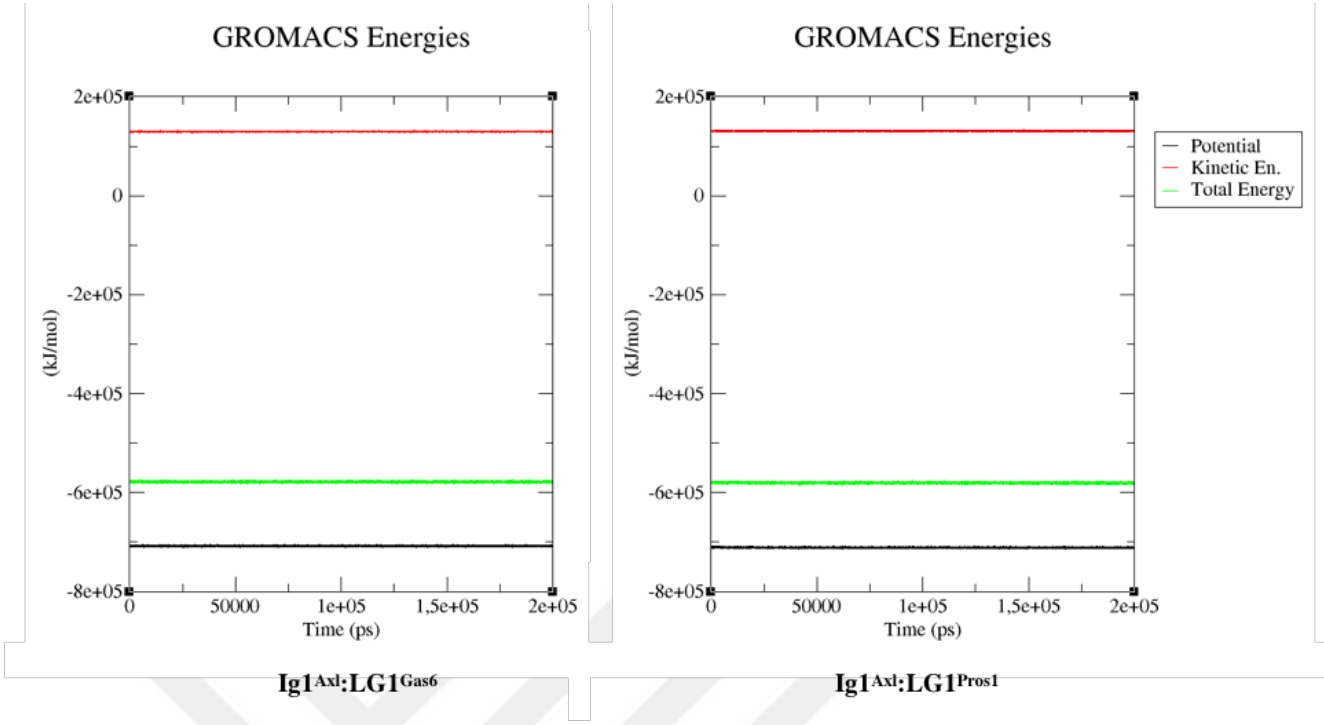


Figure B.3. Potential, kinetic and total energy over the 200 ns trajectory in $Ig1^{Axl}$:$LG1^{Gas6}$ (left) and $Ig1^{Axl}$:$LG1^{Pros1}$ (right) simulations. The values belongs to Axl-Gas6-Parallel1 and Axl-Pros1-Parallel1 simulations. Plot was generated with *xmgrace* command.

APPENDIX C.

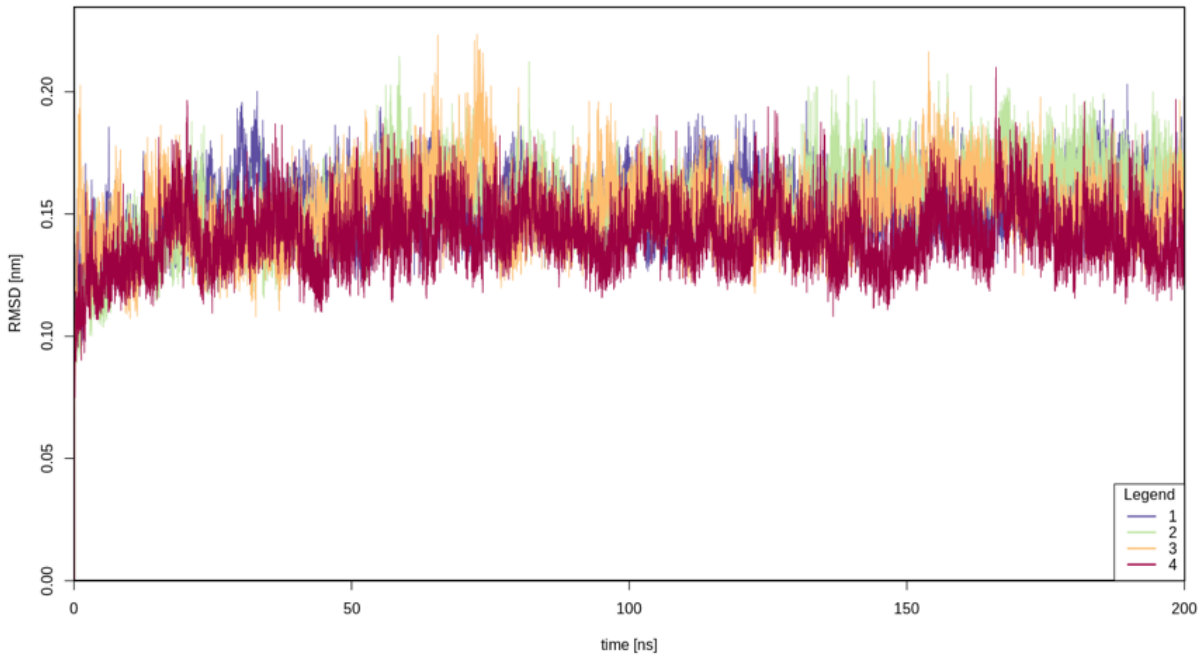


Figure C.1: Backbone atoms' RMSD values of four $Ig1^{Axl}$:$LG1^{Gas6}$ simulation over 200 ns trajectory superimposed onto first structure.

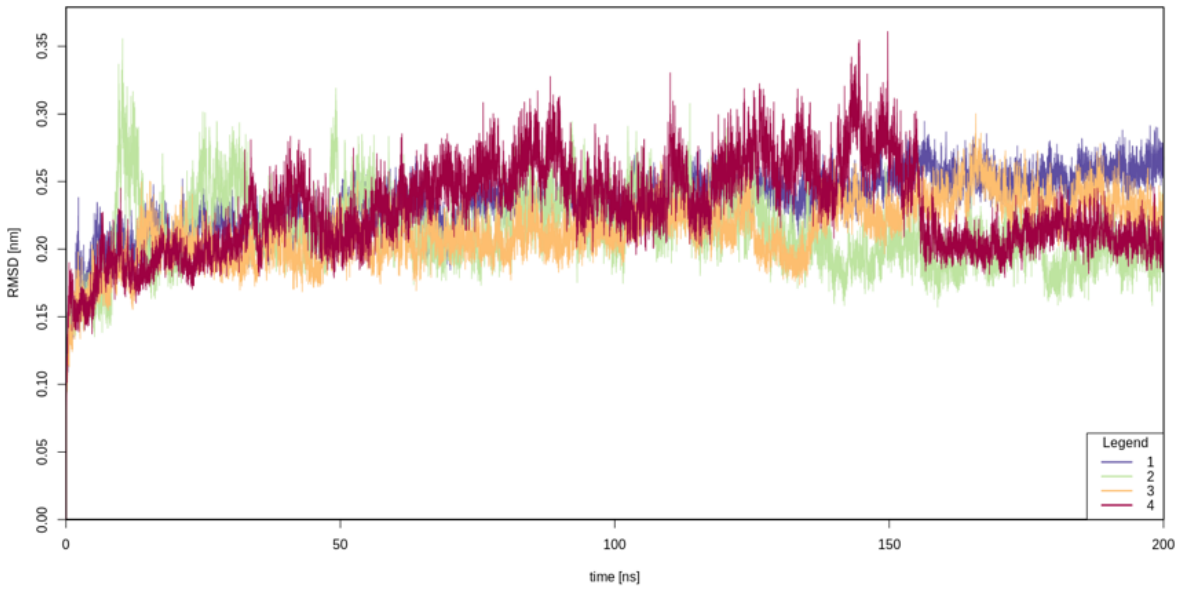


Figure C.2: Backbone atoms' RMSD values of four $Ig1^{Axl}$:$LG1^{Pros1}$ simulation over 200 ns trajectory superimposed onto first structure.

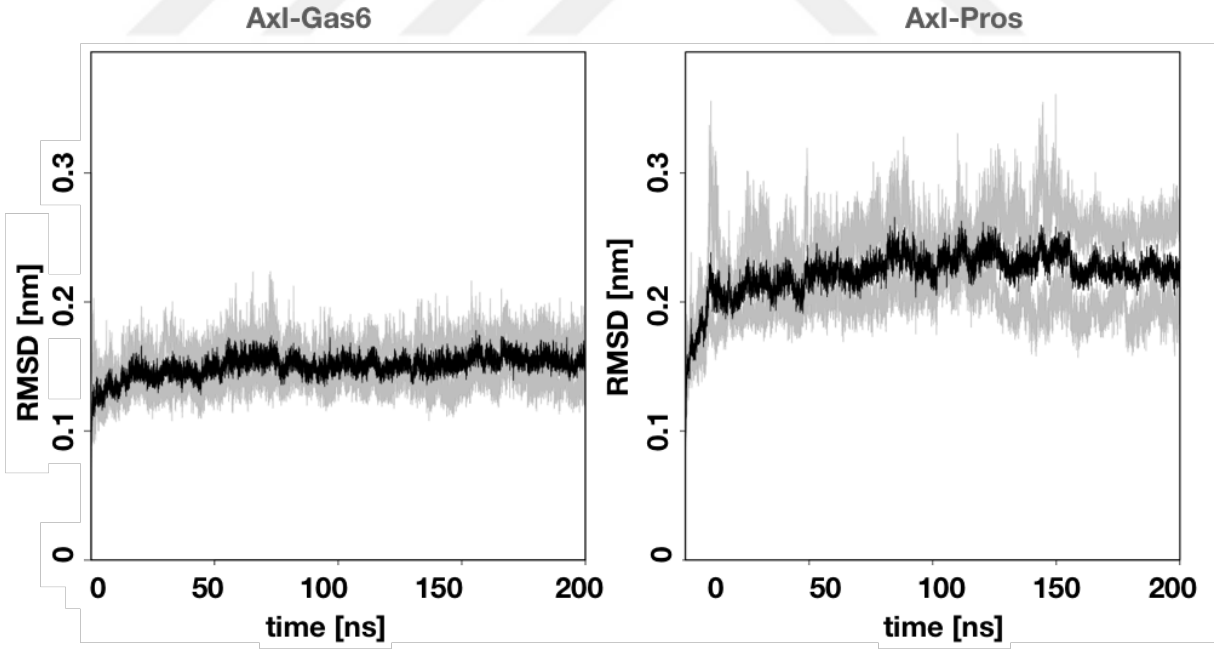


Figure C.3: Average backbone atoms' RMSD values of four $Ig1^{Axl}$:$LG1^{Gas6}$ simulation over 200 ns trajectory superimposed onto first structure. The black lines shows the average values, the gray lines represents the minimum and maximum values at each ns among the four simulations.

APPENDIX D.

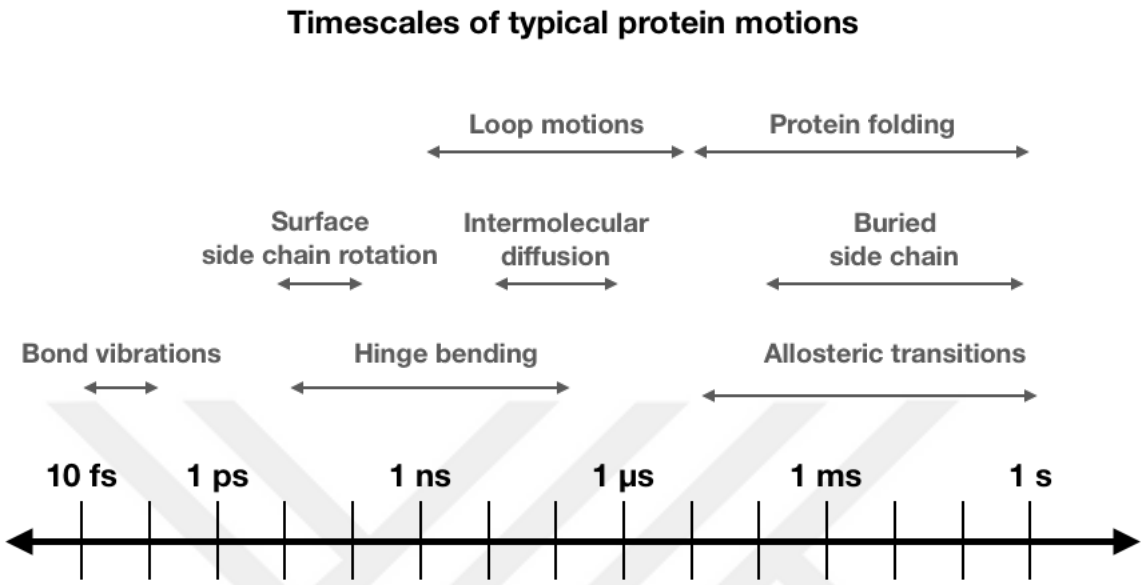


Figure D.1: The timescale in molecular dynamics simulation and required time on a typical desktop computer (adapted from (Zwier MC and Chond LT, 2010)).

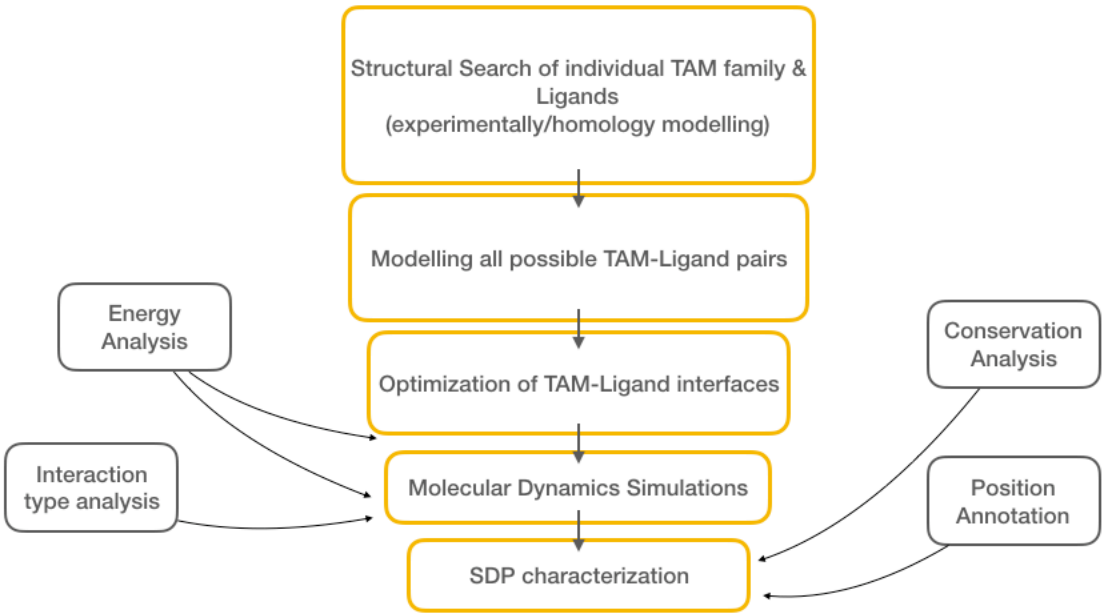


Figure D.2: The workflow of SDP characterization protocol that has been developed in this thesis.

CV

TÜLAY KARAKULAK

Personal Information

Contact Information

Identity Number	24991896008
Date of Birth	05/11/1992
Contact Address	
Telephone	
Email	tulay.karakulak@ibg.edu.tr
Web Address	

Education Information

01 February 2017 - Now (2 Year 4 Month)
Yüksek Lisans, Tezli Program, DOKUZ EYLÜL ÜNİVERSİTESİ, TÜRKİYE
İZMİR ULUSLARARASI BİYOTIP VE GENOM ENSTİTÜSÜ, GENOM BİLİMLERİ VE MOLEKÜLER BİYOTEKNOLOJİ ANABİLİM DALI
Cumulative Grade Point Average: 4.0 / 4.0

01 September 2011 - 01 June 2016 (4 Year 10 Month)
Lisans, Anadal/Normal Öğretim, İZMİR YÜKSEK TEKNOLOJİ ENSTİTÜSÜ, TÜRKİYE
FEN FAKÜLTESİ, MOLEKÜLER BİYOLOJİ VE GENETİK BÖLÜMÜ
Cumulative Grade Point Average: 3.3 / 4.0

Experience Information

01 April 2018 - Now (1 Year 2 Month) (Yarı Zamanlı)
ARAŞTIRMACI, İZMİR BİYOTIP VE GENOM MERKEZİ

Foreign Language Information

İNGİLİZCE (Reading: İyi, Writing: İyi, Speaking: İyi)

Scientific Technological Activity Fields

Keywords

Biyoinformatik
Hesaplamalı Biyoloji
Yapısal Biyoenformatik
Moleküler Dinamik Simülasyonu

R&D Competency

Conference Papers

T. KARAKULAK & E. KARACA EREK, Molecular Specificity Of Tam Family Receptors Towards Different Ligands, Sözlü Sunum, The International Symposium On Health Informatics And Bioinformatics Hibit, 25 October 2018, 27 October 2018.

T. KARAKULAK & E. KARACA EREK, Molecular Specificity Of Tam Family Receptors Towards Different Ligands, Poster Sunumu, Integrative Modelling Of Biomolecular Interactions, 02 July 2018, 06 July 2018.

TÜBİTAK Scholarships and Fundings

BIDEB Supports

TÜLAY KARAKULAK, Etkinlik Destekleri ve Eğitim Bursları Müdürlüğü, 2210-A Genel Yurt İçi Yüksek Lisans Burs Programı, Başvurusu Reddedildi, 2017 - 1.

Number of Panelist/Observer/Reporter Jobs

Panelist/Outer Consultant Count	ARDEB/BİDEB 0	TEYDEB 0	Total 0
Number of Observer/Consultant Jobs	ARDEB/BİDEB 0	TEYDEB 0	Total 0
Number of Reporter Jobs	ARDEB/BİDEB 0	TEYDEB 0	Total 0