ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

STRUCTURAL ANALYSIS OF ARTHRITOGENIC PEPTIDE AND HLA-B2705 INTERACTIONS FOR THE IMMUNE PATHOGENESIS OF ANKYLOSING SPONDYLITIS WITH MOLECULAR SIMULATIONS

SENA KIVRAK

M.Sc. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR
Assoc. Prof. Günseli Bayram Akçapınar

ISTANBUL-2022

ACIBADEM MEHMET ALI AYDINLAR UNIVERSITY
INSTITUTE OF HEALTH SCIENCES

STRUCTURAL ANALYSIS OF ARTHRITOGENIC PEPTIDE AND HLA-B2705 INTERACTIONS FOR THE IMMUNE PATHOGENESIS OF ANKYLOSING SPONDYLITIS WITH MOLECULAR SIMULATIONS

SENA KIVRAK

M.Sc. THESIS

DEPARTMENT OF BIOSTATISTICS AND BIOINFORMATICS

SUPERVISOR

Assoc. Prof. Günseli Bayram Akçapınar

ISTANBUL-2022

DECLARATION

I declare that this thesis work is my own work, I had no unethical behavior at any stages from the planning to the writing of the thesis, I obtained all the information in this thesis in accordance with academic and ethical rules, I cited all the information and comments that were not obtained with this thesis work, and I provided resources in the list of references. I also declare that there was no violation of any patents and copyrights during the study and writing of this thesis.

02/08/2022

Sena Kıvrak

PREFACE AND ACKNOWLEDGEMENT

First of all, I would like to express my deepest gratitude to my supervisor Assoc. Prof. Günseli Bayram Akçapınar who made this thesis possible and for always being kind to me. She was always there for helping me to reduce my stress with her support and guidance in the course of my thesis. I am really happy and proud to be her student.

I would also thank to my thesis committee, Prof. Uğur Sezerman and Asst. Prof. Fatma Tuba Gözet for spending their precious time and for their evaluations.

My best friend İrem, thank you for being by my side since my undergraduate years, for being the person I can talk to and discuss with when I have a problem, and for finding solutions when I deal with challenges in my thesis.

My loved one Mimar, thank you for always being there for me when I wanted to give up and when I was depressed, and for increasing my faith in myself by constantly reminding me that I can do it.

Finally, I would like to thank my family, my mother, my father, and my brother for being my family and for letting me know that they always have my back. I am very glad for their endless love.

The molecular dynamics simulations in the thesis were performed at TUBITAK ULAKBIM, High-Performance Grid Computing Center (TRUBA resources).

TABLE OF CONTENTS

LIST OF ABBREVIATIONS AND SYMBOLS

Å	Angstrom
A	Alpha
AIR	Ambiguous Interaction Restraints
β	Beta
β2m	Beta – microglobulin
APBS	Adaptive Poisson-Boltzmann Solver
AS	Ankylosing Spondylitis
CDR	Complementarity Determining Region
Cryo-EM	Cryogenic Electronic Microscopy
CTL	Cytotoxic T Lymphocytes
ER	Endoplasmic Reticulum
ERAP	Endoplasmic Reticulum Aminopeptidase
ERp57	Endoplasmic Reticulum Protein 57
HADDOCK	High Ambiguity Driven biomolecular DOCKing
HLA	Human Leukocyte Antigen
IEDB	Immune Epitope Database
BiP	Immunoglobulin Binding Protein
MD	Molecular Dynamics
MHC	Major Histocompatibility Complex
NAMD	Nanoscale Molecular Dynamics
NMR	Nuclear Magnetic Resonance
Ω	Omega
PME	Particle mesh Ewald
PBMC	Peripheral Blood Mononuclear Cells
PDB	Protein Data Bank
PLC	Peptide-loading Complex
pMHC	Peptide-Major Histocompatibility Complex
RMSD	Root Mean Square Deviations
RMSF	Root Mean Square Fluctuations
SF	Synovial Fluid

3D	Three-dimensional
TAD	Torsion Angle Molecular Dynamics
TAP	Transporter Associated with Antigen Processing
TCR	T-cell Receptor
UPR	Unfolded Protein Response
VMD	Visual Molecular Dynamics

LIST OF FIGURES

LIST OF TABLES

ÖZET

Ankilozan Spondilitin İmmün Patogenezi için Moleküler Simülasyonlarla Artritojenik Peptid ve HLA-B*27:05 Etkileşimlerinin Yapısal Analizi

Ankilozan spondilit (AS), eksenel iskelette enflamasyona neden olan otoimmün bir hastalıktır. AS ve HLA (İnsan Lökosit Antijen)-B27 arasındaki ilişkiyi gösteren çalışmalar mevcuttur. Buna rağmen, bu ilişki arkasındaki kesin mekanizma hala bilinmemektedir. HLA-B27, immün yanıtlarda antijen sunumunda rol oynar. *Self*-antijeni *non-self*-antijenden ayırt etmek için antijeni sitotoksik CD+8 T hücrelerine sunar. HLA-B*27:05, AS ile ilişkisi en çok bilinen alt tiptir. HLA-B*27:05 ve HLA-B*27:09 sadece bir amino asitte farklılık göstermesine rağmen HLA-B*27:05 AS patogenezinde rol oynarken HLA-B*27:09 ilişkili değildir. Odak noktası *self*-peptid ve *non-self*-peptidin arasındaki moleküler benzerlik olan artritojenik peptid teorisi, HLA-B27 ve AS arasındaki ilişkiyi açıklamak için yaygın hipotezlerden biridir. Bu benzerlik, *self*-peptidler üzerinde CD8+ hücresinin immün cevabıyla sonuçlanır. Bugüne kadar HLA-B27 ve artritojenik peptidlerin ilişkisini açıklayan yapısal bir çalışma bulunmamaktadır. Bu tezde, iki farklı HLA allotipinin AS patogenezi üzerindeki etkilerini anlamak üzere, daha önce tanımlanmış iki artritojenik *self*-peptid ile insan immün yetmezlik virüsüne (HIV) ait KK10 peptidi kullanılarak peptid sunumları moleküler yanaştırma ve moleküler dinamik (MD) yöntemleriyle incelendi. MD simülasyonları sonucunda, *self*-peptid ve *non-self* peptidin orta pozisyonlarında bulunan amino asitlerin benzerlik derecesi yerine bu peptidlerin terminallerinin etrafındaki yüklü amino asitlerin benzerliğinin, HLA üzerindeki sunum davranışına daha fazla katkısı olduğu gösterildi. Bu sonuç *self*-peptidlerin ve *non-self*-peptidlerin moleküler taklitçiliği üzerine olan artritojenik peptid hipotezini desteklemektedir.

Anahtar Sözcükler: Ankilozan spondilit, HLA-B*27:05, Artritojenik peptidler, Moleküler dinamik, Moleküler yanaştırma

ABSTRACT

Structural Analysis of Arthritogenic Peptide and HLA-B*27:05 Interactions for the Immune Pathogenesis of Ankylosing Spondylitis with Molecular Simulations

Ankylosing spondylitis (AS) is an autoimmune disease that causes inflammation in the axial skeleton. There are studies that have discovered the relationship between AS and HLA (Human Leukocyte Antigen)-B27. However, the exact mechanism behind the association is still unknown. HLA-B27 is involved in antigen presentation during immune responses. It presents the antigen to cytotoxic CD+8 T cells to distinguish self-antigen from non-self-antigen. HLA-B*27:05 is the most known subtype for the association with AS. Although HLA-B*27:05 and HLA-B*27:09 differ in only one amino acid, HLA-B*27:05 is implicated in the pathogenesis of AS, and HLA-B*27:09 is not. Arthritogenic peptide theory based on the molecular similarity between self-peptide and non-self-peptide is one of the common hypotheses to explain the association between HLA-B27 and AS. This similarity results in cross-recognition of CD8+ cells on self-peptides. To date, there is no structural study on the interaction of HLA-B27 and arthritogenic peptides. In this thesis, the differences between peptide presentations of HLA allotypes were studied by using two previously characterized arthritogenic self-peptides and the KK10 peptide of the human immunodeficiency virus (HIV) through molecular docking and molecular dynamics (MD) methods to understand their effect on AS pathogenesis. The similarity of charged amino acids around the termini of the self-peptides to non-self-peptide, rather than the similarity of amino acids in the middle positions, was found to contribute more to the presentation behavior on HLA-B*27:05 as observed in MD simulations in agreement with the arthritogenic peptide hypothesis on molecular mimicry of self-peptides and non-self-peptides.

Keywords: Ankylosing spondylitis, HLA-B*27:05, Arthritogenic peptides, Molecular dynamics**,** Molecular docking

1 INTRODUCTION AND AIM

Autoimmune diseases occur when the immune system attacks its body tissues (1). Ankylosing spondylitis (AS) is an autoimmune disease that causes an inflammatory response affecting the axial skeleton, especially the spinal cord (2). As the disease progresses, the joints are damaged due to inflammation and the movement of the patient is restricted, making life difficult. Although the incidence of AS varies from population to population, AS affects up to 6% of people worldwide (2). This disease is three times more common in men than in women, and younger people are also at higher risk of developing the disease. Genetic predisposition and environmental factors are thought to have an effect on triggering the disease. HLA-B27 is the most well-known gene associated with AS. The complex background and lack of sufficient information on the exact molecular mechanism of AS, has an impact on the drug design process and the implementation of more effective treatments. Unfortunately, current treatments and medications can only reduce the pain in the course of the disease (3).

Although one of the genes most associated with AS is HLA-B27, its exact role in the pathogenesis of the disease is still unknown. The protein encoded from this gene is involved in the presentation of the antigen to the immune system as a major histocompatibility complex class I (MHC-I). Depending on whether the presented antigen is self or non-self, the CD8+ T cell generates an immune response. Different hypotheses have been formulated to explain the reasons behind the relationship between AS and HLA-B27 such as the heavy chain homodimer hypothesis, the misfolding hypothesis, and the arthritogenic peptide hypothesis (4). This thesis is focused on the arthritogenic peptide hypothesis based on cross-reactivity in CD8+ T cells due to the molecular similarity of non-self-peptide and self-peptide. This hypothesis has been strengthened by studies about the cytotoxic T lymphocytes (CTL) response to self-peptides (5–7). Fully elucidating the relationship between HLA-B27 and AS remains a challenge due to the complexity of the antigen presentation mechanism including a plethora of molecules.

HLA-B27 includes several identified subtypes including polymorphism based on the differences between amino acid sequences. While some subtypes are associated with AS disease, some subtypes have been observed to be unrelated. An amino acid change at position 116 of HLA-B27 causes HLA-B*27:05 (Asp116) and HLA-B*27:09 (His116) to be differently associated with AS disease (8). Therefore, the above-mentioned difference between HLA-B*27:05 and HLA-B*27:09 during arthritogenic self-peptide presentation forms a basis for structural understanding of the disease pathogenesis within the framework of this thesis (9).

Dynamic structure of HLA-B27, its peptide binding specificity because of polymorphism and the complexity of antigen presentation mechanism and immune system cause difficulties in understanding the pathogenesis of AS disease (10–12). Although there are many studies trying to unravel the relationship between AS and HLA-B27, no structural study has been found with peptides derived from AS patients. There are many unresolved parts due to the polymorphism of MHC, but studies are leading the way in understanding the antigen presentation mechanism (10). The structural differences in the antigen presentation mechanism between HLA-B*27:05 and HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL) derived from Atagündüz's study (2005) and the KK10 (KRWIILGLNK) peptide from HIV are analyzed in this thesis with an aim to decipher the molecular mechanisms which contribute to the pathogenesis of AS disease with a specific focus on understanding the interactions between MHC class I allele, HLA-B27 and arthritogenic peptides to be presented during the antigen presentation. (6,10).

2 BACKGROUND

2.1 Ankylosing Spondylitis

Spondyloarthritis is a group of diseases resulting in pain and inflammation in a joint such as ankylosing spondylitis, reactive arthritis, and psoriatic arthritis. Ankylosing spondylitis (AS) is an autoimmune disease that causes inflammation of the axial skeleton, especially joints in the spinal cord. AS usually begins in the joint between the spine and pelvis which is sacroiliac joint. Therefore, low back pain is usually the first symptom shown in AS disease. The disease that starts in the sacroiliac joint can progress to the spine and affect it. The inflammation between vertebrae fuses the spine, and the spine becomes a fixed bone structure restricting movement (3). This reduced mobility results in problems with body functions such as breathing and makes life difficult for the patient (13).

Although the prevalence of the disease varies according to ethnic groups, it is between 0.1% and 1.4%. The risk of developing the disease is usually at younger ages and symptoms begin to appear in the 20s. AS is three times as common in men as in women and the disease shows a slow course in women. (3). The reason of AS is still a mystery. Nevertheless, ERAP1, ERAP2, and HLA-B27 are the most common genes related to AS disease and they are the most crucial factors for AS to occur (14). The association between HLA-B27 and AS was first reported in 1973 (15). However, the reason behind this association is still unknown (16). Genetic studies have illustrated that the risk of developing AS in an HLA-B27 positive child of a parent with HLA-B27 positive AS is approximately 20% (14). Also, 90-95% of AS patients are HLA-B27 positive. However, this doesn’t mean all positive cases will show AS. Besides, they are predominantly healthy people (3).

2.2 Major Histocompatibility Complex Class I

Major histocompatibility complex (MHC), known as human leukocyte antigen (HLA) in the human body, is a protein found on the cell surface and it is responsible for presentation of antigen fragments obtained from pathogens or self-body tissues (10,17). If the one to be presented is a peptide derived from a pathogen, macrophages kill the cell, and B cells provide antibodies to identify pathogens. Also, there are two important points about MHC that they have polygenicity and polymorphism. The location of MHC is on chromosome 6 and it has more than 200 genes (10). There are three classes of MHC: MHC class I (HLA-A, HLA-B, HLA-C), MHC class II, and MHC class III. They have a diversity of peptide-binding specificities because of different MHC class genes (10,14). In addition, the polymorphism of MHC class genes result in different subtypes of MHC classes within the populations (10). MHC class I has a role in immune peptide presentation, and it is located on the cell surface of all nucleated cells. It presents peptide fragments to cytotoxic CD8+ T cells to distinguish self and non-self-antigen and create an immune response (14). Thus, healthy host cells and invaded cells are separated from each other. A problem in antigen presentation may create other problems for the correct identification of antigen fragments during the immune response.

2.2.1 Structure of MHC Class I molecules

The structure of an MHC class I molecule contains two proteins which are α and β2m (β2 – microglobulin) and a peptide (Figure 1a). α1, α2, and α3 subunits create the heavy chain of the MHC class I molecule (14). There are almost four hundred amino acids for α and β subunits of an MHC. The MHC includes three disulfide bonds which are between residue 109-164 and residue 203-259 for heavy chain (α2 and α3) of MHC and residue 1-99 for β2 – microglobulin of MHC (Figure 1a) (21). α1 (residues 57-84) and α2 (residues 138-180) create a groove to be able to bind a peptide (antigen) (14,18). Most of the polymorphisms of MHC class I are especially located within these two α domains (11). The binding groove consists of two long α-helical walls and a floor of four β-sheets and includes the first 180 residues of MHC.

The macromolecular structure of MHC from its subunits is formed in the endoplasmic reticulum (ER) of cells (19). Then, the MHC-peptide complex passes through the Golgi apparatus to reach the surface of the cell for antigen presentation to the T-cell receptor (TCR) of T lymphocytes (14). TCR consists of two protein chains that are α and β including variable and constant regions (Figure 1b) (20). The variable region is composed of six loops which are called complementarity determining region (CDR) to recognize the peptide on MHC (21,22). In an MD simulation study, Tsuchiya *et al*. used five different T-cells with different levels of cross-reactivity and TCR-pMHC complexes with the same MHC class II-peptide. It is observed that these five T-cells had a similar binding affinity, but the binding site of the TCR differed. Also, six loops of C121C TCR had connections with HLA-B∗2705-KK10, especially the CDR2α, CDR3α, and CDR3β loops. The CDR2α loop was the only region of the TCR that contacted the α2 helix of HLA-B*2705, but all six loops were in contact with the α1 helix (23).

Six pockets, named as A to F, have been identified in the MHC as the binding sites of the antigen commonly for nonapeptides (Figure 1c). Thus, the peptidome of the HLA with high polymorphism is diversified according to each subtype with the change of residues in these pockets contributing to the development of MHC-induced diseases. The diversity of MHC provides the specificity of peptides binding to TCR (12). According to many studies, the B and F pockets usually bond with the antigen and play a significant role in maintaining the stability of the HLA. F pocket accommodates the C-terminus of the peptides (24,25). Crystallography studies and liquid chromatography (LC) mass spectrometry (MS) study suggested that Arg residue is preferable at position 2 of peptide bound to B pocket of HLA (9,26–28). Also, MD simulation study based on the sole difference in F pocket between HLA-B*27:05 and HLA-B*27:09 showed the importance of the difference in the flexibility of F pocket which affects the stability of HLA-B*27:05 negatively when it is unbound to a peptide (25). The stability of pMHC is an essential factor so that it can reach the surface of the cell and present the peptide (11). The chaperon proteins in the ER help to stabilize the heavy chain of HLA until the heavy chain and β2-microglobulin domains form the MHC. Studies show that in the absence of β2m, the

heavy chain can form a homodimer structure, especially by forming a disulfide bond between the Cys67 amino acids found in α1 domain (14).

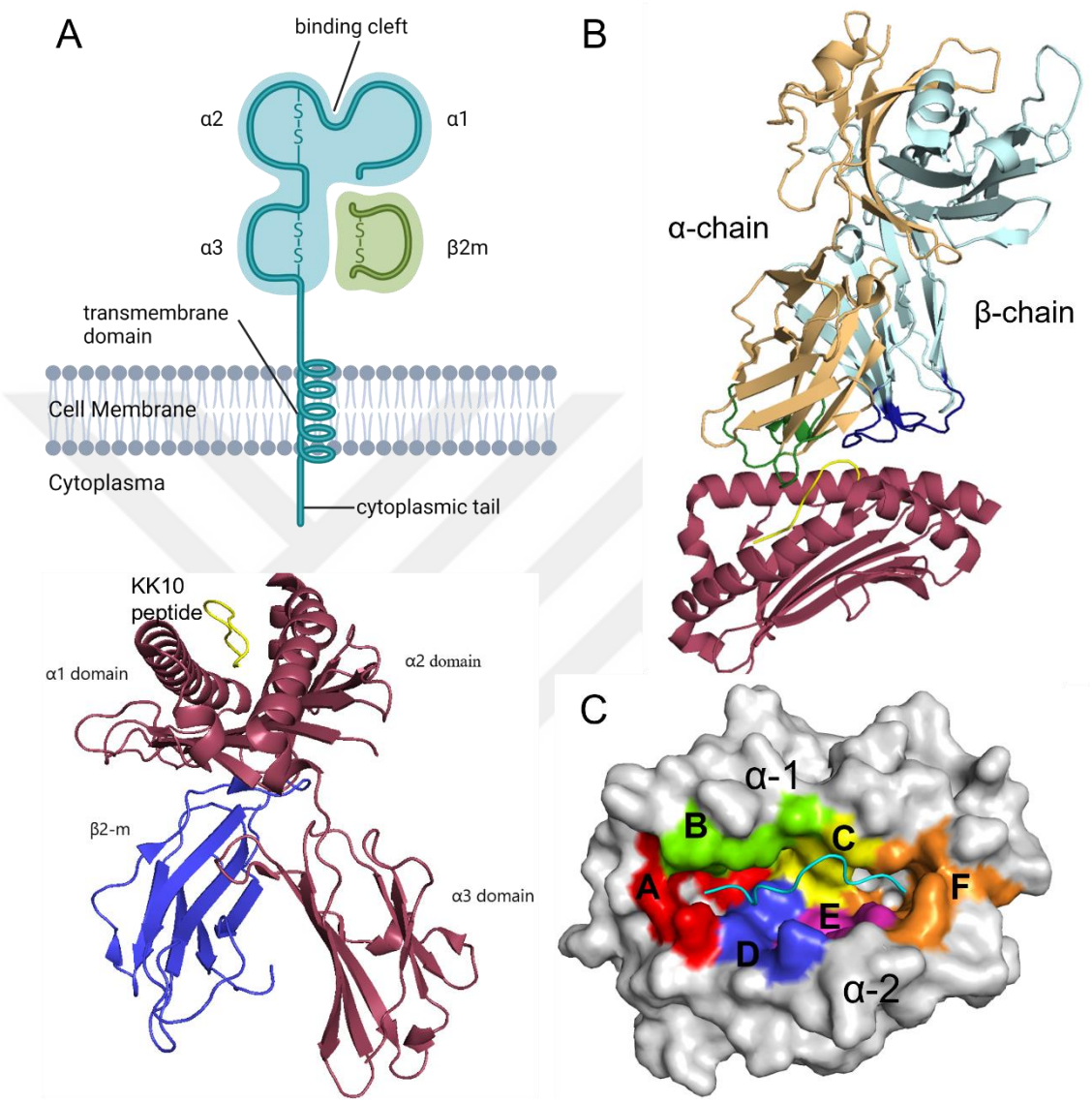

Figure 1. The structure of MHC class I complex bound and unbound to TCR and the binding pockets of MHC. (a) The graphical illustration and secondary structure of MHC class I complex. MHC class I includes three disulfide bonds in residue 101-164 and residue 203-259 for heavy chain (α1, α2, and α3) and residue 25-80 for β2-microglobulin. Created with BioRender.com. (b) The binding groove of MHC class I bound to TCR. CDR loops are shown in green and dark blue, while α and β chains of TCR are shown in light orange and light cyan, respectively. (c) The binding pockets of HLA.

The length of peptides to be presented is usually 8-10 amino acids (29). However, recent studies have showed that there are longer peptides to 11-14 amino acids to present to T cells (9,29,30). When the peptides are N- or C-terminally extended, structural studies have stated that most of the peptides prefer a shape of central bulging/canonical B27 binding rather than terminal protrusion to bind MHC for recognition of TCR (9,30,31). Mass spectrometry study of the peptides bound to HLA-B27 subtypes shows that C-terminally extended peptides are bound to HLA-B27 stronger than N-terminally extended peptides (16). Additionally, hydrophobic C-termini trimming is more efficient than charged C-termini trimming by ERAP1 (17). As a result, HLA prefers charged C-termini peptides and binds to positively charged peptides (13).

2.2.2 Presentation pathway

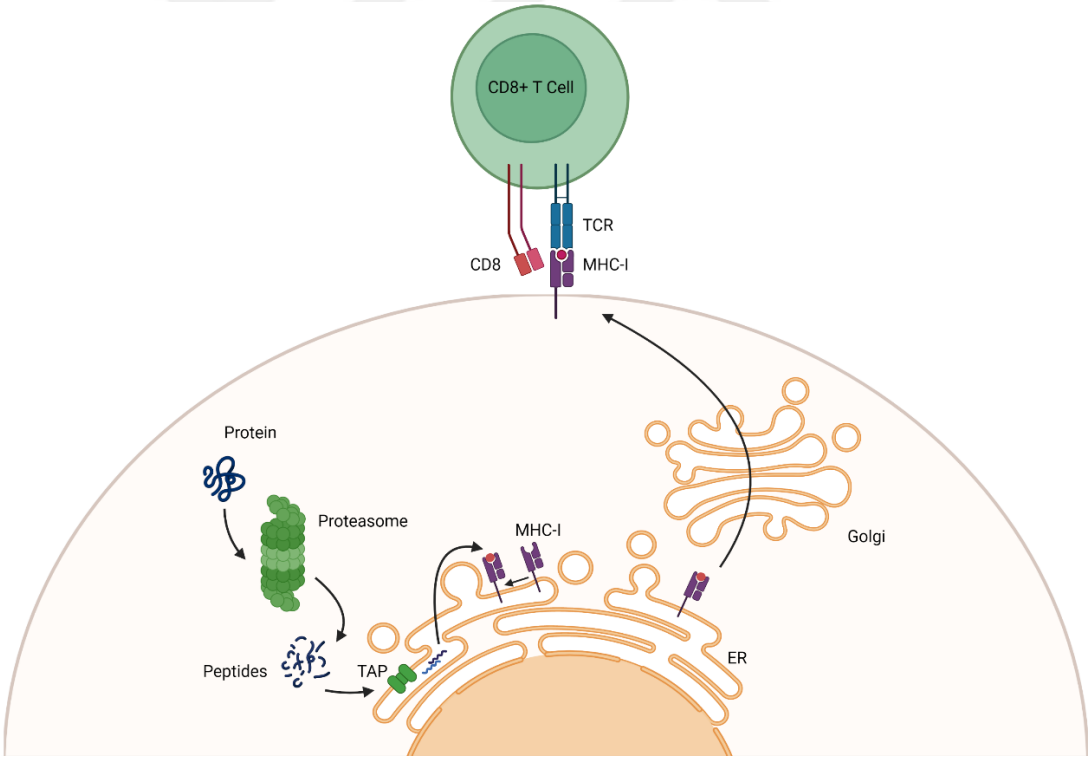

Figure 2. Antigen processing and MHC class I antigen presentation pathway. Created with BioRender.com.

Antigen processing is a pathway that prepares an antigen to be presented to the immune cells (T lymphocytes) for immune responses (Figure 2). In this process, a nucleated cell is invaded by pathogens such as viruses. Viruses produce their protein in the host cell. The viral protein is degraded into peptide fragments by the proteasome, and they are added to self-peptide pool derived from host cell proteins. The fragments are 8-16 residues in length, and they are transported to the ER lumen by ATP-binding cassette subfamily B member (TAP) protein (14). A viral peptide is loaded on MHC I by the peptide-loading complex (PLC) including MHC I heavy chain, β2m, two chaperones calnexin and calreticulin, TAP proteins (TAP1 and TAP2), and ERp57 (ER protein 57). The chaperones and ERp57 help to stabilize the heavy chain of the MHC I to prevent aggregation. Tapasin is found between MHC I and TAP protein and acts as a bridge allowing the peptide to bind to MHC (32).

In addition, endoplasmic reticulum aminopeptidase (ERAP1/ERAP2) proteins in ER trim the fragments into peptides of 8-9 amino acids in length to bind to MHC class I proteins (14). Endoplasmic reticulum aminopeptidase 1 (ERAP1) has a crucial role in peptide trimming that is essential to cleave long peptides into peptides with optimal length for MHC class I molecules (33). When a peptide fits into the binding groove of the protein, the assembly of the PLC dissociates, and MHC I protein leaves the ER with the peptide (32). It goes through the Golgi apparatus in a transport vehicle and the peptide is presented on the cell surface to the receptor of the T lymphocyte (14). TCR of the T lymphocyte recognize the peptides presented by MHC class I molecules. T lymphocytes also have a CD8 receptor that recognizes the α3 of the MHC class I molecule. Therefore, T lymphocytes bind to MHC class I on the host cell and the peptide (29). Then, the antigen is recognized as non-self-antigen which results in killing of the cell. If the peptides are produced from the host cell's protein, they are recognized as self-antigen. This whole landscape is called the MHC class I antigen presentation pathway (14).

2.3 Role of MHC in AS

2.3.1 Hypotheses

HLA-B27 is the most known MHC class I molecule related to AS disease. The distinctive features of HLA-B27 from other MHC molecules such as a tendency to misfold and peptide binding specificity have led to the emergence of some hypotheses regarding its relationship with AS (Figure 3) (34). These hypotheses try to explain the association between HLA-B27 and AS. However, the mechanism behind this association is not clear yet. From the hypotheses, the heavy chain homodimer theory and the misfolding hypothesis include the IL-23/IL-17 pathway. While the α3 domain of HLA-B27 folds fast, the α1 and α2 domains fold slower until they are bound to peptide. Immunoglobulin binding protein (BiP) which is ER chaperone is responsible for protein folding. The long-time association of BiP and HLA-B27 causes to the misfolding of HLA-B27. The misfolding of MHC class I affects ER function changing the physiological condition of ER. This situation creates stress on the ER. This stress affects the length and number of peptides (4,35). Also, this induces the unfolded protein response (UPR). Also, HLA-B27 presents the antigen on the surface but it tends to unfold resulting in the dimerization of HLA-B27. The protein-peptide complex is recognized by CD4+ cells. Unfolding of HLA-B27 causes the increase in the production of cytokines (IL-23), and this results in chronic inflammation causing the disease (34,36). Moreover, cysteine residue at the position 67 of MHC class I is prone to create a disulfide bond between heavy chains of MHC I facilitating the dimerization of MHC class I (16,37). The formation of the dimers affects the differential relation of HLA subtypes with AS (14). The two hypotheses do not include the peptide presentation function. In conjunction with the latter hypothesis, the β2m deposition hypothesis is based on the light chain of HLA. The dimerization of MHC class I gives rise to the accumulation of β2m monomers, especially in the synovia. The accumulation affects the synthesis of proteins that have a role in tissue destruction giving rise to AS (14,38).

Lastly, the arthritogenic peptide hypothesis states that the immune response occurs for specific peptide-HLA complexes (36). It is one of the most focused hypotheses for the association. According to the hypothesis, self-peptide mimics the non-self-peptide. When HLA-B27 presents the foreign antigen to cytotoxic T lymphocytes (CTLs), the CTL response occurs. Because of molecular similarity between self-peptide and the arthritogenic peptide, CD8+ cell shows cross-recognition on self-peptides. This results in over-reaction and the initiation of tissue damage (4,9).

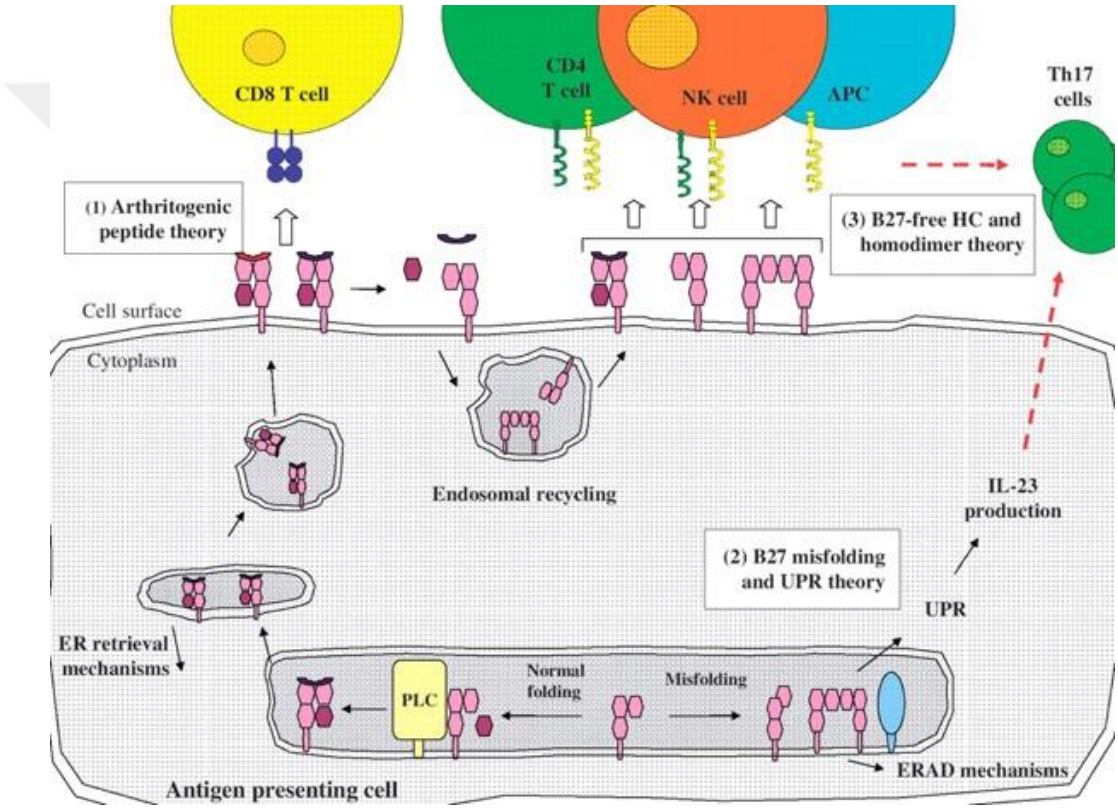

Figure 3. The most known hypotheses to explain the association between HLA-B27 and AS (4). (*Copyright*: 5305570127919)

Purcell and colleagues (2015) studied on self-peptides from eight HLA-B27 subtypes which are differentially associated with AS disease to define similarities and differences between them based on the arthritogenic peptide hypothesis by checking qualitative change of peptides between HLA-B27 subtypes which are related to AS distinctly and using high-resolution MS, but the results did not show

enough evidence for qualitative change of peptides. In addition, the consensus-binding motif of HLA-B*27:05 peptidome was considered (Table 1) (9). Then, for the next study (2016), they tried to identify candidate arthritogenic peptides presented in a small amount in HLA-B27:06 and HLA-B27:09 which are non-associated with the disease compared to disease-associated HLA-B27 allotypes which resulted in 26 putative arthritogenic peptides. Quantitative change is a more effective way to define similarities between HLA-B27 peptidome (30). Another study supporting the arthritogenic peptide theory is an X-ray crystallography study of complexes of three self-peptides and one viral peptide, which exhibited molecular mimicry as a result, with five HLA-B27 subtypes which were HLA-B*27:03, HLA-B*27:04, HLA-B*27:05, HLA-B*27:06, and HLA-B*27:09. However, these self-peptides are not arthritogenic peptides and are from cellular proteins (37). In Atagündüz *et al.* (2005) study, two candidate arthritogenic self-peptides that are ARGQPGVMG and DRASFIKNL were found by checking T-cell activation of selected peptides, by antigen-specific intracellular cytokine staining, using peripheral blood mononuclear cells (PBMC) and synovial fluid (SF) mononuclear cells of 27 HLA-B27 positive AS patients. They also used peripheral blood samples from five HLA-B27 positive healthy controls and five HLA-B27 negative AS patients, and SF samples from three HLA-B27 negative rheumatoid arthritis patients as control samples. Two computer programs based on MHC class I binding motifs were used to predict epitope and proteosome of HLA-B27. 18 cartilage proteins were identified from the human 20S proteosome, and 121 nonamer peptides were selected using a score criteria. ARGQPGVMG and DRASFIKNL were derived from type II collagen and type VI collagen, respectively. ARGQPGVMG and DRASFIKNL stimulated peripheral blood CD8 + T cells for only 1 patient (out of 20) and DRASFIKNL stimulated only SF CD8 + T cells for 4 patients (out of 7). Only these two peptides were stimulatory for T cells (6). So far, there is no combination for the structural study of HLA-B27 and arthritogenic peptides from AS patients. Mostly, the peptides derived from viruses or self-peptides from cell lines have been used to understand the association in MD simulation studies (25,39–43). Therefore, the two candidate arthritogenic self-peptides derived from AS patients from the study of Atagündüz (2005) were used in the thesis, focusing on the arthritogenic peptide hypothesis.

Table 1. The consensus-binding motif of peptides bound to the HLA-B*27:05 subtype, modified and redrawn from Purcell *et al.* study (2015) (9).

Consensus Binding Motif of HLA-B*27:05												
Nonamers												
	Amino acid positions											
Observation of the identified peptides	1	2	3	4	5	6	7	8	9			
Dominant (>30%)		R										
Strong (>20– ≤30%)	G		L						F			
Preferred (>10–≤20%)	S, A		F, V, I	P, L	L, I	I, V, L	L, V	G, S	L, R, Y			
Decamers												
	Amino acid positions											
Observation of the identified peptides	1	2	3	4	5	6	7	8	9	10		
Dominant (>30%)		R										
Strong (>20– ≤30%)	G		F						G	F		
Preferred (>10–≤20%)	R, K		L, I	D, G	L, G, P	L	G	L, I	I	Y, K, L, R		
Undecamers												
	Amino acid positions											
Observation of the identified peptides	1	2	3	4	5	6	7	8	9	10	11	
Dominant (>30%)		R										
Strong (>20– ≤30%)			F, L		P						K, F	
Preferred (>10–≤20%)	G, A, R, S		I	G, L, S	L	S, L	G	G, L	L, I	T, S, K, G, L	Y, R, L	
Dodecamers												
	Amino acid positions											
Observation of the identified peptides	1	2	3	4	5	6	7	8	9	10	11	12
Dominant (>30%)		R										
Strong (>20– ≤30%)	R		F, L	D		S		Q				F, R
Preferred (>10–≤20%)	A, K, G		W, Y	L	P, L, I	P	G, S, L		L, V, D	I, A, V	Q, L, G	K

2.3.2 Subtypes of HLA-B27

To date, there are 160 known subtypes of HLA-B27 (HLA-B*27:01 to HLA-B*27:161) based on differences between amino acid sequences (8). While some HLA-B27 subtypes are associated with AS, some subtypes are known to have no relation such as HLA-B*27:06 and HLA-B*27:09 (2,44). In two studies with 34 HLA-B27 positive patients with SpA and 26 HLA-B27 positive controls and 929 HLA-B27 positive patients with AS and 5990 HLA-B27 positive controls in the Southeast Asian population, the HLA-B*27:06 allele was found only in the control group or only 4 of 929 AS patients had it (45,46). Another study with 271 AS patients in Sardinia where is an island of Italy exhibits that the HLA-B*27:09 allele was absent in the AS patients (47). On the other hand, HLA-B*27:05 is the most mentioned subtype related to AS. However, some subtypes are common in some regions. For example, HLA-B*27:05, HLA-B*27:04, and HLA-B*27:02 seem in Caucasians, Chinese, and Mediterranean populations, respectively and they are strongly associated with AS (Figure 4a) (44). In Yao *et al.* (2010) study and Brown *et al.* (2010) study for the Chinese population, researchers investigated 317 HLA-B27 positive people in total as 172 patients with AS and 145 healthy controls and 153 patients with AS and 1545 healthy controls. According to the study of Yao and colleagues, the results of the HLA-B*27:04 allele had the highest frequency with 69.2% of patients and 53.8% of controls (48). The study by Brown shows that 105 patients and 38 controls had the HLA-B*27:04 allele (49). In addition, according to two studies in the last 20 years, HLA-B*27:05 have been the most common HLA-B27 allele in AS patients in the Turkish population (50,51). The results illustrate that 71.1% of 38 HLA-B27 positive AS patients and 68% of 47 HLA-B27 positive healthy controls had HLA-B*27:05 subtype and 65.2% of 51 HLA-B27 positive patients with AS and 54.3% of 35 HLA-B27 positive healthy controls for Birinci *et al.* (2005) study and Acar *et al.* (2012) study, respectively.

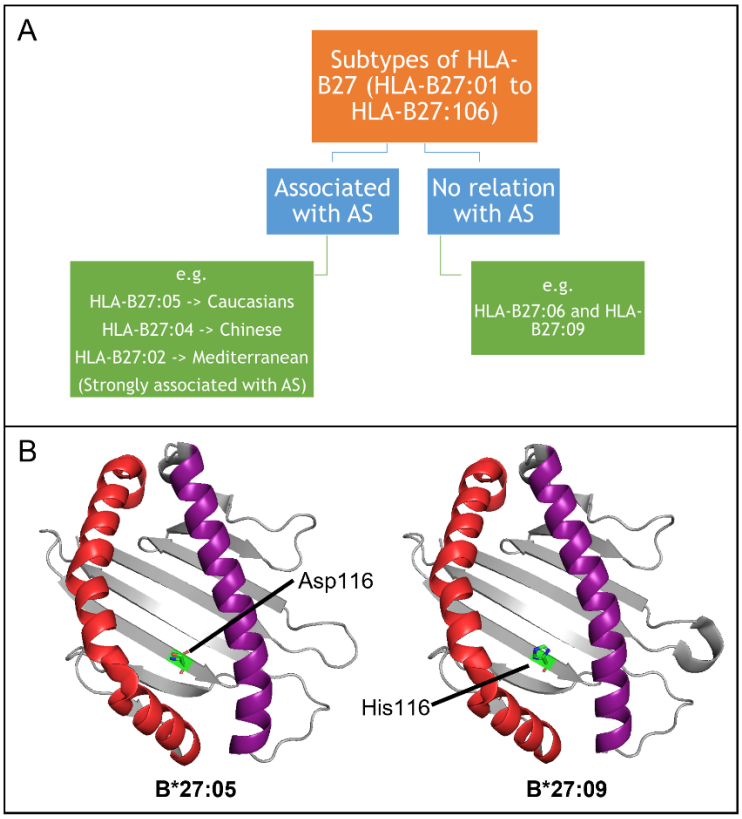

Figure 4. The association of HLA-B27 subtypes with AS disease. (a) The common HLA-B27 subtypes seen in some populations to their association with AS disease. (b) The difference between HLA-B*27:05 and HLA-B*27:09 in HLA binding groove (residues 1-180). The α1-helix and the α2-helix of the subtypes are shown in purple and red, respectively.

Structural studies show that the difference between subtypes is caused by conformational changes because of small amino acid differences between antigens of HLA – B27 subtypes and amino acid side chain differences in the binding groove (37). Besides, there is a 1 to 6 or 7 amino acid difference between subtypes of HLA, whereas there is a 15-30 amino acid difference between different B alleles. The differences between HLA-B*27:05, HLA-B*27:04, HLA-B*27:06, and HLA-B*27:09 are residues at positions 77, 114, 116, and 152. Residues at positions 114 and 116 seem more important than others since they are found on the binding floor of HLA-B27. These distinctions between subtypes result in the different types of peptides. There is only one amino acid difference between HLA-B*27:05 and HLA-B*27:09 (Figure 4b). While negatively charged Asp residue is found at position 116 of HLA-B*27:05, HLA-B*27:09 consists of positively charged His residue at

position 116 of HLA-B*27:09 (36). Therefore, HLA-B*27:05 and HLA-B*27:09 subtypes are the first preferred subtypes for the thesis since there is only one amino acid difference between these two subtypes and HLA-B*27:05 is the most common allele in the Turkish population and worldwide. All HLA-B27 alleles, their sequences, and ethnicities can be accessed through the IPD-IMGT/HLA Database (52).

2.4 Molecular Docking

Understanding the three-dimensional (3D) structure of biomolecules is an important process in order to analyze the function of the molecule. Nuclear magnetic resonance (NMR), cryogenic electronic microscopy (cryo-EM), and X-ray crystallography are three methods to visualize the 3D structure of proteins. Some problems such as the mobility of a protein for X-ray crystallography during the crystallization process and the size of a big protein for NMR are challenging to get the 3D structure of protein complexes (protein-protein, protein-ligand, and protein-nucleic acid). Also, the problem with cryo-EM is its low resolution. Therefore, the docking approach can be used to minimize these challenges (53). Molecular docking is a prediction of the best match between two or more molecules. It is used to find the best way for the interaction between molecules based on biological mechanisms using potential energy functions (54).

HADDOCK (High Ambiguity Driven biomolecular DOCKing) is a computational protein docking tool using experimental data as restraints called Ambiguous Interaction Restraints (AIR). Active and passive residues are used in the AIR. Active residues consist of known interactions between two molecules to be docked. HADDOCK accepts passive residues around selected active residues as possible contacts if the passive residues are not defined specifically. AIR calculates ambiguous intermolecular distances between an atom of an active residue of one of the proteins to be docked and an atom of an active residue or a passive residue of another protein to be docked with the threshold 3 Å. HADDOCK calculates four energies for the HADDOCK score (Equation 2.1). These are intermolecular van der

Waals energy (E_{vdw}), intermolecular electrostatic energy (E_{elec}), empirical desolvation energy (E_{desol}), and AIR energy (E_{air}) (53). E_{air} is the restraint violation energy showing consistency of experimental and calculated data (55). If a chosen active residue is not included in the interaction for docking, there will be a penalty for the HADDOCK score (53).

$$\boldsymbol{HADDOCK\ score = 1.0 * E_{vdw} + 0.2 * E_{elec} + 1.0 * E_{desol} + 0.1 * E_{air}} \quad (2.1)$$

The docking protocol of HADDOCK includes three steps which are rigid docking that allows the molecules as solid rocks, semi-flexible docking with limited flexibility of side-chains at the interface, and water refinement step which includes refinement of the structure in water molecules. Normally, at semi-flexible docking and water refinement stages, the structure is allowed to move. However, rigid docking can be performed by changing the definition of semi-flexible segments at the expert level as access level (53).

2.5 Molecular Dynamics

Studying big macromolecules with high atom numbers is hard to evaluate the dynamics of these huge systems during experimental studies and spends much time. Therefore, computational methods help to calculate quantitative features of energy surfaces of the systems (56). The study of macromolecular structure is important to understand the function of that macromolecule. When macromolecule functions, conformational changes are observed (57). MD is a computer simulation method used to find out how the particles interact with each other as time goes by (58). MD aim to provide the interaction between molecules close to reality while calculating the structural and thermodynamic features of the systems. Therefore, the conformational changes of a macromolecule can be observed. NAMD, AMBER, and GROMACS are the most popular MD tools (57). MD uses Newton's second law to guess the displacements of the particles (58). Newton's second law (Equation 2.2):

$$F = m.a \quad (2.2)$$

Here, F represents the forces acting on the particles. The forces are based on the derivative of potential energy to the changes in position (Equation 2.3). m is the mass of the system. a means the acceleration of the atoms (58). To be able to start a MD simulation, Protein Data Bank (PDB) or computational modeling studies such as homology modeling provide the initial structure of interest. Then, a topology file that includes forces acting on the particles should be prepared. Force fields are parameters that describe potential energy, and they are necessary to predict the new position of the atoms. When the forces of an atom have been determined, accelerations and velocities of the atom are calculated using Newton's law of motion for each step (57). Velocity form of the Verlet algorithm is used to integrate Newton's equation of motion (Equation 2.4) and it calculates the position (x) of an atom by using velocity (v) and acceleration (α) of the atom from previous step (57,59).

$$\boldsymbol{F_i = m_i \frac{d^2 r_i(t)}{dt^2}} \quad (2.3)$$

$$x(t + \Delta t) = x(t) + v(t) + \frac{1}{2}\alpha(t)(\Delta t)^2 \quad (2.4)$$

Molecular mechanics (MM), quantum mechanics (QM), or QM/MM force fields are used for molecular simulations and the most popular force fields are CHARMM, GROMACS, and AMBER (57,60). The total potential energy of an atom is calculated based on two forces: bonded and non-bonded forces. Bond stretching, angle bending, and bond rotation (torsion) are bonded interactions that mean atoms are covalently bonded. Non-bonded interactions usually include two forces: electrostatic potential and Lennard-Jones potential by van der Waals. These two forces are non-covalent bonds and time consuming for calculation of potential energy (60). After the topology file, a water box is generated to create a real medium for a macromolecule and the macromolecule is solvated (57). In this step, ions are also added to the system to neutralize it. Minimization of the system is the next step. A minimization step is included to get rid of high-energy interaction and relax the

system. Therefore, the most stable state of the macromolecule is found (61). Then, the initialization step is started to start the simulation smoothly. The system is equilibrated at the desired temperature (57). In MD simulations, an isolated system that is unchanged at temperature, pressure, number of particles, or volume is preferable to consider simple systems without any complications. These isolated systems are named based on the constant parameters. Microcanonical ensemble (NVE) focused on microscopic cases, canonical ensemble (NVT), and isothermal-isobaric ensemble (NPT) is constant for a number of particles (N), the size of the system (V), and the total energy (E), for a number of particles (N), volume (V), and temperature (T), and for a number of particles (N), pressure (P), and temperature (T), respectively (62). In this way, the thermodynamic properties of the system are calculated. Finally, MD simulation is produced.

3 MATERIALS AND METHODS

The steps following in the course of the thesis are explained in this section explicitly. Molecular docking and MD simulations were performed by using HADDOCK version 2.4 (https://wenmr.science.uu.nl/haddock2.4/) and NAMD version 2.14 programs, respectively (63,64). Also, VMD version 1.9.3 (Visual Molecular Dynamics), PyMOL version 2.5.0, and R version 1.4.1717 were used for the analysis of simulations and visualization (65–67). The overall summary of methodology is shown in Figure 5.

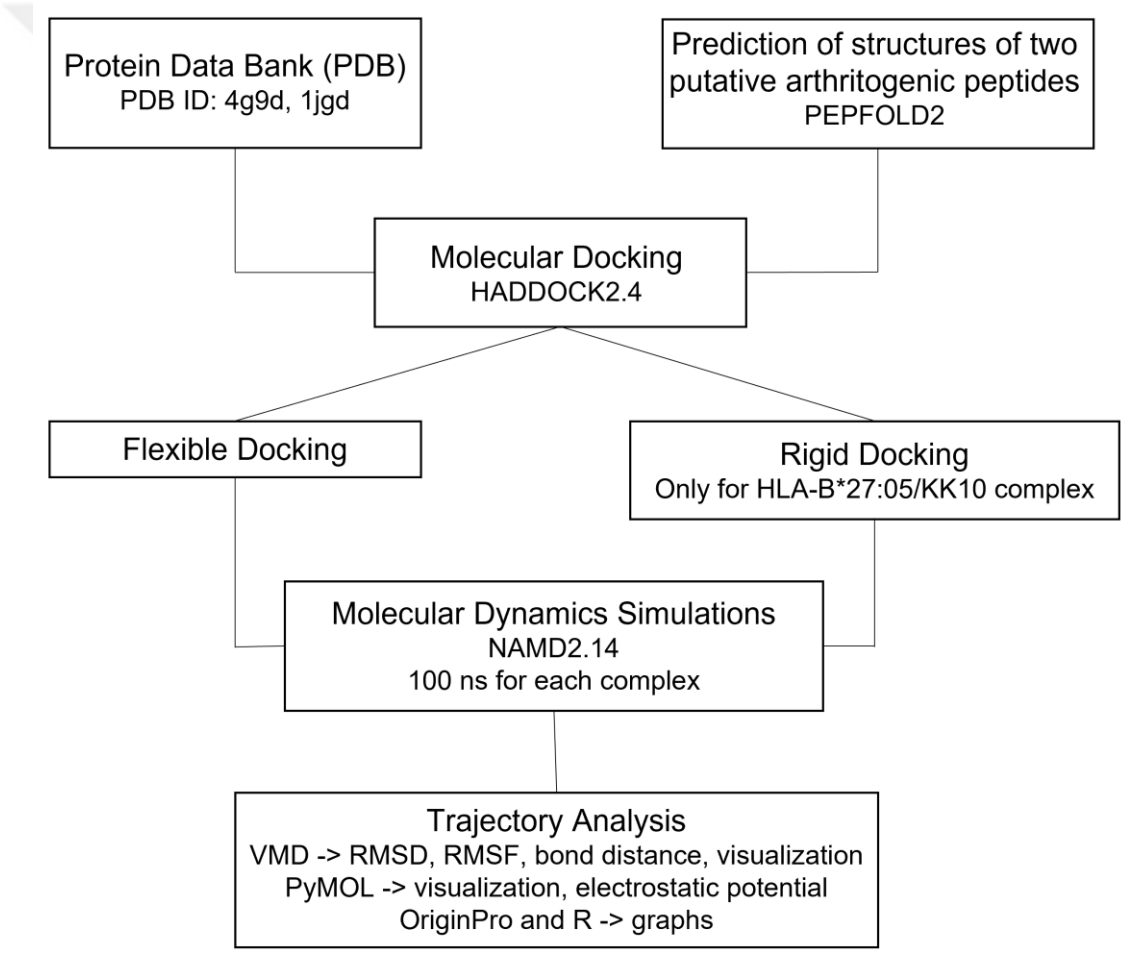

Figure 5. Schematic representation of the methodology that is followed in this thesis.

3.1 System Preparation and Molecular Docking

3.1.1 HLA-B*27:05 and the KK10 peptide

Firstly, the structures including HLA-B*27:05 with high resolution and a viral peptide were downloaded from PDB. Their amino acid sequences were examined, and the differences were specified. The structure with PDB ID 4G9D was decided to be used during the thesis study since it consists of the KK10 peptide, a viral peptide, and the structure of the HLA-B*27:05/KK10 complex with TCR is also found in PDB database (Table 2). The KK10 (KRWIILGLNK) epitope and HLA-B*27:05 protein in the structure file were separated. The structure file of 4G9D involves two conformations of some residues. Thus, the second forms were deleted to be used for HADDOCK.

Table 2. HLA-B27 subtypes that were used in the thesis.

HLA-B27 Subtypes	PDB code	Resolution of 3D Structure (Å)	Sequence of bound peptides	Ligand
HLA-B27:05 (with β2m)	4G9D	1.6 (X-ray)	KRWIILGLNK (HIV peptide)	Glycerol (GOL)
HLA-B27:09 (with β2m)	1JGD	1.9 (X-ray)	RRLLRGHNQY (s10R)	Glycerol (GOL)

To be able to set active and passive residues in the HADDOCK web server, residues within 6 Å around the peptide (PDB ID: 4G9D) were checked in PyMOL first. These residues were named as "6 Å binding site". Then, the residues that have contact with the peptide were chosen from the PDBsum website (http://www.ebi.ac.uk/thornton-srv/databases/cgi-bin/pdbsum/GetPage.pl?pdbcode=index.html) (68) and the CPORT webserver (https://alcazar.science.uu.nl/services/CPORT/) (69) was used. HADDOCK suggests CPORT to find the active residues of a protein if there is no experimental data. CPORT is a server of HADDOCK, and it helps to find active and passive residues of the protein of interest. It uses 6 interface prediction webservers to get results (69). The residues from PDBsum were named as "PDBsum

binding site". The intersection of the residues from PyMOL, CPORT, and PDBsum for active residues was used as a basis for the selection. Three articles were also evaluated to be sure of the interaction between HLA-B27 protein and a peptide to be presented to a T cell. Cys67, Ls70, Asp77, Asn97, His114, and Asp116 residues, Arg62, Asp74, Leu81, Tyr123, Thr143, Glu163, and Trp167 residues, and Tyr7, Tyr59, and Tyr171 residues are derived from aforementioned literature (70–72).

The HLA-B*27:05 protein and the KK10 peptide were docked by HADDOCK to compare the main structure and the docked result. In the first docking process, HLA-B*27:05 and the KK10 epitope were re-docked without considering flexibility. According to Alexandre Bonvin , one of the developers of HADDOCK, *et al.* study, giving flexibility to the peptide in a docking study with a protein-peptide complex is a more logical approach (73). The results of 4 flexible dockings and 2 rigid dockings of the HLA-B*27:05/KK10 complex with different parameters were compared to select best option for docking of HLA-B27 subtypes and two arthritogenic self-peptides. They were named as 4g9d-1, 4g9d-2, 4g9d-3, 4g9d-4, 4g9d-5, and 4g9d-6 to their parameters. 2 rigid dockings and 2 flexible dockings with different active residue ranges are as 4g9d-1 (pdbsum binding site residues), 4g9d-2 (6 Å binding site residues), 4g9d-3 (6 Å binding site residues), 4g9d-4 (pdbsum binding site residues). Active residues of 4g9d-5 and 4g9d-6 dockings were selected as residues within 6 Å from the peptide. Passive residues for protein were defined as automatically. All peptide residues were set as passive residues. The peptides were fully flexible for the dockings to observe larger conformational changes. Also, differently, full flexibility was given to the binding site (alpha bridge) of HLA protein for 4g9d-5 docking. For 4g9d-6, unambiguous restraint interaction which is for residues defined highly reliable predictions was used instead of ambiguous restraint interactions. The other parameters were changed as follows for all 6 dockings: The model numbers were increased to increase the accuracy of the dockings. The number of structures for rigid-body docking, the number of structures for semi-flexible refinement, and the number of structures for the final refinement were increased to 2000, 400, and 400, respectively. All the structures were refined with short molecular dynamics in explicit solvent. Also, the number of MD steps for

the rigid body high-temperature torsion angle molecular dynamics (TAD), number of MD steps during the first rigid body cooling stage, number of MD steps during the second cooling stage with flexible sidechains at the interface, and number of MD steps during the third cooling stage with fully flexible interface were changed to 2000, 2000, 4000, and 4000, respectively. With this comparison, flexible docking was used in the next steps of the study since the peptides are more conveniently attached to the binding site.

The KK10 epitope sequence was modeled to be able to compare with the main structure and decide which program will be used to predict peptides. Robetta and I-TASSER could not be used since they don't accept sequences less than 26 amino acids and 10 amino acids, respectively. PEPFOLD-2 (https://mobyle.rpbs.univ-paris-diderot.fr/cgi-bin/portal.py#forms::PEP-FOLD) (74) was used for peptide structure prediction. According to the PEP-FOLD2 result, lower sOPEP energy and higher tm value mean a better model, so PEP-FOLD2 sorts the results to sOPEP energy (74).

3.1.2 Two arthritogenic self-peptides and the KK10 (HIV) peptide

The structures of two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL) were modeled using PEP-FOLD2. They were previously identified by Atagündüz et. al. (2005) study (6). PEP-FOLD2 has produced 5 and 2 models for ARGQPGVMG and DRASFIKNL, respectively. Also, other structures of interest (Table 2) which are HLA-B*27:05, HLA-B*27:09, and the KK10 peptide were provided by PDB. Peptides in PDB structures were extracted. Only the KK10 (KRWIILGLNK) epitope on the Gag protein of human immunodeficiency virus (HIV) was used for the docking process. All produced models for these three peptides were docked with HLA-B*27:05 and HLA-B*27:09. The interface of the proteins was defined as the previously listed intersection residues. For the docking process of HLA-B27 subtypes with two arthritogenic self-peptides and the KK10 peptide, active residues of the protein were defined as the interface residues. Passive residues for protein were defined automatically. All peptide residues were selected as passive residues. The peptides were selected as fully flexible for the dockings to

observe larger conformational changes. The sampling parameters were changed as follows: The model numbers were increased to increase the accuracy of the dockings. The number of structures for rigid-body docking, the number of structures for semi-flexible refinement, and the number of structures for the final refinement were increased to 2000, 400, and 400, respectively. All the structures were refined with short molecular dynamics in explicit solvent. Also, the number of MD steps for the rigid body high-temperature TAD, number of MD steps during the first rigid body cooling stage, number of MD steps during the second cooling stage with flexible sidechains at the interface, and number of MD steps during the third cooling stage with fully flexible interface were changed to 2000, 2000, 4000, and 4000, respectively. The number of models generated was adjusted as 400. Representative complexes were selected to cluster size, cluster number, and HADDOCK score.

3.2 Molecular Dynamics Simulation

Before simulation of putative arthritogenic self-peptides with HLA-B27 subtypes, 10 parallel runs for two HLA-B27:05 and the KK10 peptide complex from the comparison of rigid docking and flexible docking were performed to control if the system gives better results whether the docking is rigid or flexible (Table 3). Then, the HLA-B27 complexes with a viral peptide (KK10) and two arthritogenic self-peptides (ARGQPGVMG and DRASFIKNL) were evaluated (Table 3). The complexes of them from docking results were used for MD simulations. Configuration files for simulations were prepared using Solution Builder of CHARMM-GUI (75). Simulations were carried out by NAMD (76). Disulfide bonds between residue 101-164 and residue 203-259 for heavy chain and residue 25-80 for β2m were added. The water box was fit to protein size. The complexes were solvated in TIP3P water (~ minimum 10 Å) and neutralized with 0.15 M NaCl. The force field was CHARMM36m (77). Particle mesh Ewald (PME) method was considered for long-range electrostatic calculations. The systems were minimized by 50,000 steps, and it was decided by controlling the total energy of the system in the output. The temperature was increased gradually in the heating step from 0 K (1 K for every 100 steps) to 310.15 K. NVT ensemble and NPT ensemble were chosen for equilibration

and production steps. 2 ns and 100 ns simulations were performed for each complex and HLA-B27 proteins without peptide for equilibration and production steps, respectively. Also, the restraints parameter was changed at the end of the equilibration configuration file to open the system gradually. Four independent simulations were performed for each complex and free HLA-B27 (Table 3).

Table 3. Content of simulations performed during the thesis.

HLA-B27 Subtypes	Peptide	Docking	# of simulations	Duration of Each Simulation	Total Duration
HLA-B27:05	ARGQPGVMG	Flexible Docking	4	100 ns (4)	400 ns
	DRASFIKNL	Flexible Docking	4	100 ns (4)	400 ns
	KRWIILGLNK	Flexible Docking	6 (comparison*) + 4	100 ns (4+6)	1000 ns
		Rigid Docking	6 (comparison*)	100 ns	600 ns
	Without peptide	Flexible Docking	4	100 ns	400 ns
HLA-B27:09	ARGQPGVMG	Flexible Docking	4	100 ns (4)	400 ns
	DRASFIKNL	Flexible Docking	4	100 ns (4)	400 ns
	KRWIILGLNK	Flexible Docking	4	100 ns (4)	400 ns
	Without peptide	Flexible Docking	4	100 ns	400 ns
Total # of simulations:			44		

* Simulations to decide between flexible docking and rigid docking

3.3 Trajectory Analysis and Visualization

Root Mean Square Deviation (RMSD) and Root Mean Square Fluctuation (RMSF) were calculated using Tcl scripting of VMD program for the heavy chain, binding site, and β2m of HLA-B27 protein. The snapshots of HLA-B*27:05 and HLA-B*27:09 with the KK10 peptide and two arthritogenic self-peptides and the empty state of the HLA-B27 subtypes were extracted from MD trajectories at 0 ns, 50 ns, and 100 ns for each simulation of four simulations using VMD software program and PyMOL program. Also, these snapshots were superimposed by PyMOL.

Furthermore, distances of p1 (Tyr59-Trp167) and pΩ (Tyr84-Ala139) sides of HLA-B27 subtypes with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 peptide, and without peptide were calculated. Moreover,

the distance between the Cα atom of Asp116 (HLA-B*27:05)/His116 (HLA-B*27:09) and PΩ of two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL) and the KK10 peptide, the distance between the Cα atom of Asp116 (HLA-B*27:05) /His116 (HLA-B*27:09) and the Cα atom of Trp60 of β2m, H-bond distance between the OD1 atom of Asp122 of the heavy chain of HLA-B27 and the NE1 atom of Trp60 of β2m, and H-bond distance between the NE2 atom of Gln96 of the heavy chain of HLA-B27 and the O atom of Trp60 of β2m were analyzed by Tcl scripting. Another calculation was for the number of H-bond interactions between the HLA-B27 subtypes and three peptides. If the distance between acceptor and donor atom was smaller than 3.5 Å, the bond was accepted as H-bond. Electrostatic potential surfaces of the HLA-peptide complexes and empty state HLA-B27 subtypes were mapped and visualized by Adaptive Poisson-Boltzmann Solver (APBS) Plugin of PyMOL. PPCHECK webserver (http://caps.ncbs.res.in/ppcheck/) (78) was used to observe the potential salt bridges with a 4 Å threshold between the binding groove of HLA-B27 complexes and peptides from MD trajectories at 0 ns, 50 ns, and 100 ns for each simulation of four simulations. Ligand-protein interaction diagrams were plotted for H-bond and hydrophobic bond by LigPlot+ program (79). The diagrams were aligned and the common residues between simulations at 0 ns, 50 ns, and 100 ns were saved.

3.4 Additional Analysis

Purcell *et al.* (2015) have the binding motifs for peptides of different lengths that bind to HLA-B*27:05 (Table 3) (9). When checking the motifs, there is Arg residue for all P2 positions. These motifs were used to check the similarities between the peptides in the study and the peptides used in the thesis.

The peptides were analyzed using Immune Epitope Database (IEDB) analysis tool to predict their antigenicity (http://tools.iedb.org/mhci/) with the NetMHCpan-4.1 prediction method (80). HLA-B*27:05 allele was selected with all length options. HLA-B*27:09 allele could not be evaluated since it was not in the options for HLA alleles. If the score value in the binding prediction results table is high, this shows

that the peptide can bind to the HLA allele better. Also, a smaller percentile rank indicates higher affinity. Moreover, Class I immunogenicity (http://tools.iedb.org/immunogenicity/) for HLA-B*27:05 was calculated for the three peptides using IEDB analysis tool (81).

The sequence of the viral peptide and two self-peptides were analyzed using BLAST to control if there is a similar sequence from *Homo sapiens* for the KK10 peptide and the sequences from pathogens for the self-peptides.

Purcell *et al.* (2015) have the lists of peptides bound to HLA-B*27:05 and HLA-B*27:09 as a result of their study beside the consensus binding motifs. HLA-B*27:05 has 2132 peptides and HLAB*27:09 has 2276 peptides. The sequences of these peptides were uploaded into the RAACBook web server (http://bioinfor.imu.edu.cn/raacbook/public/), and the alphabet was reduced. Type 55 was used to cluster the amino acid sequence of peptides. It reduces the alphabet of amino acid sequences based on the hydrophobic and polar (HP) model (82). Also, the percentages of four clusters (A, D, S, K) were calculated for the peptides bound to HLA-B*27:05 and HLA-B*27:09 from the 2015 study. Box plot was used to understand the differences in the variability of the peptides better by using a ggplot of R (83). The RAACBook web server also was used to reduce the alphabet of two arthritogenic self-peptides and the KK10 peptide.

4 RESULTS

4.1 Determining the Docking Strategy in the Thesis

4.1.1 Docking of HLA-B*27:05 and the KK10 epitope

Firstly, 2 rigid dockings and 4 flexible dockings were performed to decide which docking result to use in MD simulations (Table 4). Two rigid dockings and two flexible dockings with different active residue ranges are as follows: 4g9d-1 (pdbsum binding site residues), 4g9d-2 (6 Å binding site residues), 4g9d-3 (6 Å binding site residues), 4g9d-4 (pdbsum binding site residues). While 4g9d-1 and 4g9d-2 represent rigid docking, 4g9d-3 and 4g9d-4 state flexible docking. Additionally, for 4g9d-5, full flexibility was given not only to the peptide but also to the protein and unambiguous restraint interaction was used instead of ambiguous restraint interactions for 4g9d-6.

Table 4. The lowest HADDOCK scores for the HLA-B*27:05/KK10 complexes.

4g9d-Redock	HADDOCK Score	Cluster Number	Cluster Size
4g9d-1	-174.7	4	18
4g9d-2	-184.4	1	171
4g9d-3	-153.6	9	12
4g9d-4	-118.1	17	4
4g9d-5	-172.7	3	34
4g9d-6	-166.0	7	13

When the active residues were chosen as 6 Å binding site residues, the HADDOCK score was lower for 4g9d-2 and 4g9d-3. Also, the predicted complexes with the lowest scores were compared (Figure 6). Although 4g9d-1 and 4g9d-2 were superimposed with the reference structure perfectly, 4g9d-2 has a lower docking score. Therefore, it was selected for the MD simulation step as the complex from rigid docking. β2m and the heavy chain of the HLA-B27 complex were separated and the integrity of the complex was disrupted as the output of 4g9d-5 docking. 4g9d-4 did not match with the main PDB structure as in 4g9d-3 and 4g9d-6. Since

the score of 4g9d-6 was lower than 4g9d-3, the docking result of 4g9d-6 was chosen for the MD simulation.

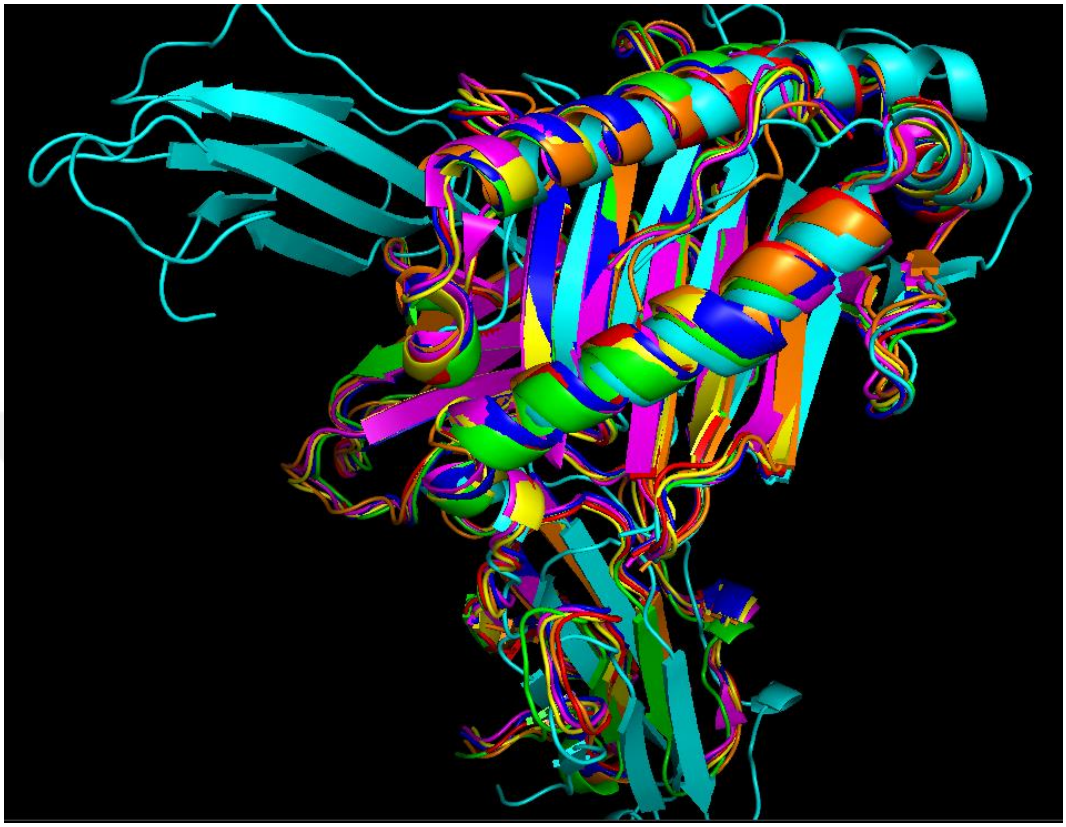

Figure 6. The structural comparison of docking results and the structure of the HLA-B*27:05/KK10 complex from the reference structure (PDB ID: 4G9D). 4g9d.pdb: magenta, 4g9d-1: green, 4g9d-2: red, 4g9d3: blue, 4g9d-4: orange, 4g9d-5: cyan, and 4g9d-6: yellow.

4.1.2 RMSD analysis of HLA-B*27:05 and the KK10 epitope

A total of 12 simulations were carried out for 4g9d-2 and 4g9d-6, 6 for each docking result. The RMSD values of the binding groove of the HLA-B*27:05 were analyzed to compare the stability of the HLA-B*27:05/KK10 complexes (Figure 7). According to the RMSD measurements, the simulation results from the flexible docking complex has lower RMSD values than the simulation results from the rigid docking complex. The range of RMSD value for parallel runs of flexible docking complex is between 1-1.5 Å, while the range of rigid docking complex is between 1-

2 Å. Since the results of 6 parallel runs from 4g9d-6 were more stable and giving flexibility to the peptide showed better results, as stated in Bonvin *et al.* (2013) (73), flexible docking was found to be reasonable method for the study of HLA-B27 subtypes with arthritogenic self-peptides.

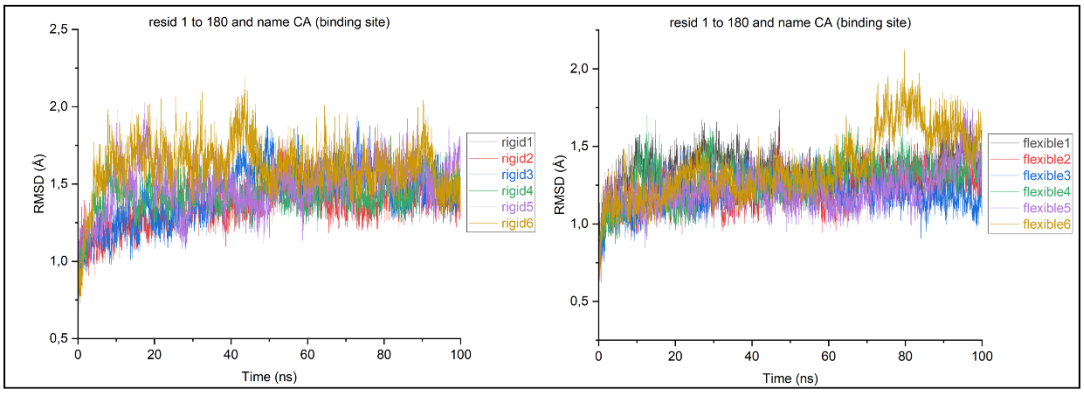

Figure 7. RMSD values of the binding groove (residues 1-180) of HLA-B27 complex.

4.2 Structural Analysis of One Viral Peptide and Two Arthritogenic Self-peptides with HLA-B27 Subtypes

This section is to show structural differences and similarities between two self-peptides which are arthritogenic self-peptides (ARGQPGVMG and DRASFIKNL) and a viral peptide from Gag protein as the KK10 (KRWIILGLNK) peptide bound to HLA-B*27:05 and HLA-B*27:09. Two HLA-B27 subtypes are differentially associated with the AS disease and differs in only one amino acid at F pocket (residue 116) of HLA-B27 complex.

4.2.1 Docking results

The HADDOCK scores of HLA-B*27:05 and HLA-B*27:09 bound to ARGQPGVMG, DRASFIKNL, and KRWIILGLNK peptide are shown in Table 5. It seems that DRASFIKNL with -130.172 for its complex with HLA-B*27:05 and -115.344 for its complex with HLA-B*27:09 demonstrated a similar HADDOCK

score to the KK10 peptide with -161.813 for HLA-B*27:05 complex and -129.807 for HLA-B*27:09 complex. Also, the size of clusters generated from the docking runs of DRASFIKNL peptide with both HLA-B*27:05 and HLA-B*27:09 complexes was larger. HADDOCK docking algorithm clusters the predicted complexes by considering RMSD cutoff given as an input parameter. It means that there are many more structures similar to the complex structure of DRASFIKNL bound to HLA-B*27:05 and HLA-B*27:09 in the docking result.

Table 5. The lowest HADDOCK scores for HLA-B*27:05 and HLA-B*27:09 bound to ARGQPGVMG, DRASFIKNL, and KRWIILGLNK peptides.

Peptide Sequence	HLA-B*27:05			HLA-B*27:09		
	HADDOCK Score	Cluster Number	Cluster Size	HADDOCK Score	Cluster Number	Cluster Size
ARGQPGVMG	-96.772	1	84	-94.718	11	4
DRASFIKNL	-130.172	1	126	-115.344	1	84
KRWIILGLNK	-161.813	12	9	-129.807	13	8

To analyze the docking results structurally, they were aligned in PyMOL (Figure 8). The KK10 peptide was bound to HLA-B*27:05 and HLA-B*27:09 differently. While the structure of the KK10 peptide with HLA-B*27:05 was bulging out, the KK10 peptide with HLA-B*27:09 had non-canonical shape. Also, it seems that the DRASFIKNL and KK10 peptides were tethered with N- and C-terminus of them. ARGQPVMG peptide was located closer to the binding groove of HLA-B*27:05 when compared with the HLA-B*27:09/ARGQPVMG complex.

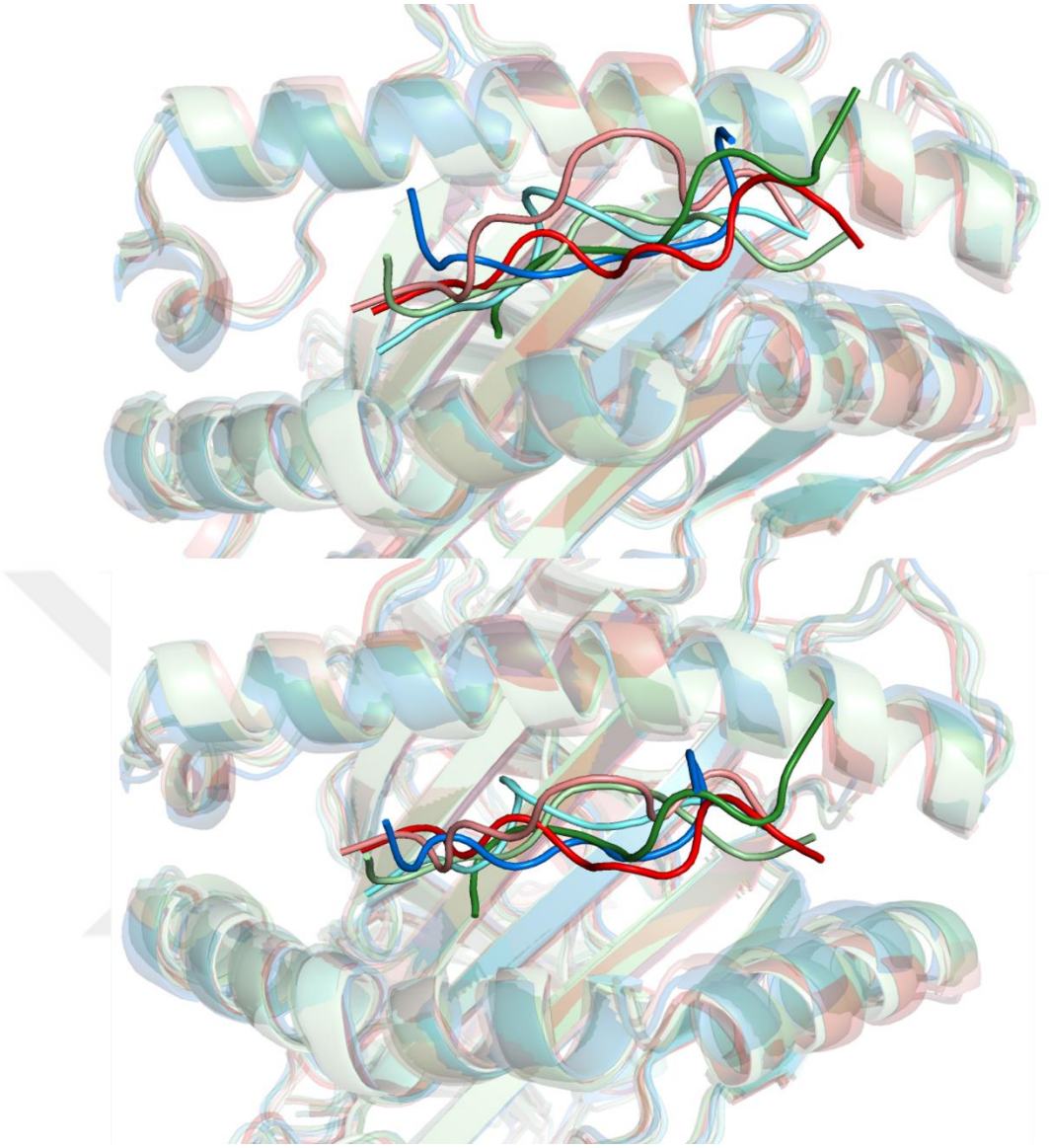

Figure 8. The side and top view of superimposition of 6 HLA-B27 and peptide complexes of docking results. HLA-B*27:05/KK10: salmon, HLA-B*27:09/KK10: red, HLA-B*27:05/ARGQPGVMG: pale green, HLA-B*27:09/ARGQPGVMG: forest, HLA-B*27:05/DRASFIKNL: aqua marine, and HLA-B*27:05/DRASFIKNL: marine.

4.2.2 Visualization of MD simulations

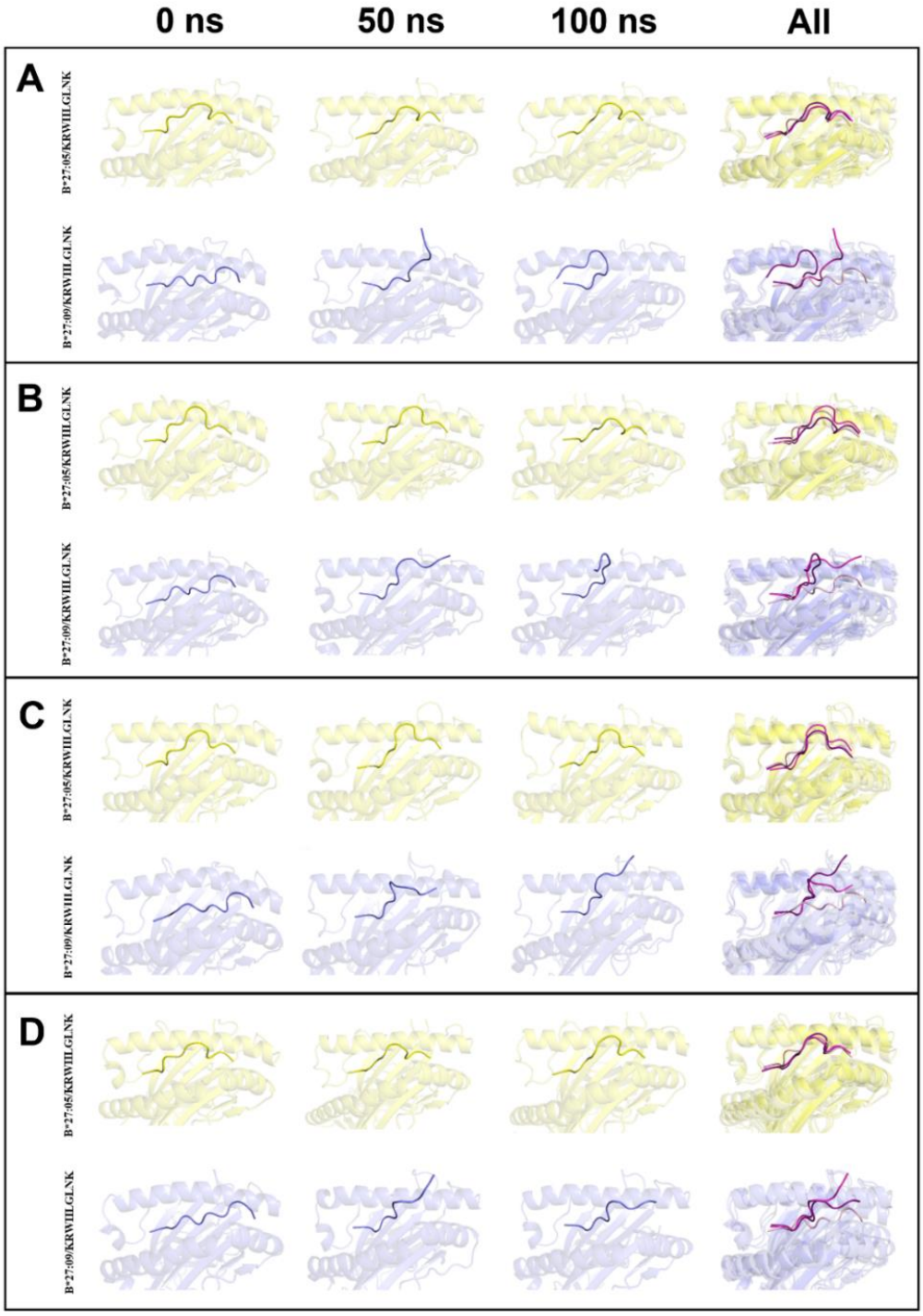

Figure 9. The snapshots of HLA-B*27:05 and HLA-B*27:09 with KRWIILGLNK (HIV) peptide (PDB ID: 4g9d, 1jgd) at 0 ns, 50 ns, and 100ns extracted from four MD trajectories. Four independent MD simulations are shown (A, B, C, and D). As time progresses, the color becomes darker (All part). 0 ns: light pink, 50 ns: light magenta, 100 ns: deep purple.

The docking results show only one snapshot of the protein-peptide complex. Therefore, long-term analysis of the protein of interest or complex structure using MD simulation offers a broader perspective on the structure of the protein. With this aim, 24 MD simulations were performed for HLA class I complexes bound to peptides which are HLA-B*27:05/KK10, HLA-B*27:09/KK10, HLA-B*27:05/ARGQPGVMG, HLA-B*27:09/ARGQPGVMG, HLA-B*27:05/ DRASFIKNL, and HLA-B*27:09/DRASFIKNL. Furthermore, 8 simulations were also carried out for HLA-B*27:05 and HLA-B*27:09 in unbound state to compare the dynamic of HLA-B27 molecules to understand the effect of the peptide binding. The results of MD simulations were visualized to check an observable change in the structures of the complexes.

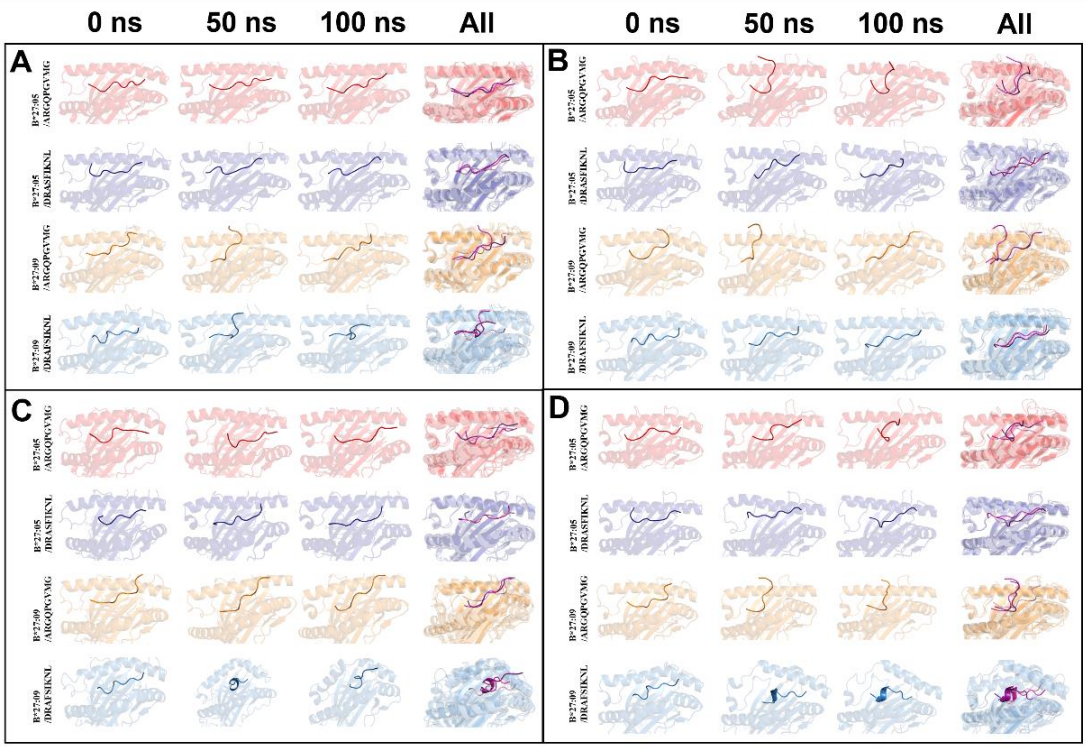

Figure 10. The snapshots of HLA-B*27:05 and HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL) at 0 ns, 50 ns, and 100ns extracted from four MD trajectories. Four independent MD simulations are shown (A, B, C, and D). As time progresses, the color becomes darker for peptide molecule (All part). 0 ns: light pink, 50 ns: light magenta, 100 ns: deep purple.

The comparison of binding modes for the complexes are shown in Figure 9 and Figure 10. While the structures of the KK10 peptide were similar in all 4 simulations, the other two self-peptides had more conformational changes. The structures of the KK10 peptide were stable for the complexes with HLA-B*27:05. On the other hand, for HLA-B*27:09, the C-terminus of the KK10 peptide moved away from the binding groove for all 4 simulations. As a result of the 3rd and 4th simulations, the DRASFIKNL peptide disrupted the protein structure of HLA-B*27:09. The N-terminus of the self-peptides appear to be bound in all simulations, although there were fluctuations in the C-terminus.

In addition to the comparison of the structure of peptides, the peptide-bound and unbound states of HLA-B*27:05 and HLA-B*27:09 were examined by taking snapshots at 100 ns of each simulation, and how much the HLA protein and peptide changed for 4 simulations by aligning the snapshots taken at 0, 50, and 100 ns (Figure 11). When looking at the empty state of HLA-B*27:05 and HLA-B*27:09, the α2 chain of HLA-B*27:05 and the floor of four β-sheets were more movable. Moreover, the KK10 peptide is not constant for its complex with HLA-B*27:09. Although the binding of the KK10 peptide was stabilized in both subtypes, the empty state of HLA-B*27:05 showed more mobility, but with the binding of KK10, it becomes more stabilized than HLA-B*27:09. DRASFIKNL peptide caused stabilization of α1 and α2 domains of HLA-B*27:05, while at HLA-B*27:09 it resulted in further activation of α domains.

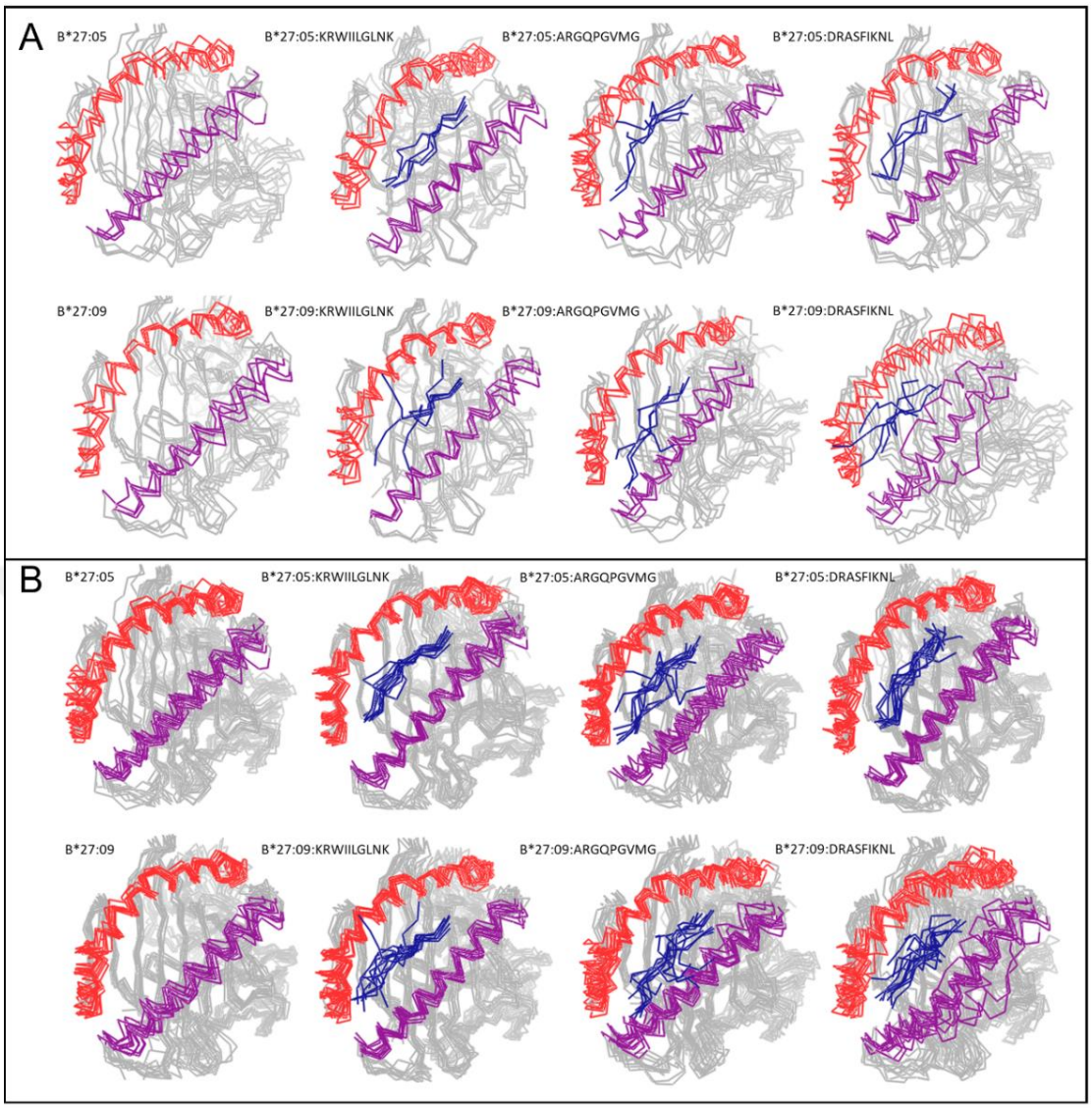

Figure 11. The snapshot of HLA-B*27:05 and HLA-B*27:09 with and without a peptide. (A) Snapshots are taken from MD simulations at 100ns for each simulation of four simulations, and they were superimposed. (B) Snapshots are taken from MD simulations at 0 ns, 50 ns, and 100 ns for each simulation of four simulations, and they were superimposed. The α1-helix, the α2-helix, and peptides are shown in purple, red, and dark blue, respectively.

4.2.3 Trajectory analyses of MD simulations

The binding grooves of the subtypes were analyzed to understand the structural and dynamical differences between HLA*B-27:05 and HLA*B-27:09 (Figure 12 and Figure 13). The unbound HLA-B*27:09 was more stable than the unbound HLA-B*27:05. The HLA-B*27:05/KK10 complex had more protein stability and its

RMSD values were smaller than the RMSD value of the HLA-B*27:09/KK10 complex. Although the complex with peptide ARGQPGVMG showed similar RMSD values for HLA-B*27:05 and HLA-B*27:09, the DRASFIKNL peptide represented similar stability as in the protein/the KK10 peptide complex simulations.

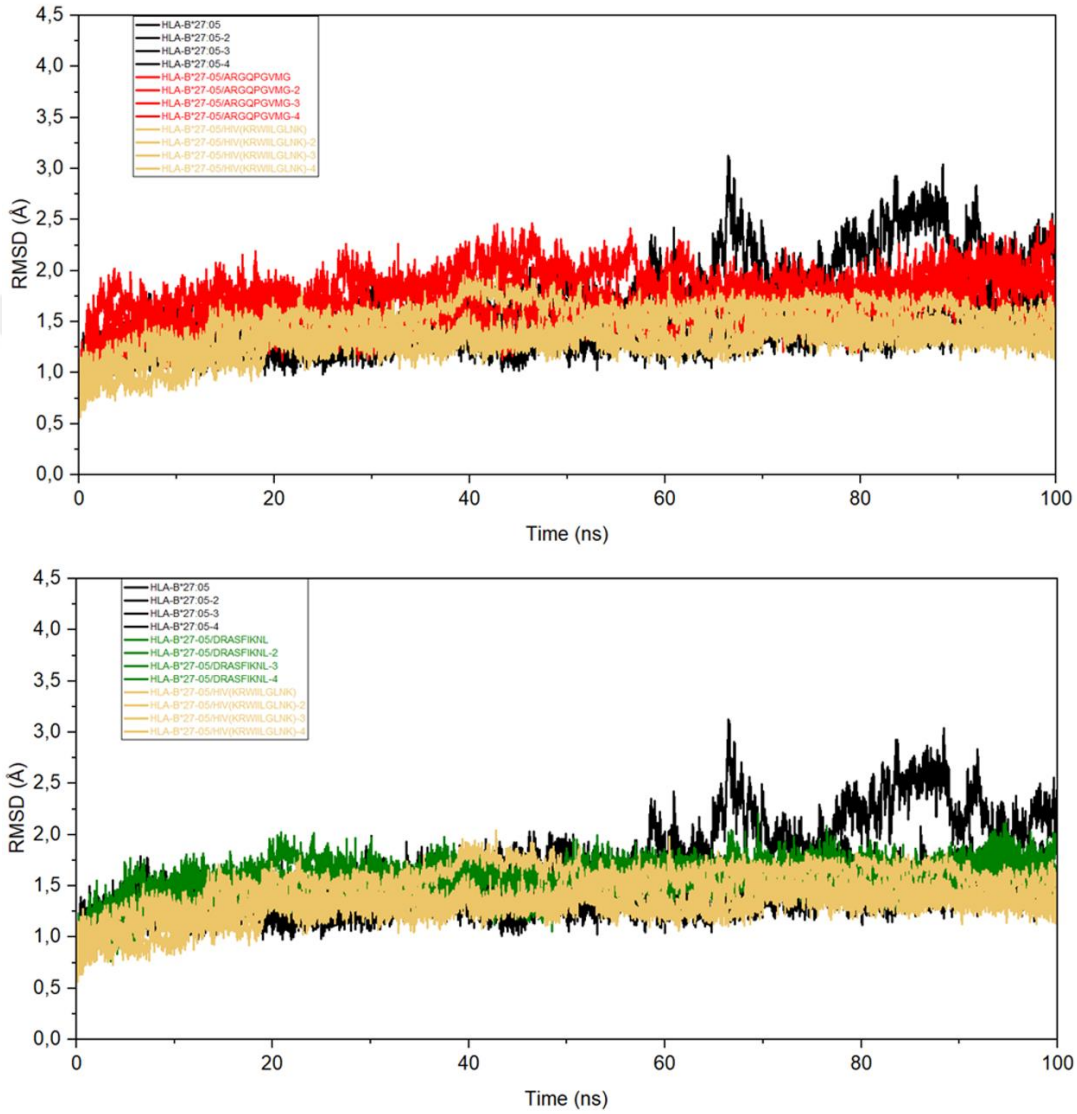

Figure 12. RMSD values of the binding grooves (residues 1-180) of HLA-B*27:05 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

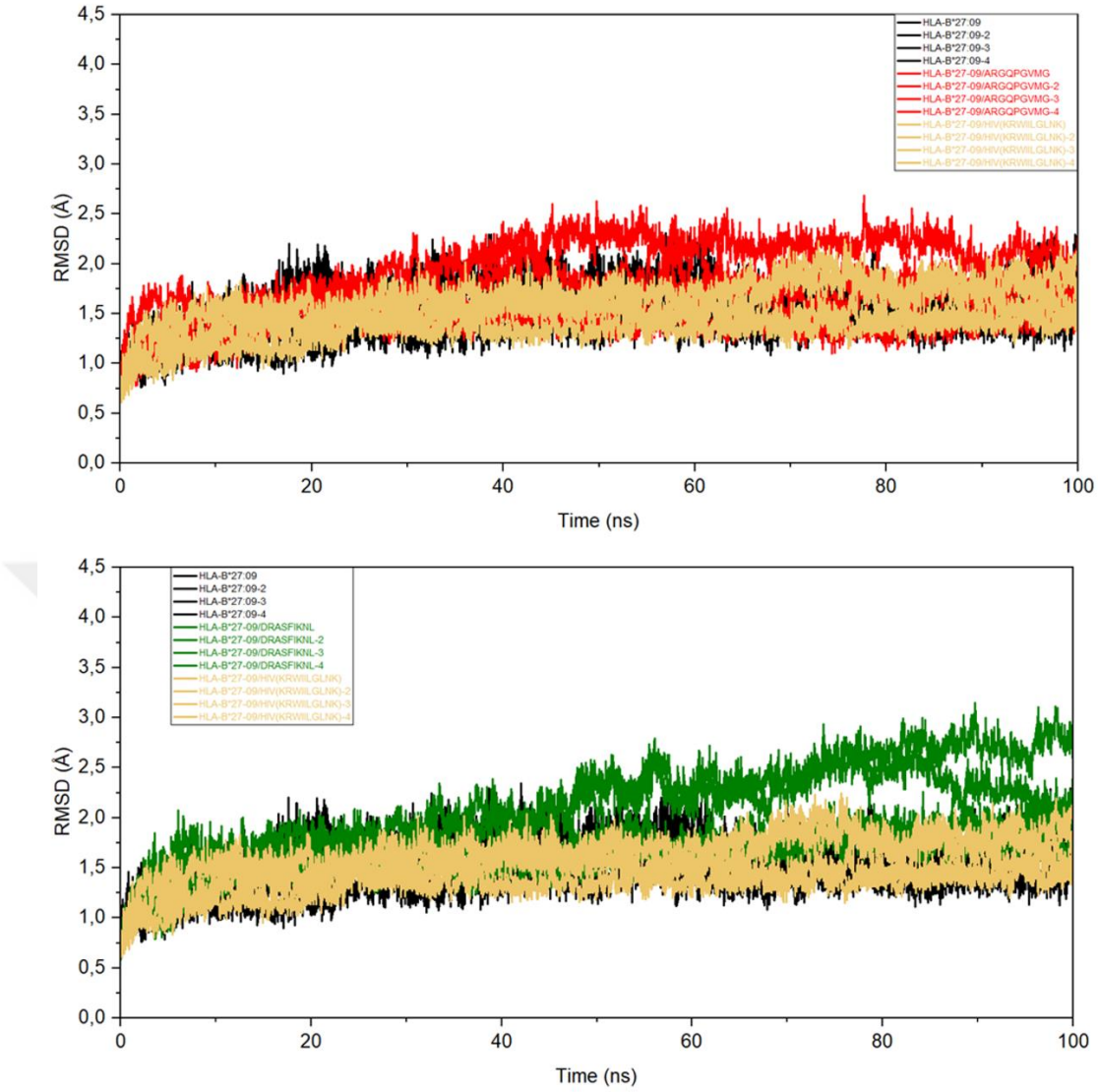

Figure 13. RMSD values of the binding grooves (residues 1-180) of HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

The RMSF values of the heavy chain and β2m of HLA-B27 subtypes with two self-peptides and one viral peptide were analyzed to understand the structural stability between HLA-B*27:05 and HLA-B*27:09 (Figure 14-17). Regions exhibiting the highest fluctuations were identified in the loop regions of the HLA-B27 proteins and they are located around residues 20, 90, 225, and 250 of the heavy chain (Figure 14 and Figure 15). The HLA-B*27:09/DRASFIKNL complex showed high fluctuations at many points due to the conformational change of the protein in

the 3rd and 4th parallel simulations (Figure 15). When the KK10 peptide is bound to the HLA-B*27:05 subtype, fluctuations in the α-2 domain decreased (Figure 14). Like the heavy chain of HLA-B27, β2m of HLA-B27 showed fluctuations for loop regions.

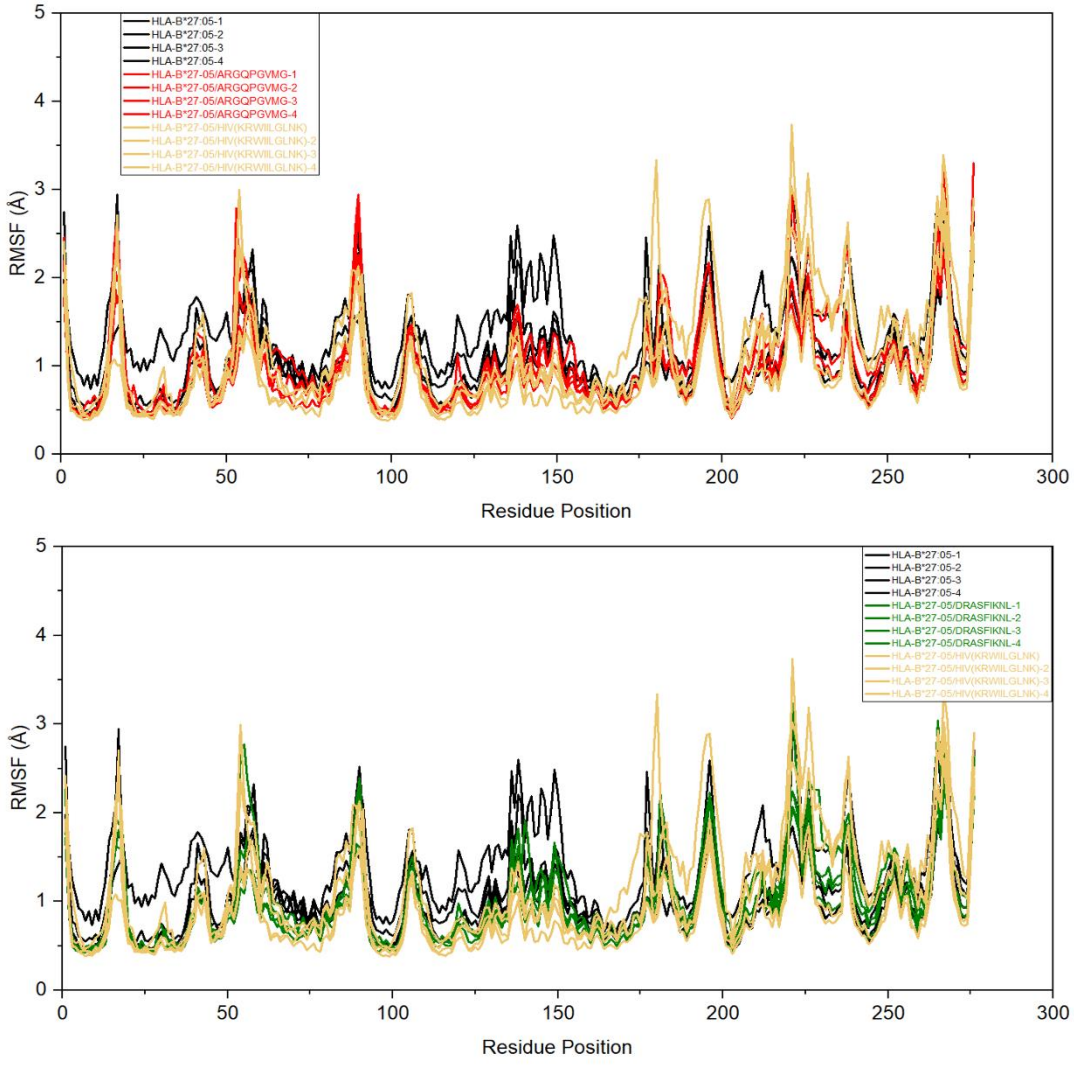

Figure 14. RMSF values of the heavy chain (α1, α2, and α3 domains) of HLA-B*27:05 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

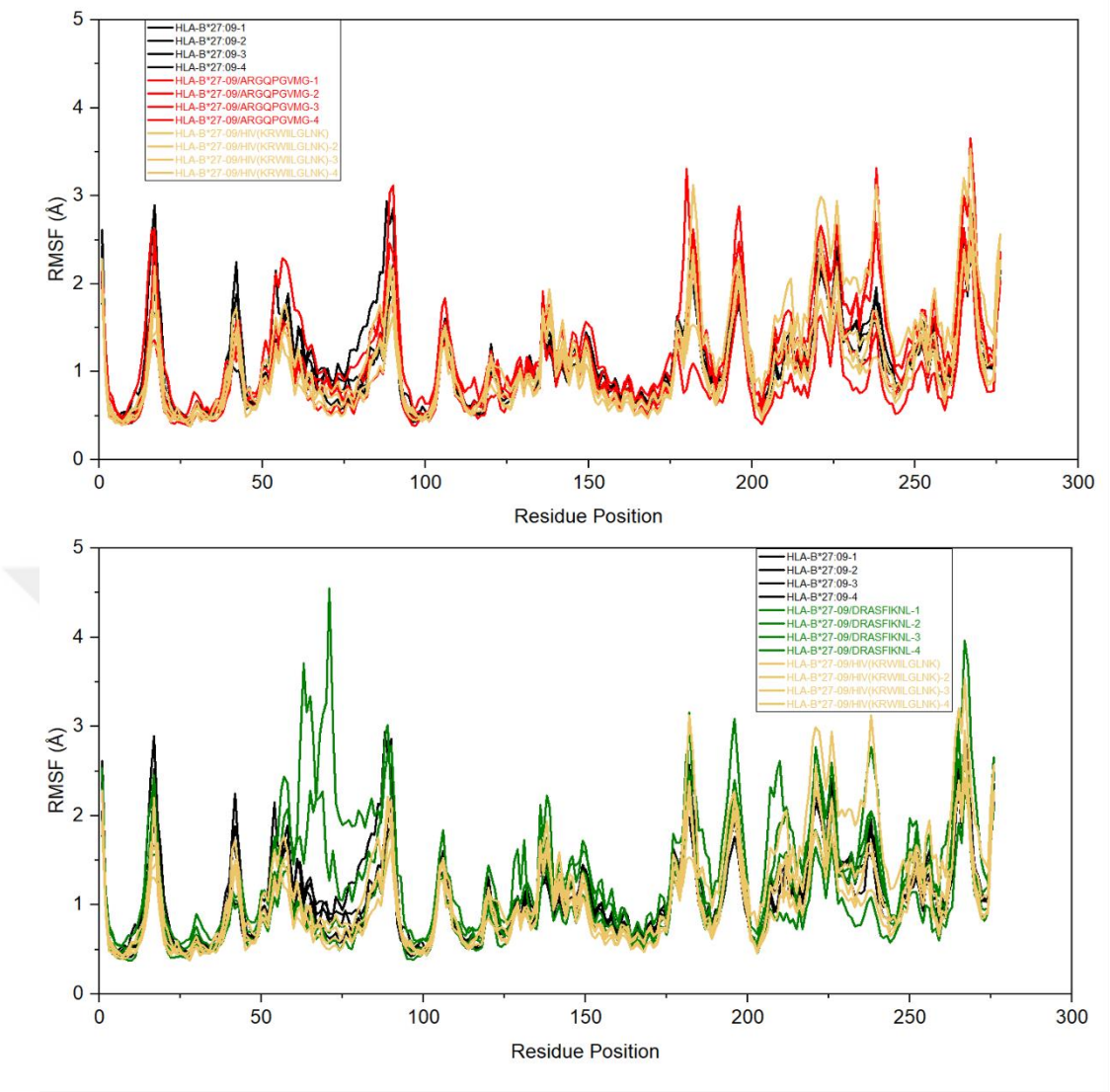

Figure 15. RMSF values of the heavy chain (α1, α2, and α3 domains) of HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

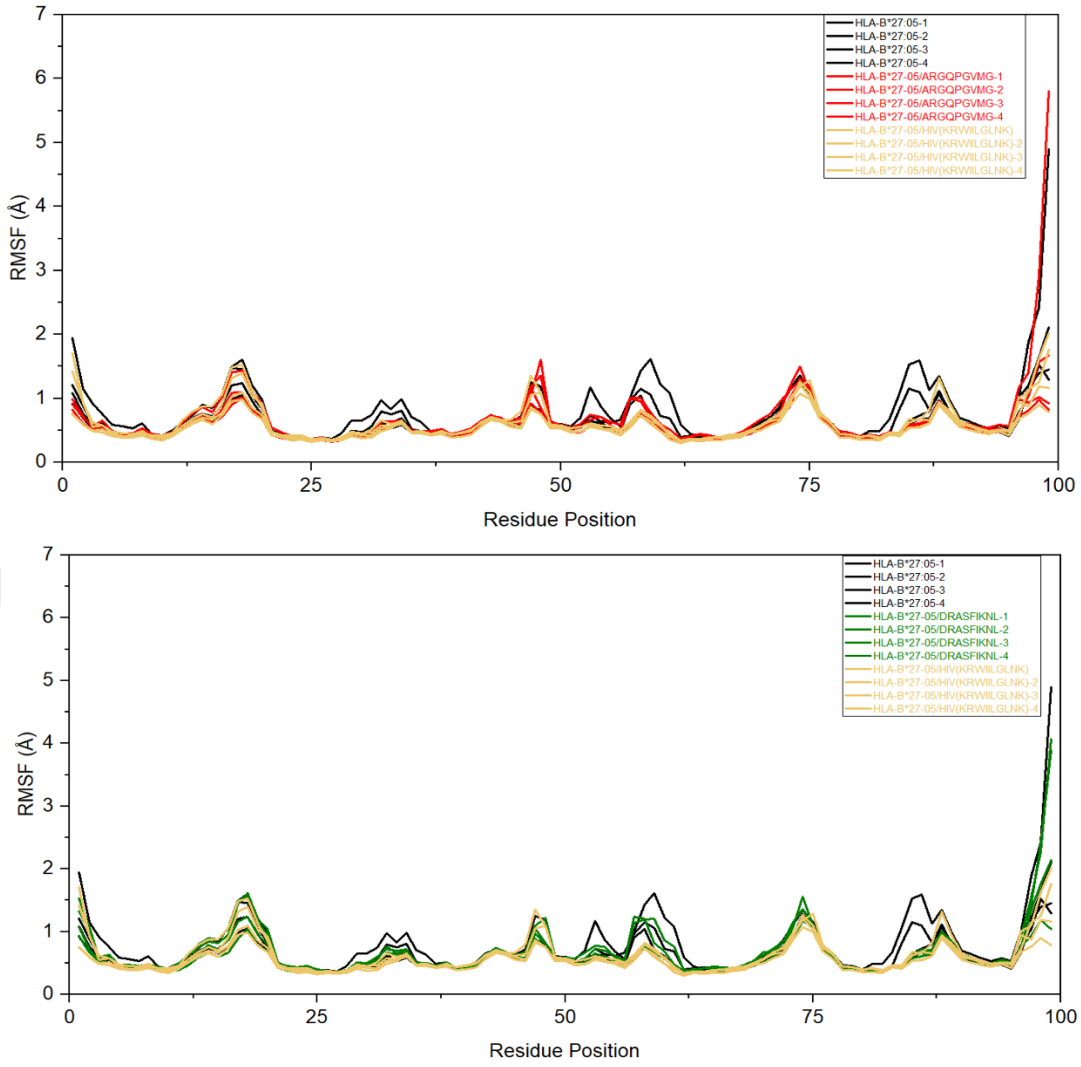

Figure 16. RMSF values of β2m of HLA-B*27:05 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

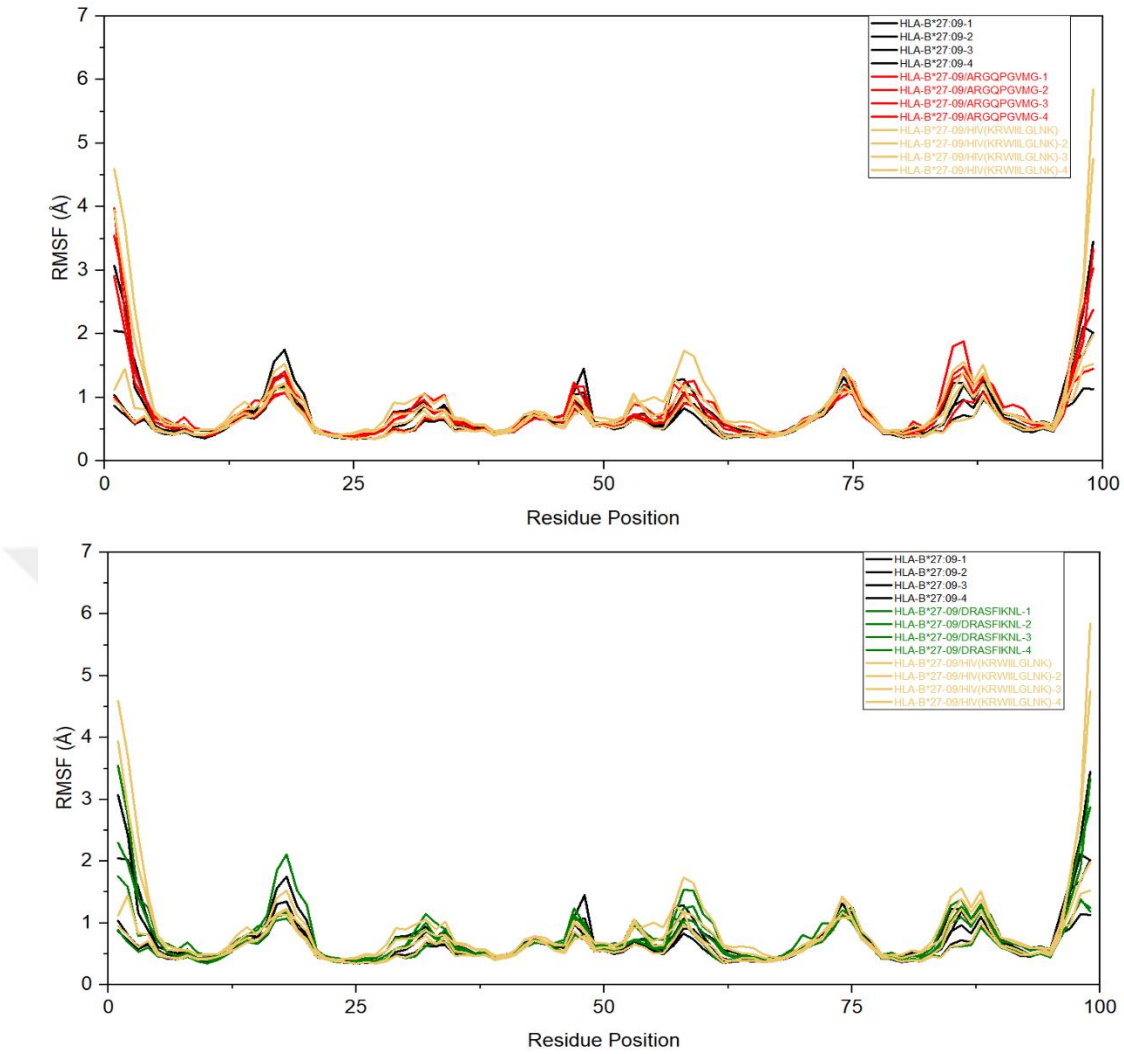

Figure 17. RMSF values of β2m of HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

Tyr59-Trp167 (p1) and Tyr84-Ala139 (pΩ) are opposing residues on the two ends of the protein's binding groove. The distances between these residues were calculated to observe whether the binding groove shows opening and closing movement when the HLA-B27 is bound to the peptide (Figure 18-21). Distances of p1 and pΩ sides were more compatible with each other for the self-peptides and the viral peptides in complex with HLA-B*27:05 (Figure 18 and Figure 20). On the other hand, for HLA-B*27:09 complexes, the distance between the α domains in the p1 and pΩ part of the peptide were fluctuating ranging from 10 to 18 Å for p1

distance and 6 to 14 Å for pΩ distance as well as in the empty state (Figure 19 and Figure 21).

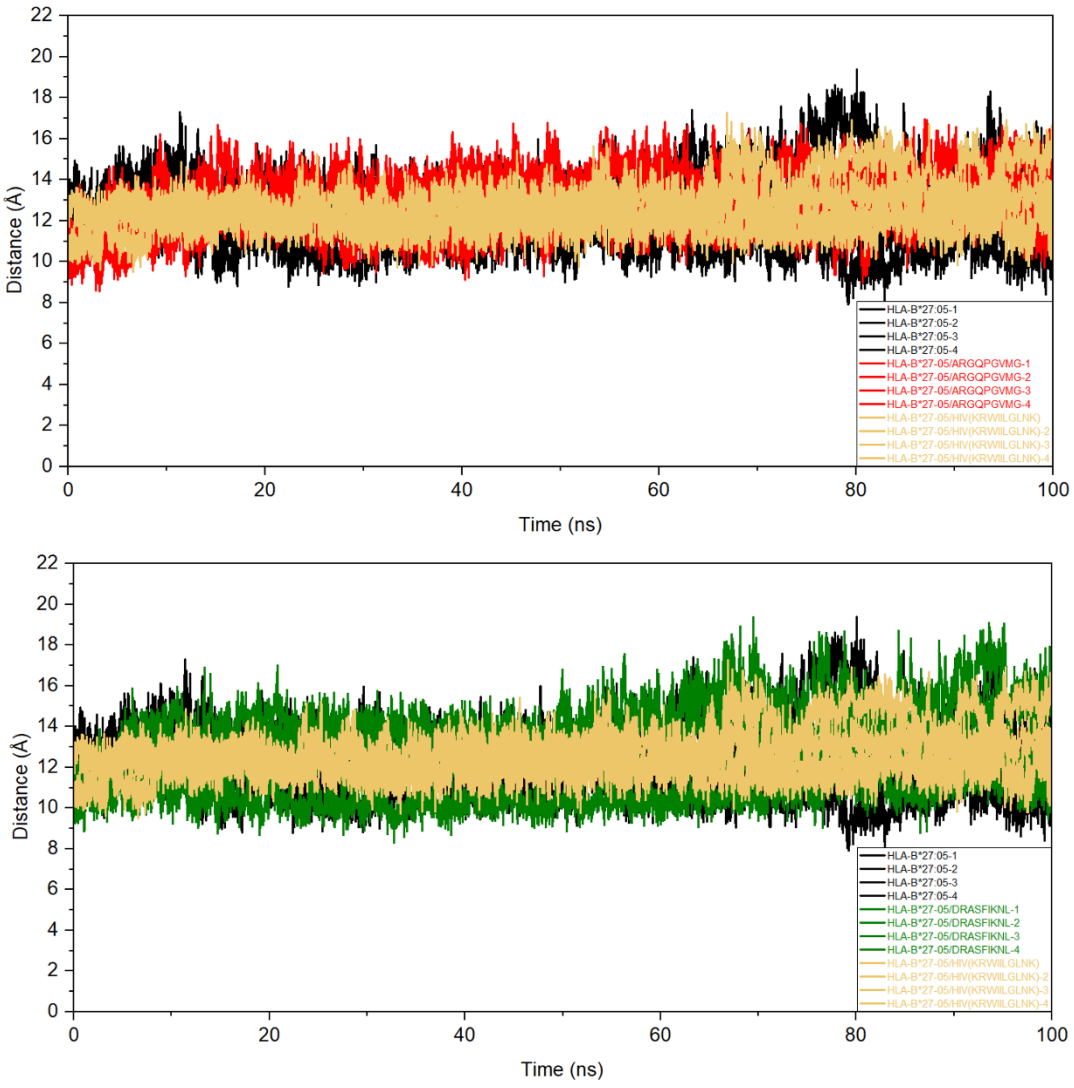

Figure 18. Distance graphs of p1 (Tyr59-Trp167) side of HLA-B*27:05 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

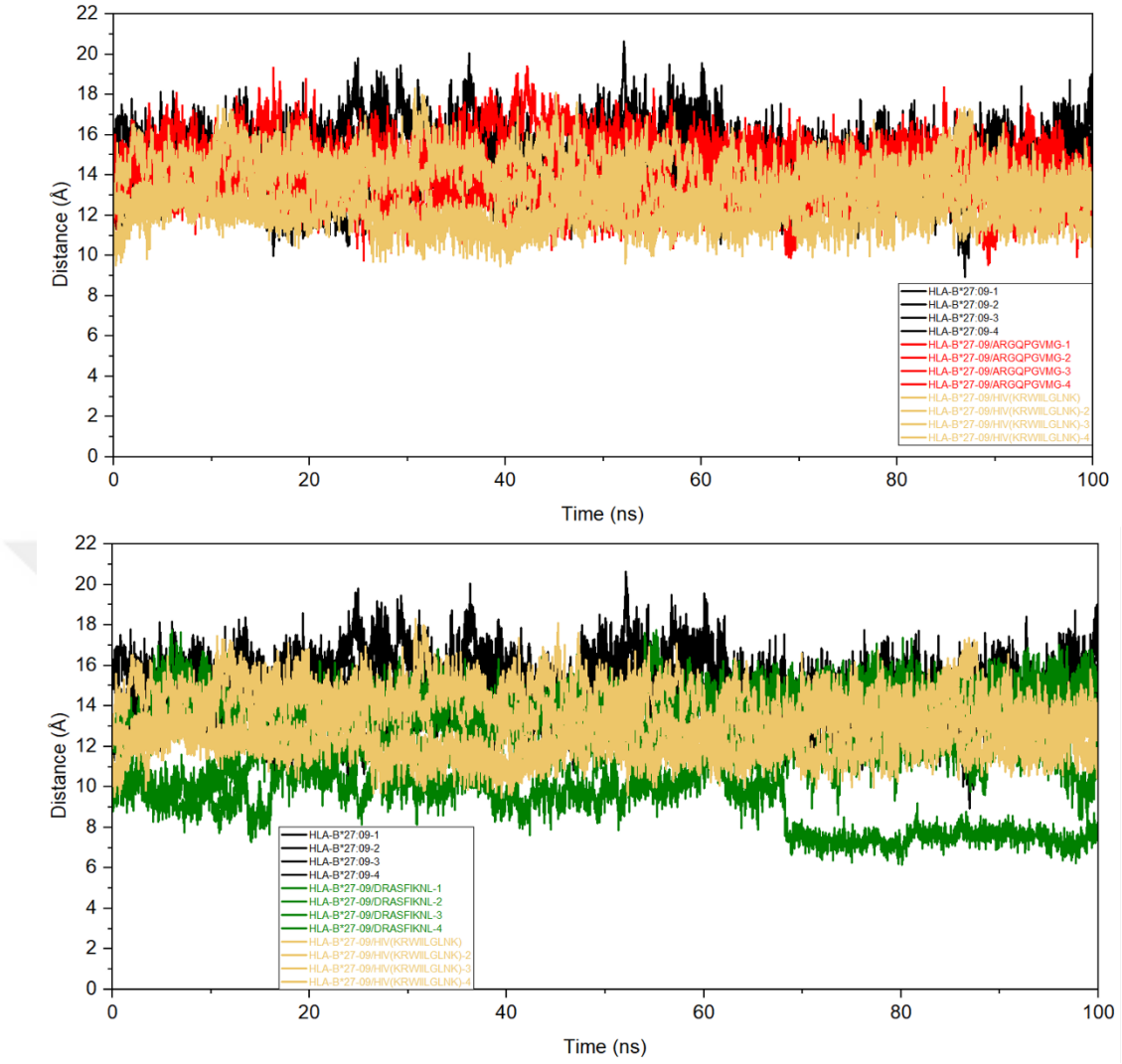

Figure 19. Distance graphs of p1 (Tyr59-Trp167) side of HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

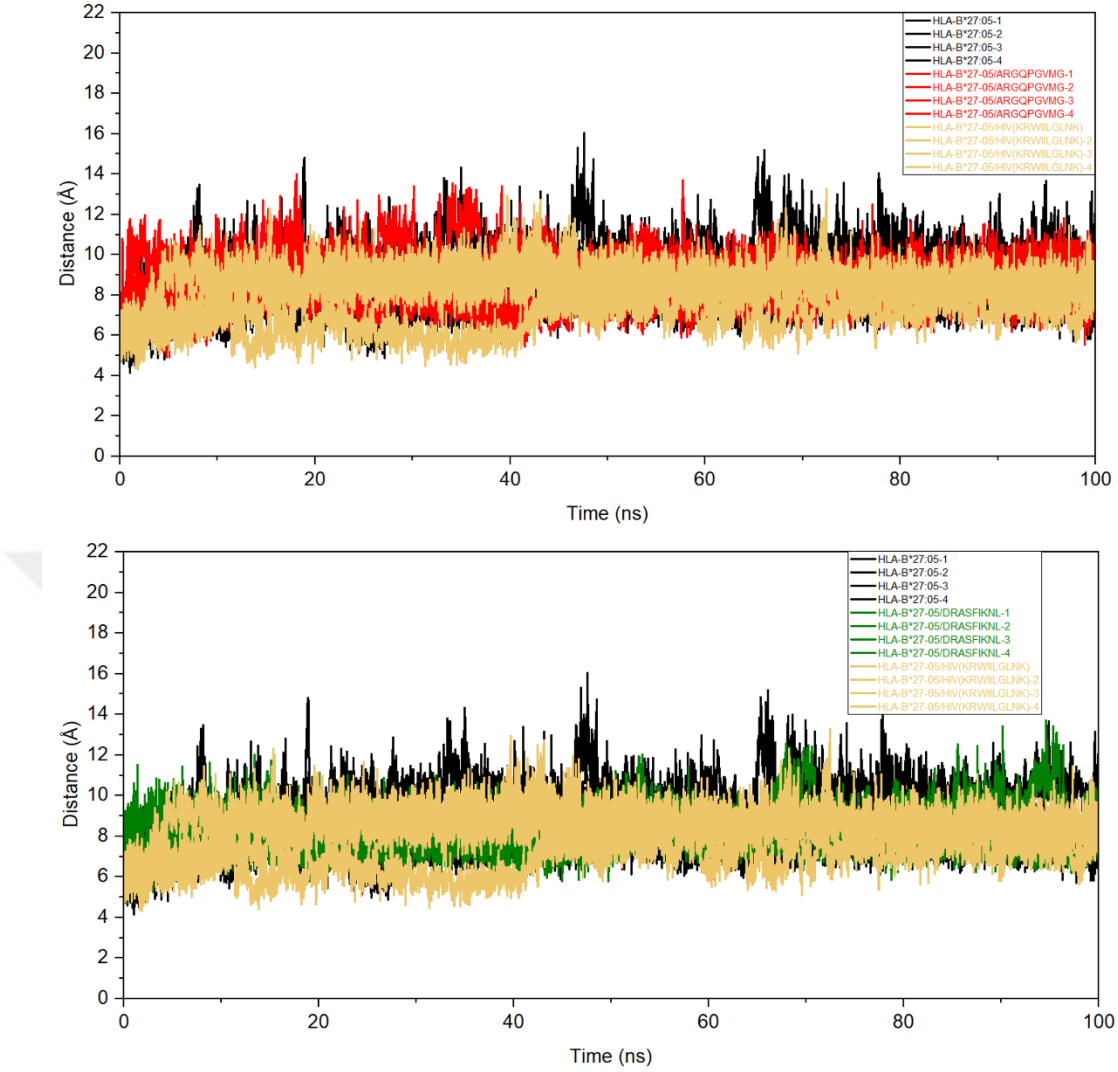

Figure 20. Distance measurements of pΩ (Tyr84-Ala139) side of HLA-B*27:05 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

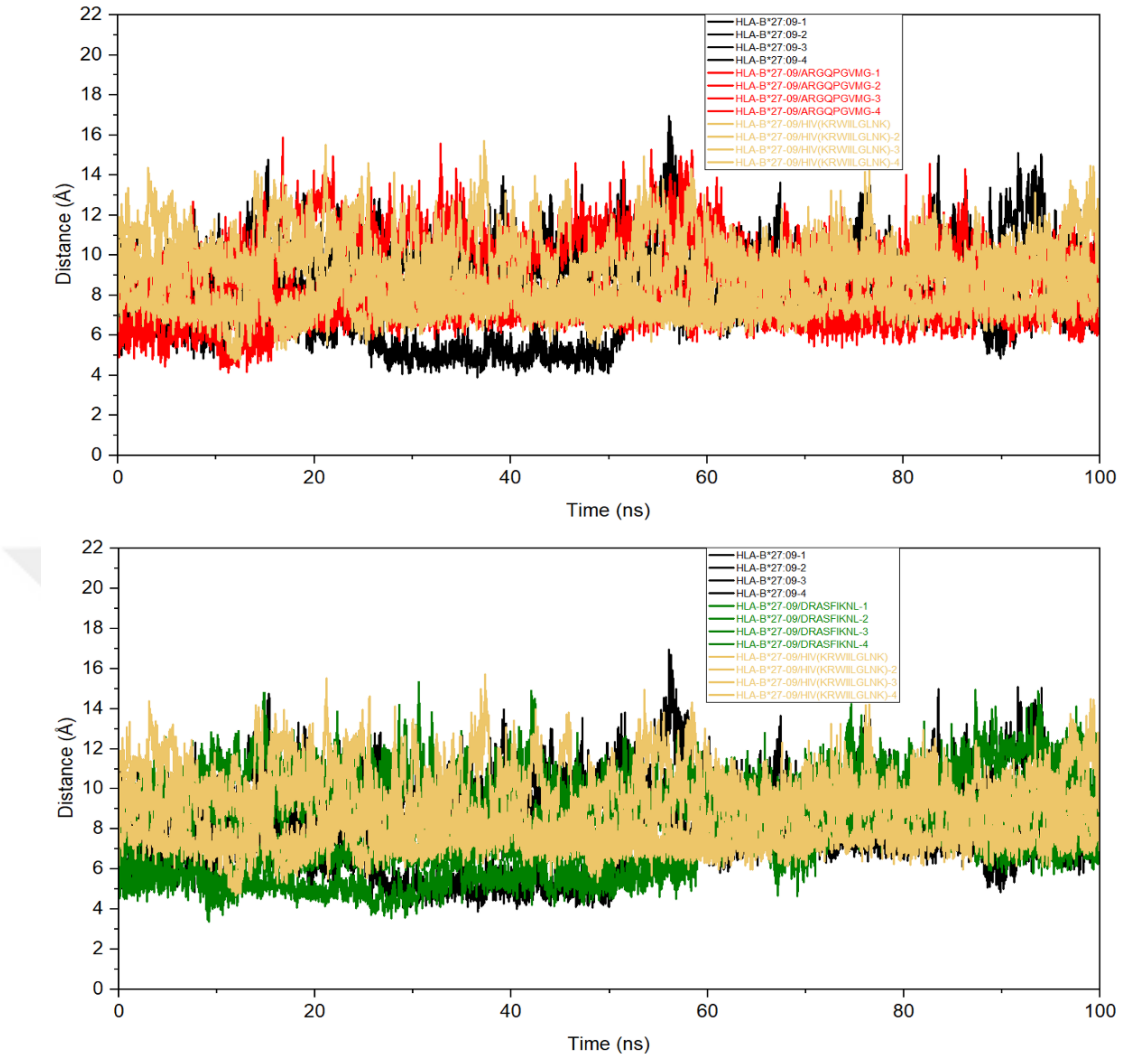

Figure 21. Distance graphs of pΩ (Tyr84-Ala139) side of HLA-B*27:09 with two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL), the KK10 (KRWIILGLNK) peptide, and without peptide from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

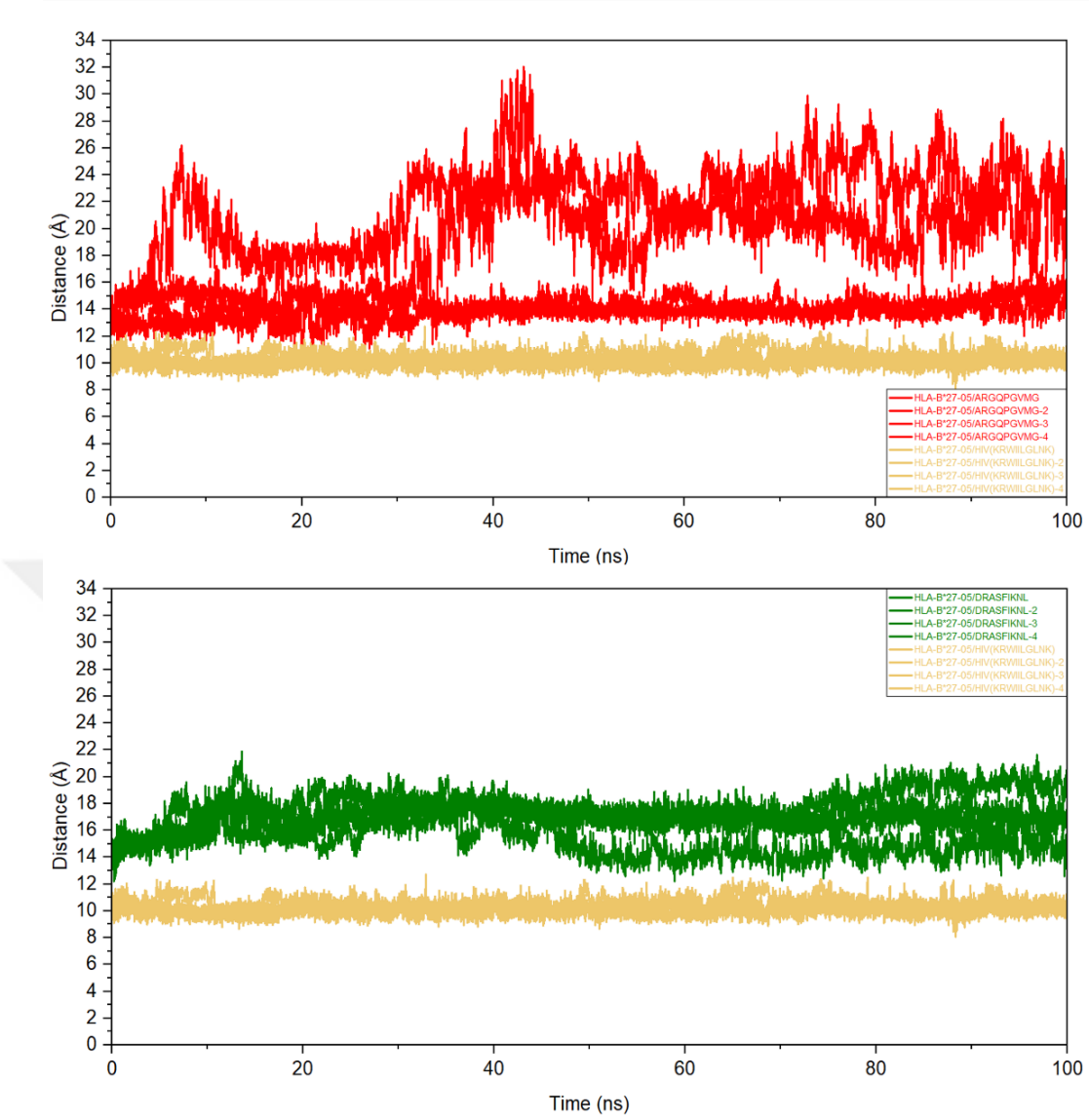

Figure 22. Graphs of the distance between the Cα atom of Asp116 (HLA-B*27:05) and PΩ of two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL) and the KK10 (KRWIILGLNK) peptide from MD trajectories. Twelve independent MD simulations are shown in the graphs.

The mobility of the C-terminus was observed in the snapshots of the conformations. Therefore, by calculating the distance between Asp116/His116 in the binding floor and the C-terminus of the peptide, the difference between the HLA-B27 complexes can be shown by observing how far the end side of the peptide moves (Figure 22 and Figure 23). For peptide ARGQPGVMG, it was observed that the measured distance changes of HLA-B*27:05/ARGQPGVMG in two simulations

were similar to the results of HLA-B*27:05/KK10 complex, but on the other hand, there was a mismatch in the two simulations of the HLA-B*27:05/ARGQPGVMG complex when compared with the KK10 peptide bound to HLA-B*27:05 (Figure 22). Also, the HLA-B*27:05/KK10 complex was stable like the HLA-B*27:05/DRASFIKNL complex.

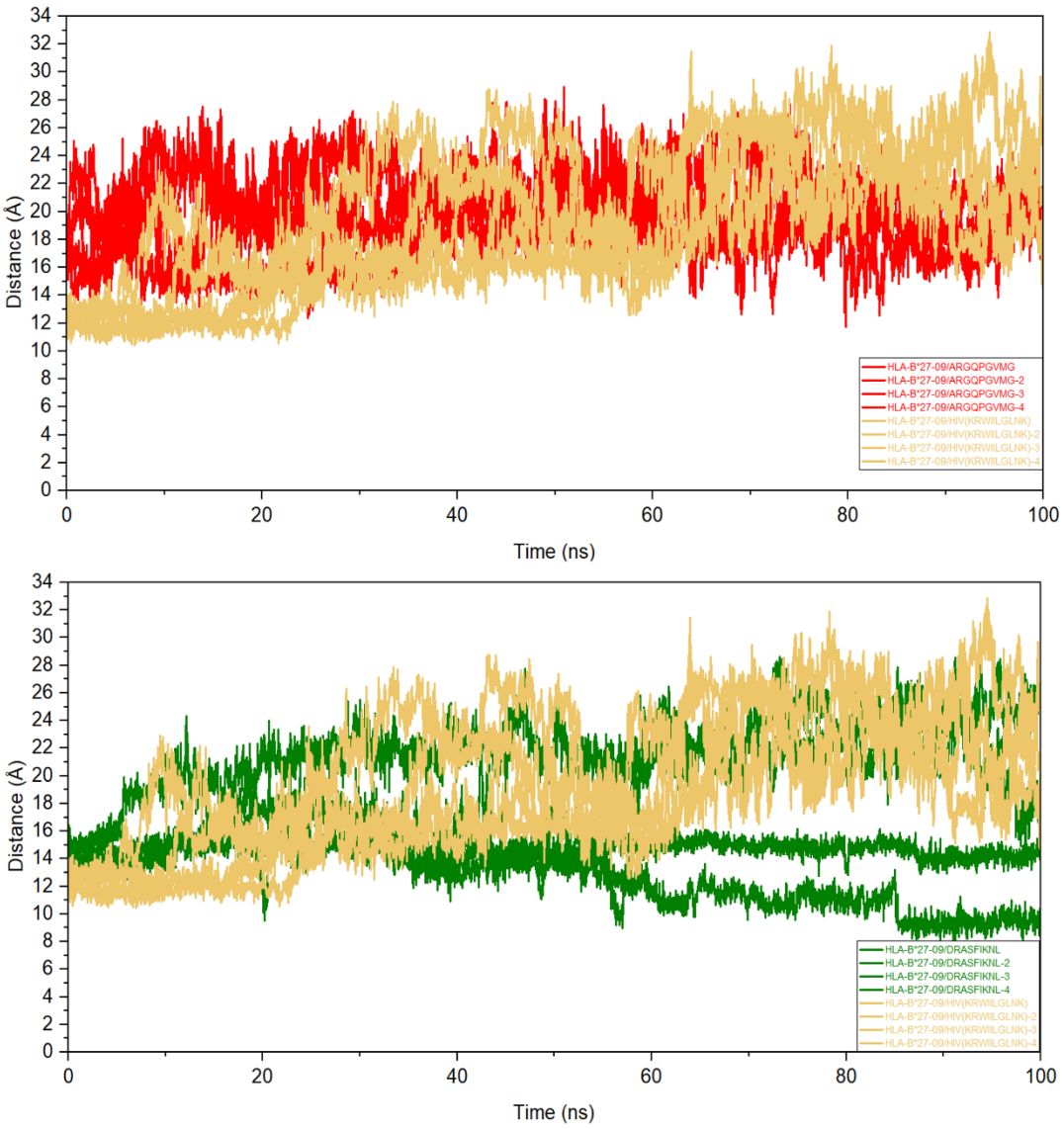

Figure 23. Graphs of the distance between the Cα atom of His116 (HLA-B*27:09) and PΩ of two arthritogenic self-peptides (ARGQPGVMG-DRASFIKNL) and the KK10 (KRWIILGLNK) peptide from MD trajectories. Twelve independent MD simulations are shown in the graphs.

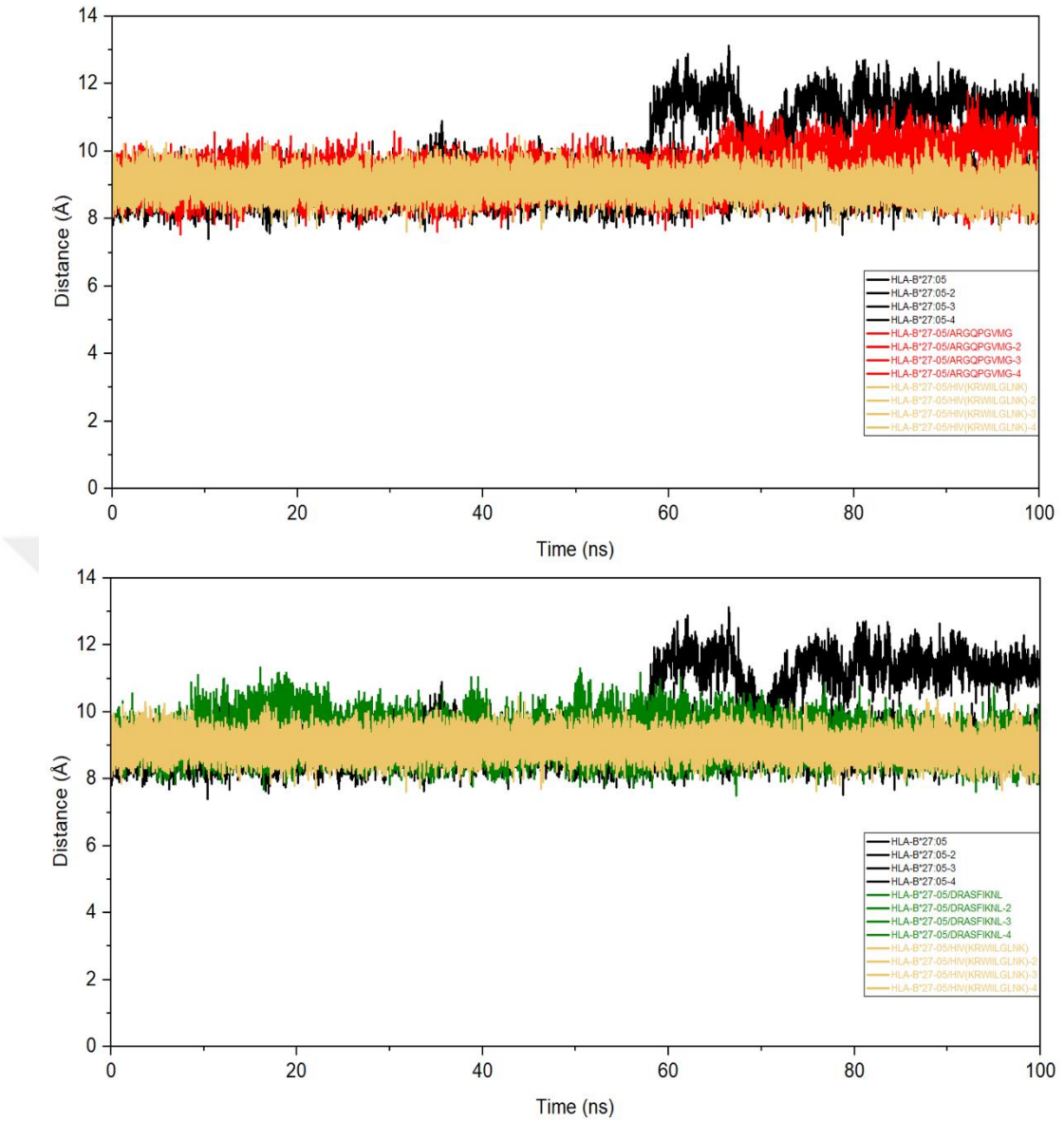

Figure 24. Graphs of the distance between the Cα atom of Asp116 (HLA-B*27:05) and the Cα atom of Trp60 of β2m from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

It was observed that Trp60 of β2m may have a role to increase the stability of the peptide. Asp122 and Gln96 of HLA-B27 are highly conserved and forms H-bond with Trp60 of β2m (84). Therefore, the distances between the Cα atom of Asp116/His116 and the Cα atom of Trp60 of β2m and H-bond distance for the OD1 atom of Asp122 and the NE1 atom of Trp60 of β2m and for the NE2 atom of Gln96 and the O atom of Trp60 of β2m were calculated to understand if there is an effect of

Trp60 in dynamics of antigen presentation (Figure 24-29). For the results of HLA-B*27:09/KK10 complexes, only one of the KK10 simulations resulted in larger fluctuations (4th simulation), others were stable (Figure 25). For the distance between the Cα atom of Asp116 (HLA-B*27:05) /His116 (HLA-B*27:09) and the Cα atom of Trp60 of β2m, the complexes with HLA-B*27:05 seemed to be more stable, but there is no significant difference with the HLA-B*27:09 complexes. The distance range for unbound state of HLA-B*27:05 except for the third parallel simulation was smaller than the distance range of unbound HLA-B*27:09 for H-bond distance between the OD1 atom of Asp122 of the heavy chain of HLA-B27 and the NE1 atom of Trp60 of β2m (Figure 26 and Figure 27). The third parallel simulation of empty HLA-B*27:05 showed high fluctuations after 50 ns (Figure 26). The H-bond distance between the NE2 atom of Gln96 of the heavy chain of HLA-B27 and the O atom of Trp60 of β2m varied between 6 to 8 Å during simulation, except for the second simulation of the HLA-B*27:05/KK10 complex and appears to be consistent and stable for the other three parallel simulations of HLA-B*27:05/KK10 complex (Figure 28). Also, the distance between Gln96 and Trp60 for HLA-B*27:09/KK10 had a smaller range for the three parallel simulations to HLA-B*27:05/KK10 (Figure 29).

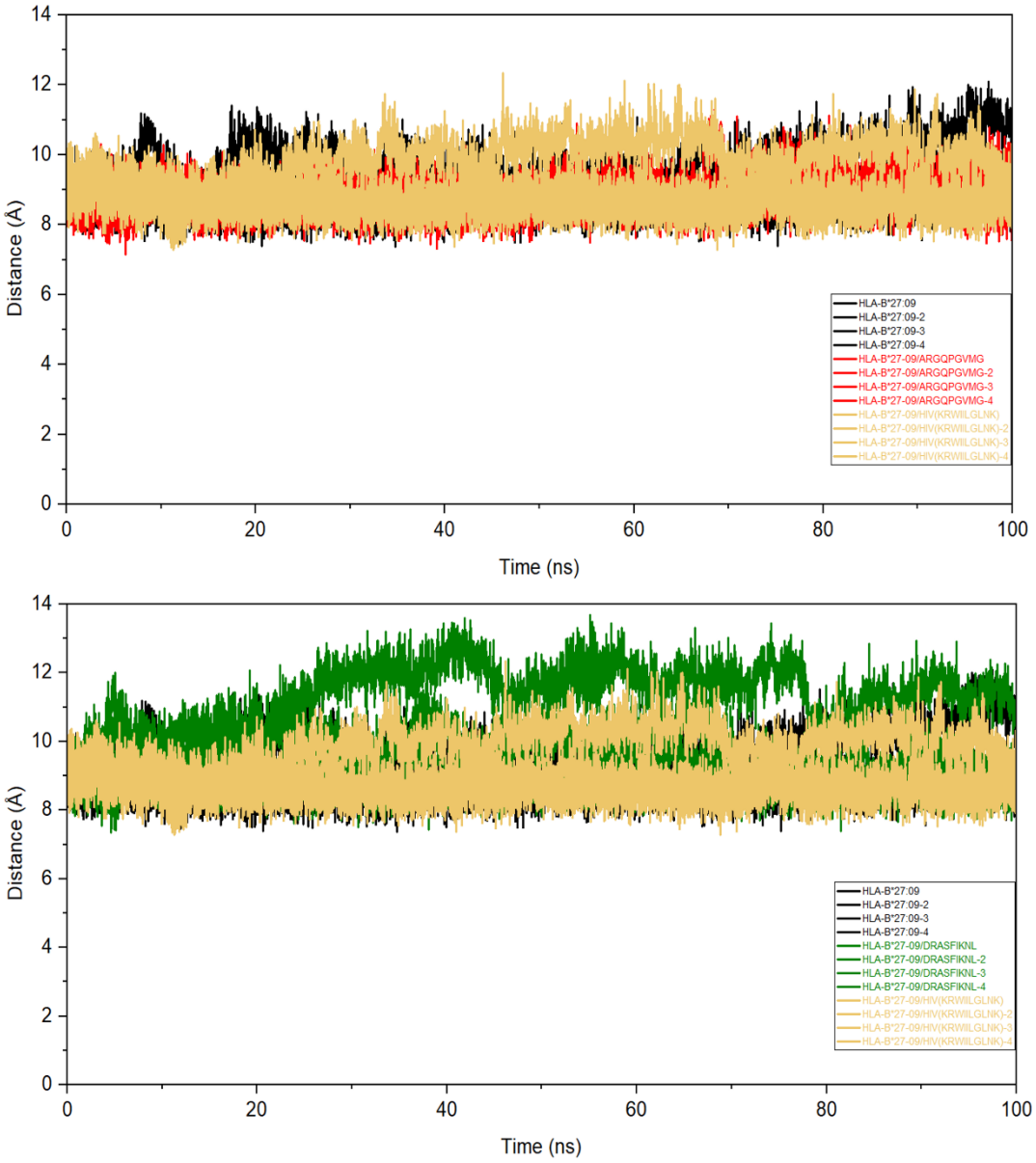

Figure 25. Graphs of the distance between the Cα atom of His116 (HLA-B*27:09) and the Cα atom of Trp60 of β2m from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

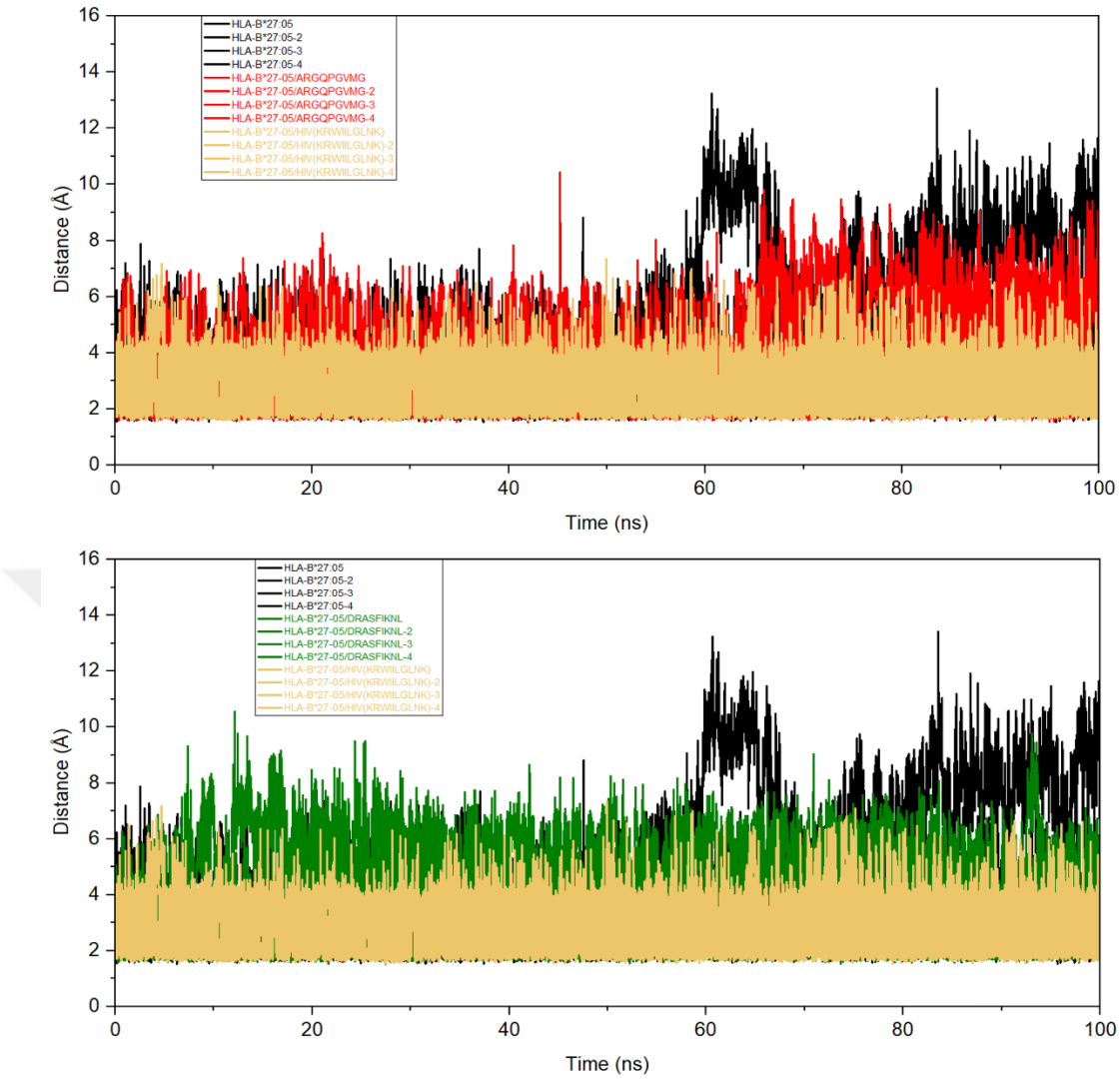

Figure 26. Graphs of H-bond distance between the OD1 atom of Asp122 of the heavy chain of HLA-B*27:05 and the NE1 atom of Trp60 of β2m from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

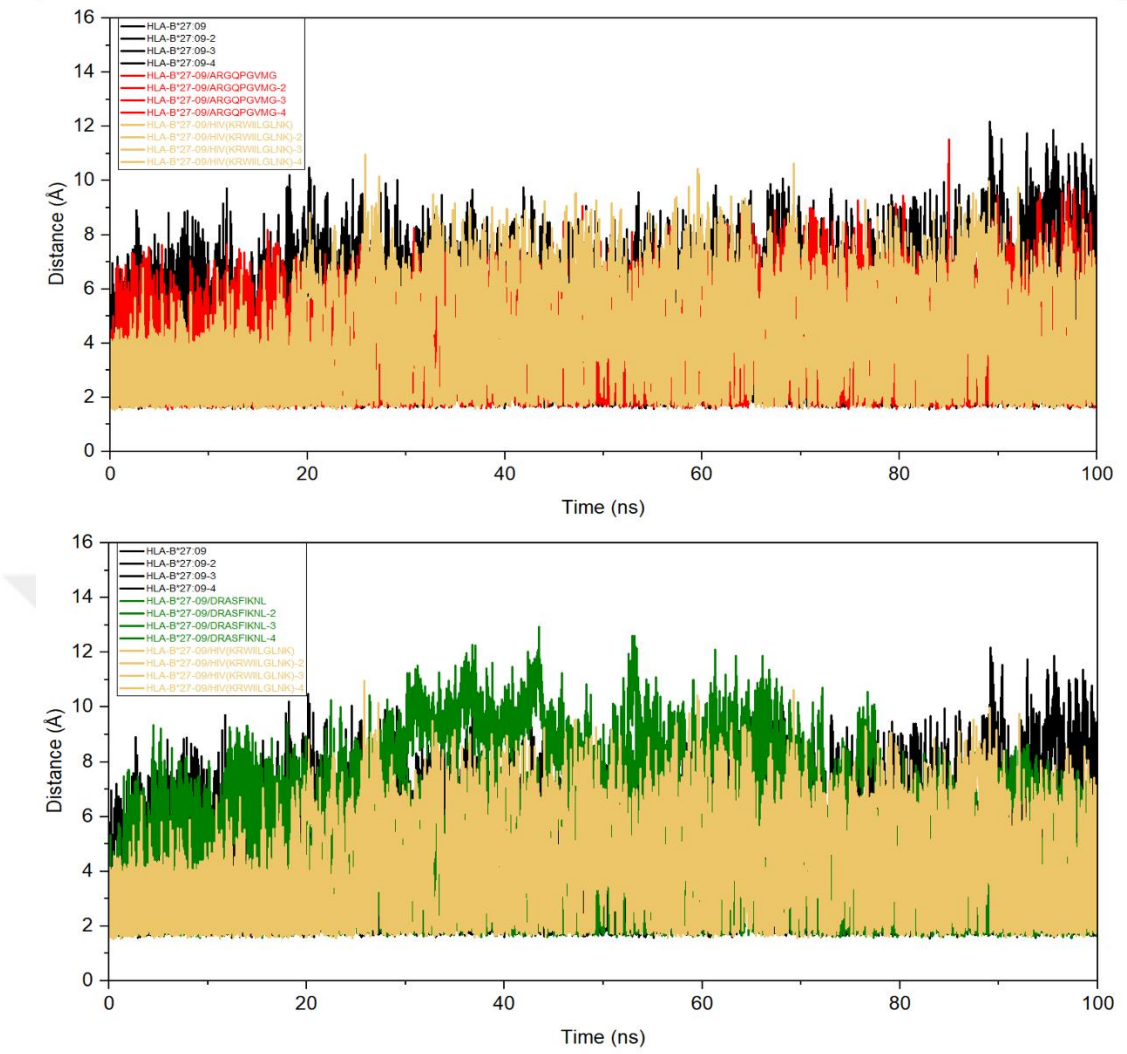

Figure 27. Graphs of H-bond distance between the OD1 atom of Asp122 of the heavy chain of HLA-B*27:09 and the NE1 atom of Trp60 of β2m from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

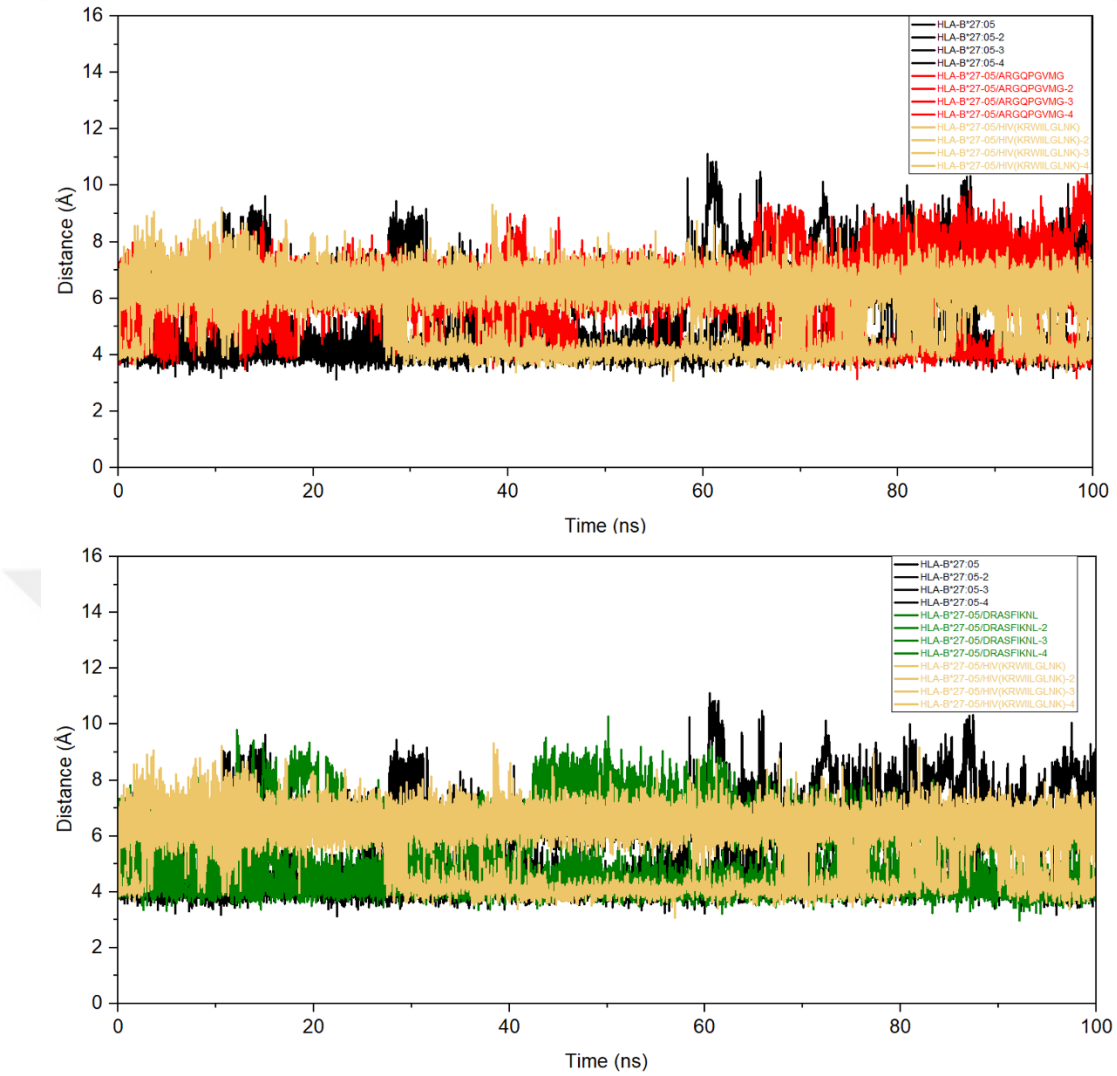

Figure 28. Graphs of H-bond distance between the NE2 atom of Gln96 of the heavy chain of HLA-B*27:05 and the O atom of Trp60 of β2m from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

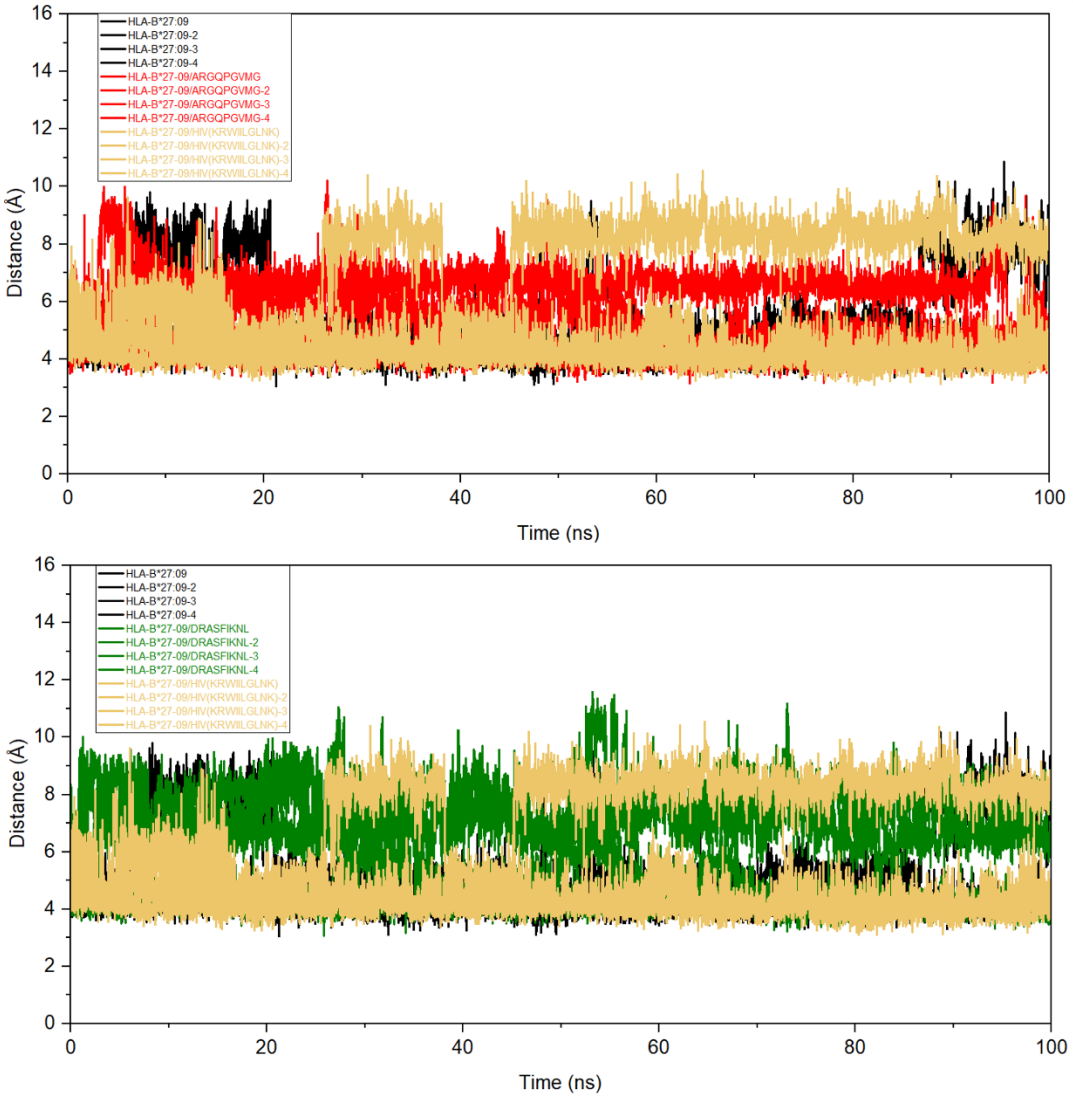

Figure 29. Graphs of H-bond distance between the NE2 atom of Gln96 of the heavy chain of HLA-B*27:09 and the O atom of Trp60 of β2m from MD trajectories. Sixteen independent MD simulations are shown in the graphs.

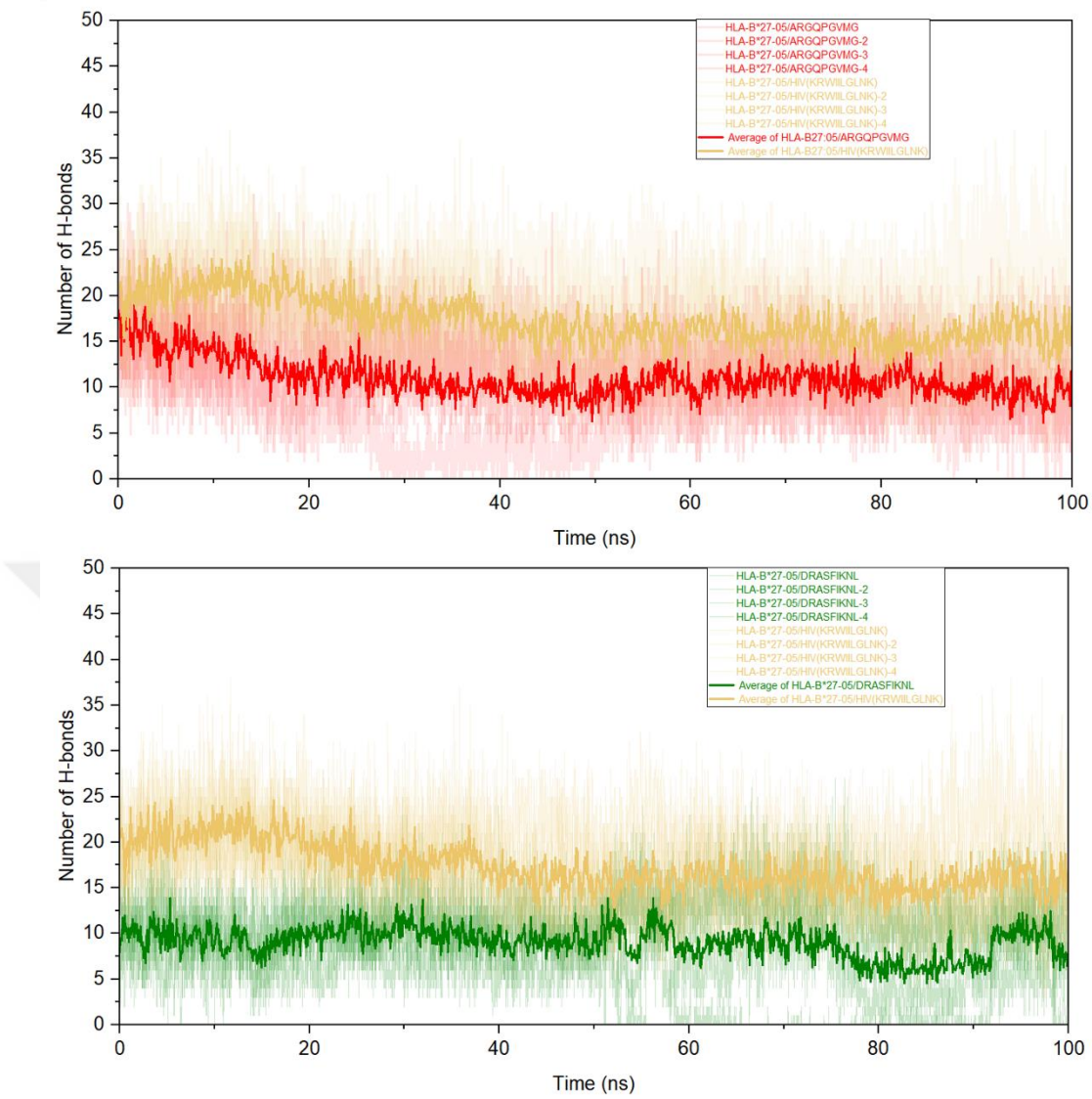

Figure 30. Graphs of the number of H-bond interactions between three peptides and HLA-B*27:05. Twelve independent MD simulations are shown in the graphs.

Since H-bond is a key non-bonded interaction for the stabilization of the protein, the number of H-bonds was evaluated for the HLA-B*27:05/KK10, the HLA-B*27:09/KK10, the HLA-B*27:05/ARGQPGVMG, the HLA-B*27:09ARGQPGVMG, the HLA-B*27:05/DRASFIKNL, and HLA-B*27:09/DRASFIKNL to understand if H-bond plays a role in the stability of HLA-B27/peptide complexes (Figure 30 and Figure 31). Higher number of H-bonds was measured in HLA-B*27:05 complexes than in HLA-B*27:09 complexes. By calculating the average values, the number of H-bonds in the HLA-B*27:09

complexes showed a decrease except for the HLA-B*27:09/ARGQPGVMG complex. It was detected that the KK10 bound complex had more H-bonds than the other two self-peptides bound complexes. Thus, the number of H-bonds for two self-peptides bound HLA-B27 was similar to the HLA-B*27:09/KK10 complex. For the HLA-B*27:05 in complex with ARGQPGVMG peptide, the number of H-bonds was initially close to the complex with the KK10 peptide, but then the H-bond number decreased for the HLA-B*27:05/ARGQPGVMG until 50 ns of simulation. Nevertheless, the HLA-B*27:05/ARGQPGVMG maintained the number of H-bonds more stably throughout the simulation than the HLA-B*27:05/DRASFIKNL.

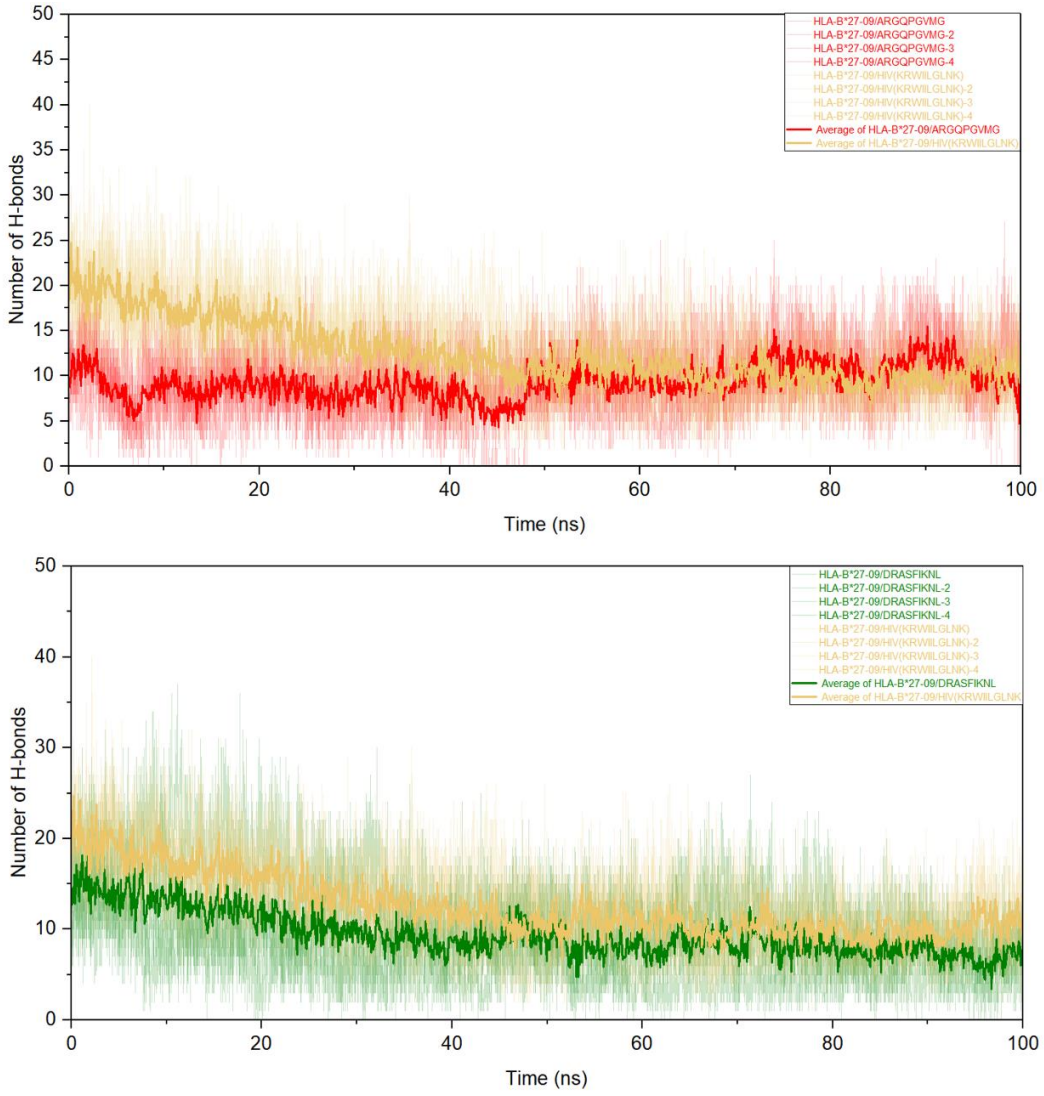

Figure 31. Graphs of the number of H-bond interactions between three peptides and HLA-B*27:09. Twelve independent MD simulations are shown in the graphs.

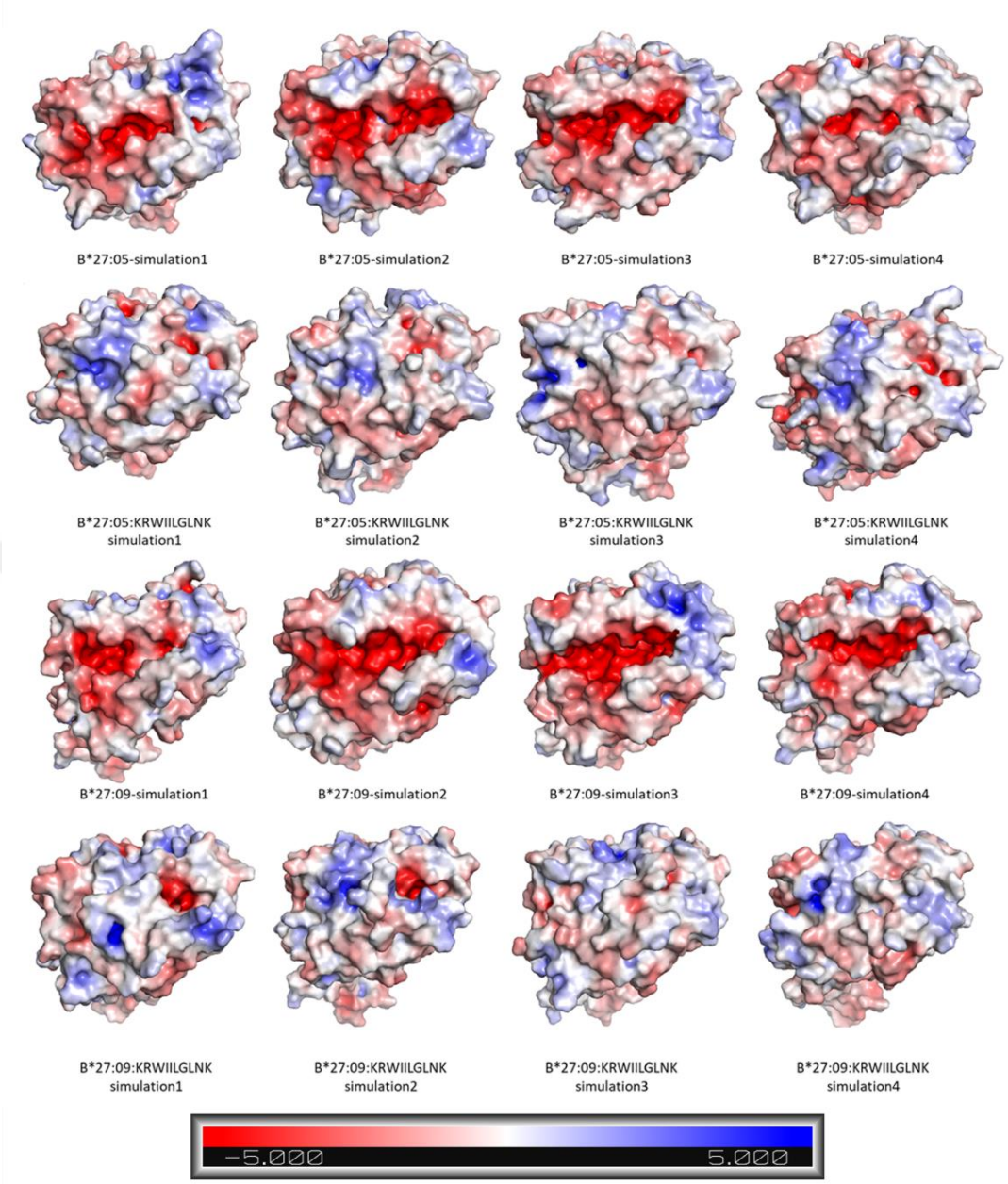

Figure 32. Electrostatic surface potential of HLA-B*27:05 and HLA-B*27:09 bound to the KK10 peptide and unbound form for the final coordinates of 100 ns MD simulation. Range is shown in the corresponding color bar. Red, blue, and white represent negative potential, positive potential, and neutral potential, respectively.

Besides the H-bond, electrostatic interactions also play a role in both the function and stability of the protein. Thus, electrostatic potential surfaces of HLA-B27 subtypes bound and unbound peptides were mapped. The results of electrostatic surface potential measurements are shown in Figure 32, Figure 33, and Figure 34. In

addition to the protein-peptide complexes, the unbound state of HLA-B27 subtypes were visualized. The binding groove with the binding pockets of the empty HLA-B*27:05 and empty HLA-B*27:09 showed more negative potential. When the KK10 peptide was bound to the HLA-B*27:05 protein, the binding pockets to which the p1 region of the peptide was bound demonstrated positive potential. Also, the overall negative potential reduced when the KK10 peptide was bound to the HLA protein. The negative potential for the HLA-B*27:05/ARGQPGVMG decreased slightly compared to the empty state of the HLA protein and the HLA-B*27:09/ARGQPGVMG. The HLA-B*27:09/ARGQPGVMG complex had more neural binding groove. It was noted that when the DRASFIKNL peptide was bound to the HLA-B*27:05 protein, the negative potential was reduced in common in all three parallel simulations of the HLA-B*27:05/DRASFIKNL.

When the HLA-B*27:05 complexes bound to the ARGQPGVMG and DRASFIKNL peptide were compared, there were more identifiable blue (positive potential) and red (negative potential) colors in the binding groove in the complex with the ARGQPGVMG peptide. In addition, the DRASFIKNL bound complex and the KK10 bound complex illustrated more similar results for the HLA-B*27:05 subtype than the ARGQPGVMG bound complex and the KK10 bound complex. When the two self-peptides and the viral peptide were bound to HLA-B*27:09, the HLA-B*27:09/KK10 complex had more positive potential on the binding groove. Although the 3rd simulations of the HLA-B*27:09/KK10 and the HLA-B*27:09/ DRASFIKNL complexes exhibited similar electrostatic potential surfaces, this result was not valid for the other three parallel simulations. The two self-peptides showed similarity within themselves for their HLA-B*27:09 bound complexes.

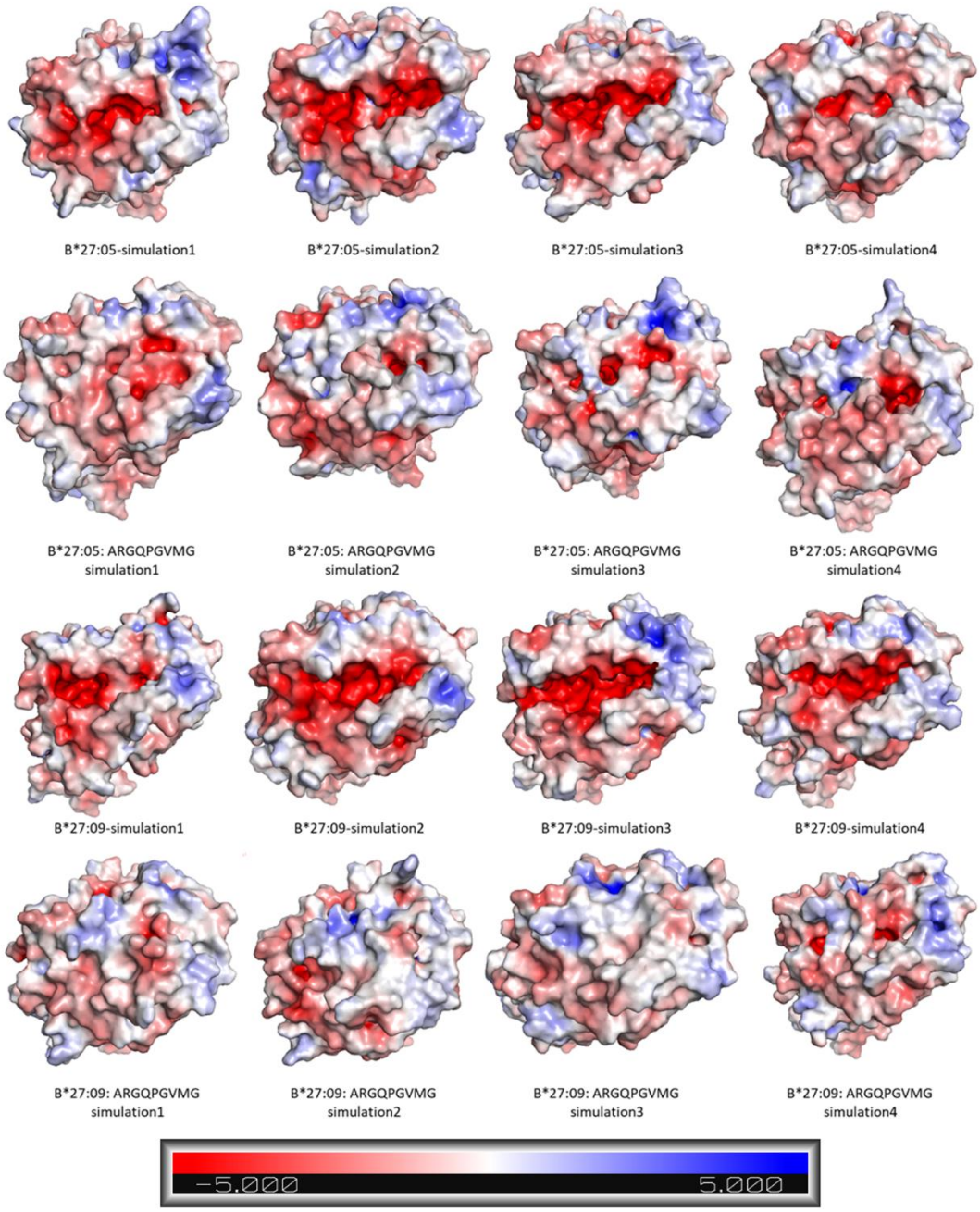

Figure 33. Electrostatic surface potential of HLA-B*27:05 and HLA-B*27:09 bound to the ARGQPGVMG peptide and unbound form for the final coordinates of 100 ns MD simulation. Range is shown in the corresponding color bar. Red color, blue color, and white color represent negative potential, positive potential, and neutral potential, respectively.

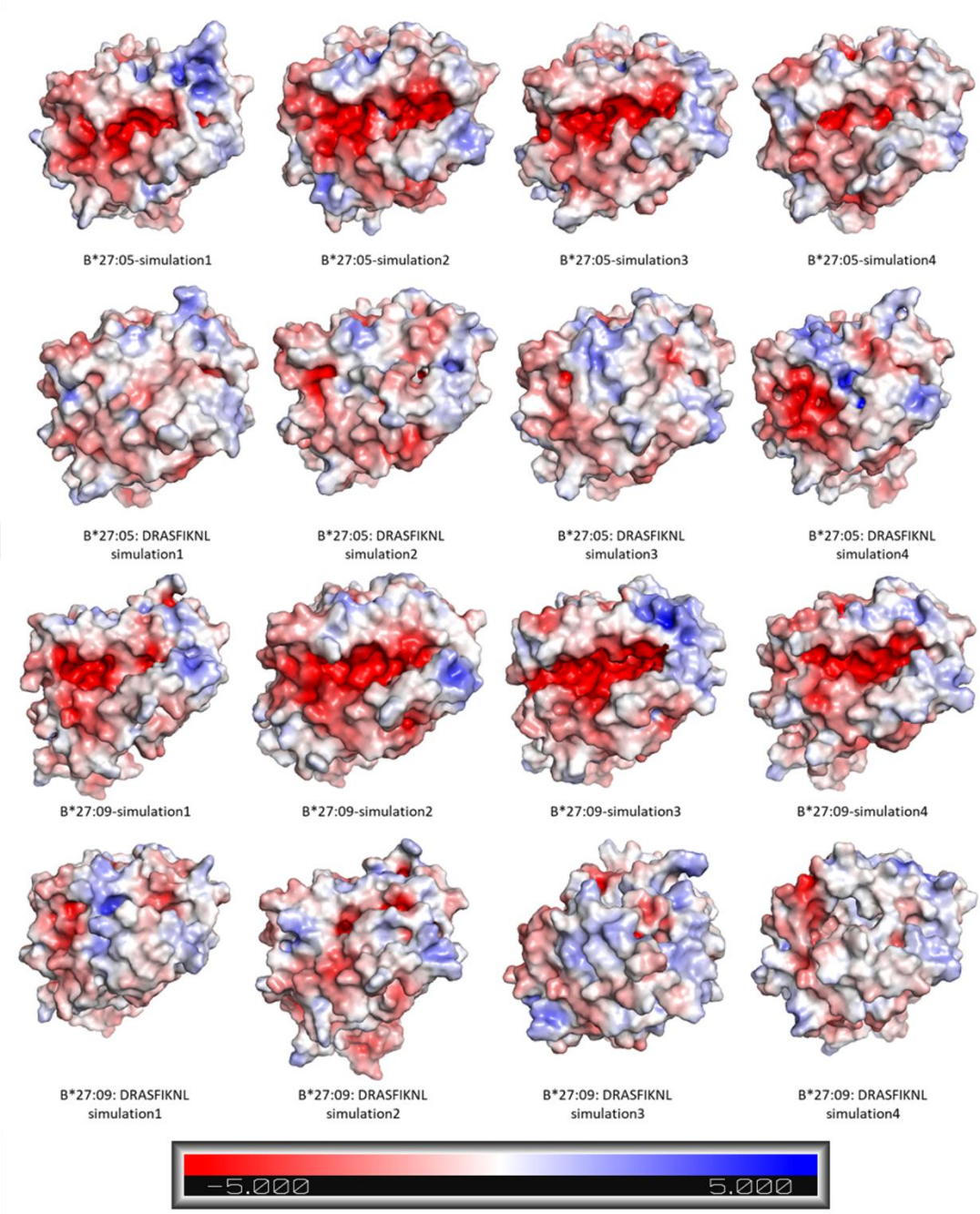

Figure 34. Electrostatic surface potential of HLA-B*27:05 and HLA-B*27:09 bound to the DRASFIKNL peptide and unbound form for the final coordinates of 100 ns MD simulation. Range is shown in the corresponding color bar. Red color, blue color, and white color represent negative potential, positive potential, and neutral potential, respectively.

Table 6. The potential salt bridges between the binding groove of HLA-B27 and peptides obtained from PPCHECK webserver.

Potential Salt Bridges													
		0 ns				50 ns				100 ns			
HLA Subtype	Peptide Sequence	Residue 1 (Binding Groove)		Residue 2 (Peptide)		Residue 1 (Binding Groove)		Residue 2 (Peptide)		Residue 1 (Binding Groove)		Residue 2 (Peptide)	
		Res Num	Res Name	Res Num	Res Name	Res Num	Res Name	Res Num	Res Name	Res Num	Res Name	Res Num	Res Name
B*27:05	ARGQPGVMG	45	GLU	2	ARG	45	GLU	2	ARG	45	GLU	2	ARG
		63	GLU	2	ARG	63	GLU	2	ARG				
	DRASFIKNL	63	GLU	2	ARG	63	GLU	2	ARG	63	GLU	2	ARG
		77	ASP	7	LYS	77	ASP	7	LYS	74	ASP	7	LYS
		116	ASP	7	LYS					77	ASP	7	LYS
		163	GLU	2	ARG								
	KRWIILGLNK (HIV)	45	GLU	2	ARG	45	GLU	2	ARG	45	GLU	2	ARG
		77	ASP	10	LYS	116	ASP	10	LYS	116	ASP	10	LYS
		116	ASP	10	LYS	163	GLU	1	LYS	163	GLU	1	LYS
		163	GLU	1	LYS								
B*27:09	ARGQPGVMG	45	GLU	2	ARG	45	GLU	2	ARG	45	GLU	2	ARG
	DRASFIKNL	76	GLU	7	LYS	163	GLU	2	ARG	No common interaction			
		163	GLU	2	ARG								
	KRWIILGLNK (HIV)	45	GLU	2	ARG	45	GLU	2	ARG	45	GLU	2	ARG
		63	GLU	2	ARG	163	GLU	1	LYS	163	GLU	1	LYS
		77	ASP	10	LYS								
		163	GLU	1	LYS								

Salt bridge interactions between the binding groove of HLA-B27 and the peptides in 3 of 4 simulations are shown in Table 6. The bond between Glu45 (B pocket) of the heavy chain of HLA-B27 and Arg2 of three peptides and Glu63 (B pocket) of heavy chain of HLA-B27 and Arg2 of three peptides appeared to be conserved for most HLA-B27/peptide complexes. Compared to the salt bridges formed with HLA-B*27:09, the DRASFIKNL peptide formed more salt bridges with HLA-B*27:05. It was also observed that the DRASFIKNL peptide lost its bonds with HLA-B*27:09. The positively charged amino acid Lys10 in the C-terminus of the KK10 peptide and the negatively charged Asp116 in the F pocket of the binding groove of HLA-B27 generated a salt bridge and this bond seems to be preserved throughout the simulation. A salt bridge was formed between the positively charged

residue Lys7 of the DRASFIKNL peptide and the negatively charged Asp77 in the F pocket of the binding groove of HLA-B27

The potential preserved bonds throughout the simulation were examined by comparing common H-bonds and hydrophobic interactions at 0, 50, and 100 ns. The common interacting residues were shown according to their binding pocket in Table 7 and Table 8. The residues in both tables were those that change between the H-bond and the hydrophobic interaction throughout the simulation. Considering the hydrogen bonds, the preserved residues are mostly found in the B and F pockets. Glu45 and Glu65 in the B pocket of the binding groove were conserved for H-bonds in the majority of the simulations. The HLA-B*27:05 complexes contained more amino acids forming H-bonds and hydrophobic interactions were compared to the HLA-B*27:09 complexes. Tyr159 in A pocket of the binding groove was seen as common for both HLA-B*27:05 and HLA-B*27:09 complexes. When the interactions formed in HLA-B*27:05 and HLA-B*27:09 complexes were compared, it was determined that amino acids in the E and F pocket of the binding groove, especially those in the F pocket, formed fewer hydrogen bonds and hydrophobic interactions for the HLA-B*27:09 complexes. In the two tables, the residues of the HLA-B*27:05 bound to the DRASFIKNL peptide were similar to the residues of the HLA-B*27:05/KK10 complex compared to the HLA-B*27:05/ARGQPGVMG complex, especially for F pocket. Additionally, E pocket showed a consistent result for the number of hydrophobic interactions of the HLA-B*27:05/KK10 and HLA-B*27:05 complexes

Table 7. LigPlot+ results of the residues of the binding groove of HLA complexes that form H-bond with three peptides.

HLA Subtype	Peptide Sequence	Simulation Number	H-bond					
			0, 50 , 100 ns					
			Pockets					
			A	B	C	D	E	F
B*27:05	ARGQPGVMG	1	Tyr7, Tyr159 (50, 100), Glu163	Glu45, Glu63		Tyr99	Trp147	Lys146
		2		Glu45, Glu63 (0, 50)				
		3				Tyr99 (50, 100)	Trp147	Thr143
		4	Tyr159 (100)	Glu45, Glu63 (0, 50)				
	DRASFIKNL	1	Tyr159 (0, 100), Glu163 (50, 100)	Glu63, Ile66	Lys70		Val152	Glu76, Lys146
		2	Tyr159 (0, 100)			Tyr99 (0, 100)		Asp77
		3	Glu163 (0)	Glu63	Asp74 (50, 100)			Asp 77, Lys146 (100)
		4		Glu63 (0, 50)	Thr73 (50)			Asp77 (0, 100)
	KRWIILGLNK (HIV)	1	Tyr7, Trp167	Glu63, Ile66		Tyr99	Trp147 (50), Val152, Leu156	
		2	Tyr159 (0, 50), Glu163	Thr24, Glu45	His9 (100)	Tyr99	Trp147	Asp77 (0,50), Asp116
		3	Tyr159	Thr24, Glu45, Glu63 (0, 100)	Thr73 (0)	Tyr99 (0, 100)	Trp147	Asp77, Asp116, Lys146 (0, 50)
		4	Tyr159, Glu163	Thr24, Glu45, Glu63	His9 (0, 100)	Tyr99	Trp147	Asp77, Asp116, Lys146 (100)
B*27:09	ARGQPGVMG	1	Tyr7, Trp167	Glu63 (0, 50), Ile66	Thr73 (0, 100)	Tyr99 (50)		
		2	Tyr159 (50, 100)	Thr24, Glu45, Glu63	His9	Gln155 (50)		
		3		Glu45, Glu63 (50, 100)	Thr73 (50, 100)	Tyr99 (0, 50)		
		4		Glu45, Glu63 (50, 100)		Tyr99 (0)		
	DRASFIKNL	1	Tyr159 (50, 100)	Glu63, Ile66				
		2	Tyr159 (100)	Glu63 (50, 100)				Thr143 (0)
		3			Lys70 (100)			
		4	Tyr159 (0), Glu163 (0, 50)				Trp147	Asp77, Thr143 (0, 50)
	KRWIILGLNK (HIV)	1	Tyr7	Glu63 (50, 100), Ile66	His9 (0, 50), Lys70	Gln155	Trp147 (50, 100), Leu156	
		2	Tyr159, Glu163	Thr24, Glu45, Glu63 (0)		Tyr99 (100), Gln155 (100)		
		3	Tyr159, Glu163	Thr24, Glu45, Glu63 (0)		Tyr99 (0)		
		4	Tyr159	Thr24, Glu45, Glu63 (0)		Tyr99 (50, 100)	Trp147 (0)	

Table 8. LigPlot+ results of the residues of the binding groove of HLA complexes that form a hydrophobic interaction with three peptides.

HLA Subtype	Peptide Sequence	Simulation Number	Hydrophobic bond					
			0, 50 , 100 ns					
			Pockets					
			A	B	C	D	E	F
B*27:05	ARGQPGVMG	1	Tyr7, Tyr159, Glu163					Thr80, Tyr123, Lys146
		2	Tyr159	Glu63		Tyr99		
		3	Tyr159			Tyr99	Thr143, Val152, Leu156	Asp77, Lys146
		4	Tyr159, Glu163	Glu63, Cys67		Tyr99		
	DRASFIKNL	1	Tyr159, Glu163		Thr73	Tyr99	Trp147, Val152	
		2	Tyr159, Glu163		Thr73	Tyr99	Trp147, Val152	
		3	Tyr159, Glu163	Glu63, Ile66	Thr73		Val152, Leu156	Glu76, Lys146
		4		Glu63, Ile66, Cys67			Trp147	Glu76, Asp77, Lys146
	KRWIILGLNK (HIV)	1	Tyr7, Trp167	Ile66, Cys67			Thr143, Trp147, Val152, Leu156	Tyr123, Lys146
		2	Tyr7, Tyr159	Ile66, Cys67	His9		Thr143, Leu156	Asp77, Tyr84, Tyr123, Lys146
		3	Tyr7, Trp167	Glu63, Ile66, Cys67	Thr73	Tyr99	His114, Thr143, Val152	Tyr123, Lys146
		4	Tyr7, Trp167	Ile66, Cys67	His9, Thr73		Thr143, Leu156	Lys146
B*27:09	ARGQPGVMG	1	Tyr7, Trp167	Glu63, Ile66, Cys67	Lys70, Thr73	Tyr99		
		2	Tyr7, Tyr159	Ile66, Cys67		Gln155		
		3		Glu63, Ile66	Lys70, Thr73	Tyr99		
		4		Glu63		Tyr99		
	DRASFIKNL	1	Tyr159	Glu63, Ile66			His114	
		2	Tyr7, Tyr159	Arg62, Glu63		Tyr99	Thr143	Asp77
		3	Tyr159		Lys70	Gln155	Leu156	
		4	Tyr159, Glu163				Thr143, Val152, Leu156	
	KRWIILGLNK (HIV)	1	Tyr7	Glu63, Ile66	His9, Lys70, Thr73	Gln155	His114, Trp147, Leu156	
		2	Tyr7	Glu63, Ile66		Ty99, Gln155	Leu156	
		3	Tyr7	Glu63, Ile66	Lys70	Tyr99		
		4	Tyr7	Glu63		Tyr99, Gln155	Trp147, Val152	

4.2.4 Additional analyses

When the peptides were evaluated to the found consensus binding motifs for HLA-B*27:05, the ARGQPGVMG peptide showed two preferred residues and one strong residue which are Ala at p1, Arg at p2, and Val at p7 position of the peptide. Like the ARGQPGVMG peptide, DRASFIKNL had two preferred residues and one strong residue that are Arg at p2, Ile at p7, and Leu at the pΩ position of the peptide. Also, the KK10 peptide exhibited five preferred residues and one strong residue which are Lys, Arg, Leu, Gly, Leu, and Lys at p1, p2, p6, p7, p8, and pΩ positions of the peptide, respectively.

Table 9. Immunogenicity prediction of the peptides specified to HLA-B*27:05.

Peptide	Length	Score
DRASFIKNL	9	-0.11297
ARGQPGVMG	9	-0.16322
KRWIILGLNK	10	0.27526

Table 10. The predictions of the MHC-I binding for HLA-B*27:05.

Peptide	Length	Score	Percentile Rank
DRASFIKNL	9	0.30774	0.64
ARGQPGVMG	9	0.02239	3
KRWIILGLNK	10	0.79704	0.1

Immunogenicity measurements and predictions of the MHC-I binding of the peptides used in this thesis are shown in Table 9 and Table 10. A higher score of immunogenicity shows a higher possibility for immune response. The KK10 and ARGQPVMG had the highest and lowest scores for immunogenicity, respectively. For the binding prediction results, high score exhibits that the peptide can bind to the HLA allele better. Also, a smaller percentile rank indicates higher affinity. The KK10 peptide had the highest score and the smallest percentile rank. Further, the

DRASFIKNL peptide had higher score and smaller percentile rank than the ARGQPVMG peptide.

Regions of similarity were searched for the peptides to identify the presence of similarity between self- and non-self-peptides. According to the BLAST results, the KK10 peptide matched with the RWIILG sequence from retinoic acid early transcript 1E from *Homo sapiens* with %100 identity and 60% query cover. Retinoic acid early transcript 1E includes genes associated with MHC class I (85). Similar regions were found for the DRASFIKNL with bacteria with 100% identity and %100 query cover. Although the ARGQPGVMG peptide matched several hits, the DRASFIKNL peptide demonstrated a higher number of hits with 100% identity and %100 coverage.

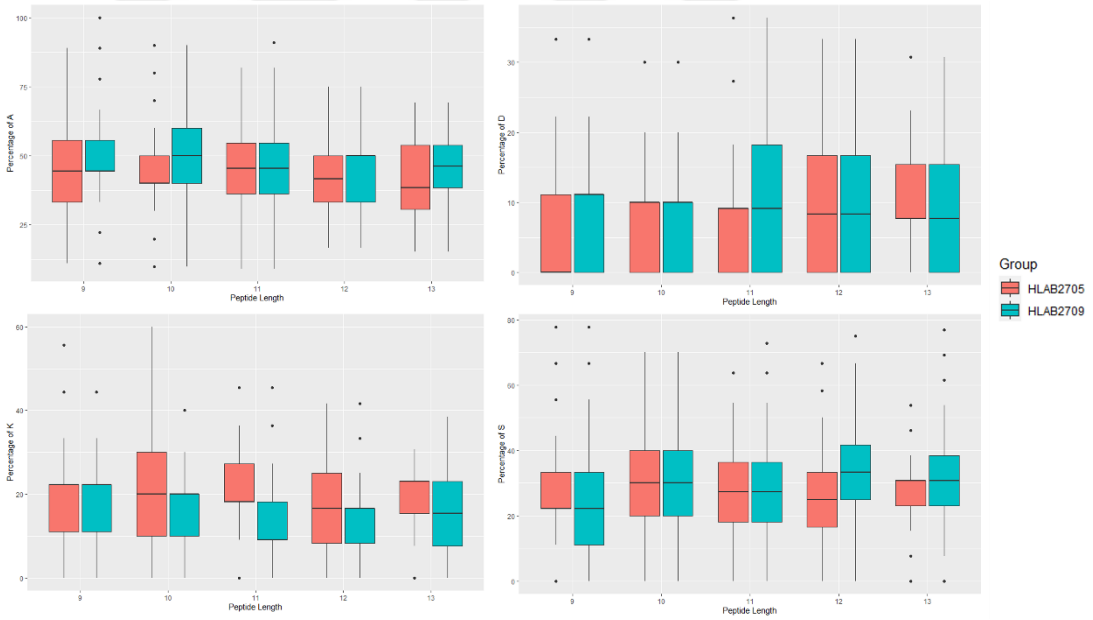

Figure 35. The comparison of the percentages of different kinds of residues between the peptides bound to HLA-B*27:05 and HLA-B*27:09. Nonpolar amino acids = A, negatively charged amino acids = D, positively charged amino acids = K, polar amino acids = S.

The percentages of four clusters based on the hydrophobic and polar model in the RAACBook were calculated for 2132 peptides bound to HLA-B*27:05 and 2276 peptides bound to HLAB*27:09 to check the chemical similarities of the peptidome

of the two subtypes by using a simplified representation (Figure 35). The percentage of K presented a higher percentage range for the peptides with 10, 11, and 12 amino acid lengths bound to HLA-B*27:05. The overall percentages did not show a significant difference in average and range for the two HLA-B27 subtypes. In addition, two arthritogenic self-peptides and the KK10 peptide, their amino acid sequence was reduced using RAACBook (Table 11). p5, p6, and p7 for ARGQPGVMG and p6, p7, and p8 for the KK10 peptide were similar, respectively. Also, the ARGQPGVMG and KK10 peptides were same at p8 and p9 positions. The DRASFIKNL and the KK10 peptides were comparable at p2 and p3 positions as in their N-termini. They had positively charged K residue around their C-termini.

Table 11. The reduced sequences for two arthritogenic self-peptides and the KK10 peptide.

Peptide Sequence	Peptide Length	Reduced Sequence
ARGQPGVMG	9	AKSSASAAS
DRASFIKNL	9	DKASAAKSA
KRWIILGLNK	10	KKAAAASASK

5 DISCUSSION

Ankylosing spondylitis (AS) is one of the autoimmune diseases that cause inflammation on the skeleton, especially for joints in the spine. Several studies have indicated the strong association between HLA-B27 and AS disease in the pathogenesis of AS because of the role of MHC class I in antigen presentation. There are many HLA-B27 subtypes based on differences in the protein sequence, which results in different AS associations. HLA-B*27:05 and HLA-B*27:09 subtypes differ in only one amino acid at position 116 (D116H) of the heavy chain of HLA-B27, although HLA-B*27:05 is the most observed subtype associated with AS and HLA-B*27:09 is not found to be related to the disease. The arthritogenic peptide hypothesis based on the molecular mimicry between self-peptide and non-self-peptide is one of the several hypotheses aimed to explain the association between HLA-B27 and AS. In this thesis, the structure of the HLA-B27:05 and HLA-B*27:09 complexes with two arthritogenic self-peptides and a viral peptide was investigated to elucidate the presence of a common structural behavior between self-peptide and viral peptide on HLA-B*27:05 for antigen presentation mechanism based on the arthritogenic peptide hypothesis by using molecular docking and MD simulations for each bound and unbound HLA-B27 complexes.

The key finding of this thesis was similar structural behavior of the DRASFIKNL self-peptide and the KK10 (KRWIILGLNK) peptide in their complexes with HLA-B*27:05 compared to the ARGQPVMG self-peptide and the KK10 peptide. By applying reduced alphabet method to the three peptide sequences, it was observed that the DRASFIKNL and the KK10 peptides had more common residues based on their chemical properties. Although the amino acids in the middle position of the KK10 and ARGQPGVMG peptides are compatible, the similarity of charged amino acid sequences at the termini of the peptide resulted in similar structural behavior of the peptides for their complexes with HLA-B27 subtypes according to trajectory analyses.

Initially, the HLA-B*27:05 and HLA-B*27:09 subtypes were docked with three peptides to be used in MD simulations. The predicted protein-peptide complexes obtained from the HADDOCK presents a snapshot of the HLA-B27/peptide complexes. The HADDOCK scores suggested that there was a difference in binding between HLA-B*27:05 and HLA-B*27:09 complexes for the DRASFIKNL and KK10 peptides, while the HADDOCK scores of the ARGQPGVMG peptides with HLA-B*27:05 and HLA-B*27:09 were similar. Also, the ARGQPGVMG peptide bound to HLA-B*27:05 and the KK10 peptide bound to HLA-B*27:09 were located close to the binding groove, whereas the KK10 and the DRASFIKNL peptides were bulging out on HLA-B*27:05 peptide. Thus, the superimposition of the structures supports the similar behavior of the DRASFIKNL self-peptide and the KK10 peptide as the viral peptide.

The structures of the complexes were analyzed further by using MD simulations. The visualization of conformations offers to observe the movement of the peptides throughout the simulation. The KK10 peptide on HLA*B27:05 was the most stable peptide for all 4 simulations. The unbound states of HLA-B*27:05 and HLA-B*27:09 were also simulated. The binding groove structure of empty HLA*B27:09 was less fluctuating than empty HLA-B*27:05. The RMSD of empty HLA-B*27:05 were higher than empty HLA-B*27:09 for independent two MD simulations by Abualrous *et al.* (2015) study (25). When one of the three peptides bound to HLA-B*27:05, the α-2 domain of HLA-B*27:05 demonstrated stability. The MD study conducted by Hein *et al.* (2018) also showed that the residual fluctuations, measured via RMSF, were decreased when a peptide bound to HLA-B*27:05 (41). The DRASFIKNL peptide contributed the stability of HLA-B*27:05 on a large scale compared to the effect of the ARGQPGVMG peptide and the DRASFIKNL peptide increased the mobility of α domains of its complex with HLA-B*27:09. The RMSD values were calculated to examine the stability of bound and unbound HLA-B27 complexes and measurements were supported the conformation snapshots of the complexes. Also, the RMSD values of the HLA-B27:05/DRASFIKNL complex were more compatible with the values of the HLA-B*27:05/KK10 when compared to the HLA-B*27:05/ARGQPGVMG complex, which suggests the similar behavior

opinion. On the other hand, the RMSF calculations were unable to discriminate between the effect of the two arthritogenic self-peptides.

Tyr59-Trp167 (p1) and Tyr84-Ala139 (pΩ) are located on the α-1 and α-2 domains as opposing residues on the ends of the binding groove of HLA-B27. The distances between these residues were calculated to understand the opening and closing movement of the structure of HLA-B27. The HLA-B*27:05/peptide complexes exhibited less opening and closing movements with narrower distance range for p1 when compared to the pΩ side, which shows that while the p1 side of the HLA-B*27:05/KK10 complex is more stable than p1 side of the two self-peptides bound to HLA-B*27:05, measured distances for pΩ side were similar for all three peptide complexes with HLA-B*27:05. Therefore, the analysis of these distances was consistent with the conformation snapshots, which shows the effect of pΩ side movement on the α domains of HLA-B27 protein. The mobility of the C-terminus of the peptide was analyzed further to compare how far the C-terminus moves from the binding groove by calculating the distance between the Cα atom of Asp116 (HLA-B*27:05) /His116 (HLA-B*27:09) and PΩ of the peptides. The HLA-B*27:05 complexes with the KK10 and DRASFIKNL peptide were more stable than other complexes.

The effect of peptides on β2m was observed by calculations of the distance between the Cα atom of Asp116 (HLA-B*27:05) /His116 (HLA-B*27:09) and the Cα atom of Trp60 of β2m, H-bond distance between the OD1 atom of Asp122 of the heavy chain of HLA-B27 and the NE1 atom of Trp60 of β2m, H-bond distance between the NE2 atom of Gln96 of the heavy chain of HLA-B27 and the O atom of Trp60 of β2m. These distances were used since Trp60 of β2m showed an important role in the increase of the stability of the peptide in the NMR study (2013) by Beerbaum and colleagues (84). For the distance between the Cα atom of Asp116 (HLA-B*27:05)/ His116 (HLA-B*27:09) and the Cα atom of Trp60 of β2m, there was no significant result that shows the effect of the peptide in the complex stability. For the other two bond distances, only the KK10 complexes with HLA-B27 subtypes showed an effect on β2m except for the H-bond distance between the NE2 atom of

Gln96 of the heavy chain of HLA-B27 and the O atom of Trp60 of β2m when compared to the results of two self-peptides. Besides β2m calculations, the number of H-bonds was estimated for the role of H-bond in the stabilization of protein structure. A higher number of H-bonds for HLA-B*27:05 complexes supports the conformation snapshots that showed the movement of the three peptides was stable on HLA-B*27:05 rather than on HLA-B*27:09. In addition, the similarity of the number of hydrophobic interactions for the E pocket of HLA-B*27:05/KK10 and HLA-B*27:05/ DRASFIKNL complexes indicated the effect of the E pocket for the presentation mechanism in the pathogenesis of AS disease. Another analysis was the electrostatic potential surface to observe the change in electrostatic interactions. The empty states of two HLA-B27 subtypes exhibited high negative potential on the binding groove. However, these negativities increased when the HLA-B27 subtypes were bound to the peptides for almost all simulations. Since the binding groove of HLA-B27 is mostly negative, HLA-B27 may prefer peptides with a higher number of positively charged residues. The electrostatic potential surfaces of the HLA-B*27:05/DRASFIKNL and HLA-B*27:09/KK10 complexes were fit compared to the HLA-B*27:05/ARGQPGVMG complex. These results may show the similarity of the effect of a self-peptide and viral peptide on HLA-B27 subtype, which may result in a similar CTL response.

Potential salt bridge interactions, H-bond interactions, and hydrophobic interactions were investigated to observe the conserved residues throughout the simulation. Positively charged arginine in the 2nd position of the peptides formed a salt bridge with negatively charged glutamic acid at the 45th and 63rd position of the heavy chain, and this strengthens the bond between the peptide and HLA protein. Although in some studies, they come through other positively charged amino acids such as Lys in some studies, the second position of the HLA-B27:05 peptide generally prefers arginine (86,87). In the study of Narzi *et al.* (2008), Glu45/Glu63 residues showed an important role in decreasing the interaction energy of Arg2 of the peptide and stabilize the peptide (87). Also, according to the studies of Chen *et al.* (2017) and Colbert *et al.* (2013), the B pocket of the binding groove folds slowly increasing the possibility of misfolding (14,34). A peptide binding to Glu45 or Glu63

may reduce the misfolding problem. Two positively charged Lys10 of the KK10 peptide and Lys7 of the DRASFIKNL peptide also formed salt bridges with Asp77 and Asp116 of the F pocket in the binding groove, respectively. It seems that although Asp77 was bound to Lys7, it can also be affected by His116 in the course of protein peptide interaction. It might result in the C terminus to distancing from the binding floor. While there are Asp74, Asp77, and Asp116 triple groups in the F pocket of HLA-B*27:05, which provides a long-term interaction for the peptides with positively charged residues at their C-termini compared to HLA-B*27:09 including positively charged His116 (88). Compared to the salt bridges formed with HLA-B*27:09, the DRASFIKNL peptide formed more salt bridges with HLA-B*27:05 and formed salt bridges at the beginning of the simulation for the HLA-B*27:09/DRASFIKNL disappeared throughout the simulation. The preserved residues were found mostly in the B and F pockets of the binding groove. The study of Nguyen *et al.* (2021) also exhibited the key role of B and F pockets in the stabilization of HLA protein (12). The analysis of H-bond and hydrophobic interactions also illustrates the similar behavior of the DRASFIKNL and KK10 peptide bound to HLA-B*27:05, especially for the F pocket while this behavior was not found for the HLA-B*27:09 complexes. According to the interaction analyses, the KK10 peptide had the highest number of interactions with HLA-B*27:05 which was consistent with the results of immunogenicity analysis and the MHC-I binding prediction. The KK10 peptide had the highest score and the smallest percentile rank which illustrates the KK10 peptide can bind to the HLA-B*27:05 better. The DRASFIKNL peptide had higher score and smaller percentile rank than the ARGQPVMG peptide. The DRASFIKNL peptide had more possibility to bind to HLA-B*27:05 subtype better.

In this thesis, some limitations were encountered. TCR should be considered for the integrity of the TCR/pMHC complex, and it should be evaluated with the interaction analysis and the electrostatic potential surface to check if these peptides are able to bind to TCR. However, TCR has polymorphism. Additionally, which TCR will be used for the docking and MD methods is a challenging. Experimental studies are needed for patient-specific TCR repertoire. Besides the arthritogenic

peptides from AS patients, the sequence of TCR can also be obtained to use in the future study. 100 ns of simulations is short for big molecules. Therefore, additional repeat runs of MD simulations are needed to understand and further verify the structural and dynamical differences. Also, a non-arthritogenic peptide should be used to as a control and compare with the arthritogenic peptides in the further study.

To sum up, even if the sequence similarity is low for the studied peptides, the similarity of charged amino acids around termini in the DRASFIKNL and the KK10 peptides may cause them to have similar effects on the stability of the protein peptide complex. Having similar structural features may cause an immune response to the self-peptide recognized by the TCR. In addition, the fact that these peptides come from HLA-B27 positive AS patients shows that it may cause misidentification of the self-peptide. Also, the existence of bacterial sequence hits of the BLAST run is consistent with the possibility of the DRASFIKNL peptide's effect on AS disease.

6 CONCLUSION

In this thesis, the MD simulations of the HLA-B*27:05 and HLA-B*27:09 with bound and unbound states were considered with 4 parallel runs of 100 ns simulations. The structural analysis including RMSD, RMSF and, the distances were performed to compare bound and unbound forms of HLA-B27 subtypes and to compare the results of two arthritogenic self-peptides with a viral peptide. Peptide DRASFIKNL, a self-peptide, showed similar results with peptide KK10, a viral peptide. Even though it does not bind as well as the KK10 peptide, it is stated in Atagündüz's study (2005) that this peptide generated a CTL response like peptide ARGQPGVMG. Thus, this thesis supports the arthritogenic peptide hypothesis by showing similar results between the DRASFIKNL peptide and the KK10 peptide. It also showed that the similarity of charged amino acid sequence with the viral peptide at the termini of the peptide had a greater effect than the similarity of polarity of positions. Also, RMSD measurements and the visualization results showed that the empty HLA-B*27:05 has more mobility than the empty HLA-B*27:05. With the binding of peptide, HLA-B*27:05 became more stable than the HLA-B27:09 complexes bound to the three peptides except the HLA-B*27:05/ARGQPGVMG. In conclusion, it is illustrated that the DRASFIKNL peptide has a possibility of being an immunogenic peptide.

7 REFERENCES

1. Marrack P, Kappler J, Kotzin BL. Autoimmune disease: why and where it occurs. Nat Med. 2001 Aug;7(8):899–905.
2. Zhu W, He X, Cheng K, Zhang L, Chen D, Wang X, et al. Ankylosing spondylitis: etiology, pathogenesis, and treatments. Bone Res. 2019 Aug 5;7(1):1–16.
3. Braun J, Sieper J. Ankylosing spondylitis. The Lancet. 2007 Apr 21;369(9570):1379–90.
4. McHugh K, Bowness P. The link between HLA-B27 and SpA--new ideas on an old problem. Rheumatology. 2012 Sep 1;51(9):1529–39.
5. Hermann E, Büschenfelde KHM zum, Fleischer B, Yu DTY. HLA-B27-restricted CD8 T cells derived from synovial fluids of patients with reactive arthritis and ankylosing spondylitis. The Lancet. 1993 Sep 11;342(8872):646–50.
6. Atagunduz P, Appel H, Kuon W, Wu P, Thiel A, Kloetzel PM, et al. HLA-B27-restricted CD8+ T cell response to cartilage-derived self peptides in ankylosing spondylitis. Arthritis Rheum. 2005 Mar;52(3):892–901.
7. FRAUENDORF E, VON GOESSEL H, MAY E, MÄRKER-HERMANN E. HLA-B27-restricted T cells from patients with ankylosing spondylitis recognize peptides from B*2705 that are similar to bacteria-derived peptides. Clin Exp Immunol. 2003 Nov;134(2):351–9.
8. Khan MA. An Update on the Genetic Polymorphism of HLA-B*27 With 213 Alleles Encompassing 160 Subtypes (and Still Counting). Curr Rheumatol Rep. 2017 Mar 1;19(2):9.
9. Schittenhelm RB, Sian TCCLK, Wilmann PG, Dudek NL, Purcell AW. Revisiting the Arthritogenic Peptide Theory: Quantitative Not Qualitative Changes in the Peptide Repertoire of HLA–B27 Allotypes. Arthritis Rheumatol. 2015;67(3):702–13.
10. Charles A Janeway J, Travers P, Walport M, Shlomchik MJ. The major histocompatibility complex and its functions. Immunobiol Immune Syst Health Dis 5th Ed [Internet]. 2001 [cited 2022 May 18]; Available from: https://www.ncbi.nlm.nih.gov/books/NBK27156/
11. Madden DR. The three-dimensional structure of peptide-MHC complexes. Annu Rev Immunol. 1995;13:587–622.
12. Nguyen AT, Szeto C, Gras S. The pockets guide to HLA class I molecules. Biochem Soc Trans. 2021 Nov 1;49(5):2319–31.
13. Miller JM, Sproule BJ. Pulmonary Function in Ankylosing Spondylitis. Am Rev Respir Dis. 1964 Sep 1;90(3):376–82.
14. Chen B, Li J, He C, Li D, Tong W, Zou Y, et al. Role of HLA-B27 in the pathogenesis of ankylosing spondylitis. Mol Med Rep. 2017 Apr;15(4):1943–51.
15. Brewerton DA, Hart FD, Nicholls A, Caffrey M, James DCO, Sturrock RD. ANKYLOSING SPONDYLITIS AND HL-A 27. The Lancet. 1973 Apr 28;301(7809):904–7.

16. Barnea E, Melamed Kadosh D, Haimovich Y, Satumtira N, Dorris ML, Nguyen MT, et al. The Human Leukocyte Antigen (HLA)-B27 Peptidome in Vivo, in Spondyloarthritis-susceptible HLA-B27 Transgenic Rats and the Effect of Erap1 Deletion. Mol Cell Proteomics MCP. 2017 Apr;16(4):642–62.
17. Jung JH, Bang CH, Seok H, Choi SJ, Song GG. Clinical Findings of Ankylosing Spondylitis With and Without Human Leukocyte Antigen (HLA)-B27 and HLA-B5. 2019;48(10):9.
18. Sorrentino R, Böckmann RA, Fiorillo MT. HLA-B27 and antigen presentation: At the crossroads between immune defense and autoimmunity. Mol Immunol. 2014 Jan 1;57(1):22–7.
19. Li XC, Raghavan M. Structure and function of major histocompatibility complex class I antigens. Curr Opin Organ Transplant. 2010 Aug;15(4):499–504.
20. Richman SA, Aggen DH, Dossett ML, Donermeyer DL, Allen PM, Greenberg PD, et al. Structural features of T cell receptor variable regions that enhance domain stability and enable expression as single-chain VαVβ fragments. Mol Immunol. 2009 Feb;46(5):902–16.
21. Ladell K, Hashimoto M, Iglesias MC, Wilmann PG, McLaren JE, Gras S, et al. A Molecular Basis for the Control of Preimmune Escape Variants by HIV-Specific CD8+ T Cells. Immunity. 2013 Mar 21;38(3):425–36.
22. Wong WK, Leem J, Deane CM. Comparative Analysis of the CDR Loops of Antigen Receptors. Front Immunol [Internet]. 2019 [cited 2022 May 24];10. Available from: https://www.frontiersin.org/article/10.3389/fimmu.2019.02454
23. Tsuchiya Y, Namiuchi Y, Wako H, Tsurui H. A study of CDR3 loop dynamics reveals distinct mechanisms of peptide recognition by T-cell receptors exhibiting different levels of cross-reactivity. Immunology. 2018;153(4):466–78.
24. Parker KC, Biddison WE, Coligan JE. Pocket Mutations of HLA-B27 Show That Anchor Residues Act Cumulatively to Stabilize Peptide Binding. Biochemistry. 1994 Jun 21;33(24):7736–43.
25. Abualrous ET, Fritzsche S, Hein Z, Al-Balushi MS, Reinink P, Boyle LH, et al. F pocket flexibility influences the tapasin dependence of two differentially disease-associated MHC Class I proteins. Eur J Immunol. 2015;45(4):1248–57.
26. Madden DR, Gorga JC, Strominger JL, Wiley DC. The three-dimensional structure of HLA-B27 at 2.1 Å resolution suggests a general mechanism for tight peptide binding to MHC. Cell. 1992 Sep 18;70(6):1035–48.
27. Hülsmeyer M, Fiorillo MT, Bettosini F, Sorrentino R, Saenger W, Ziegler A, et al. Dual, HLA-B27 Subtype-dependent Conformation of a Self-peptide. J Exp Med. 2004 Jan 19;199(2):271–81.
28. Bowness P, Zaccai N, Bird L, Jones EY. HLA-B27 and disease pathogenesis: new structural and functional insights. Expert Rev Mol Med. 1999 Oct;1(16):1–10.
29. Burrows SR, Rossjohn J, McCluskey J. Have we cut ourselves too short in mapping CTL epitopes? Trends Immunol. 2006 Jan;27(1):11–6.

30. Schittenhelm RB, Sivaneswaran S, Sian TCCLK, Croft NP, Purcell AW. Human Leukocyte Antigen (HLA) B27 Allotype-Specific Binding and Candidate Arthritogenic Peptides Revealed through Heuristic Clustering of Data-independent Acquisition Mass Spectrometry (DIA-MS) Data. Mol Cell Proteomics. 2016 Jun 1;15(6):1867–76.
31. Tedeschi V, Alba J, Paladini F, Paroli M, Cauli A, Mathieu A, et al. Unusual Placement of an EBV Epitope into the Groove of the Ankylosing Spondylitis-Associated HLA-B27 Allele Allows CD8+ T Cell Activation. Cells. 2019 Jun;8(6):572.
32. Praest P, Liaci AM, Förster F, Wiertz EJHJ. New insights into the structure of the MHC class I peptide-loading complex and mechanisms of TAP inhibition by viral immune evasion proteins. Mol Immunol. 2019 Sep 1;113:103–14.
33. Alvarez-Navarro C, López de Castro JA. ERAP1 structure, function and pathogenetic role in ankylosing spondylitis and other MHC-associated diseases. Mol Immunol. 2014 Jan 1;57(1):12–21.
34. Colbert RA, Tran TM, Layh-Schmitt G. HLA-B27 misfolding and ankylosing spondylitis. Mol Immunol. 2014 Jan 1;57(1):44–51.
35. Yazar M. The relation between ER stress and HLA-B27 misfolding. Med Res Innov [Internet]. 2017 [cited 2022 Jun 8];1(3). Available from: http://www.oatext.com/the-relation-between-er-stress-and-hla-b27-misfolding.php
36. Powis SJ, Colbert RA. HLA-B27: The Story Continues to Unfold. Arthritis Rheumatol Hoboken NJ. 2016 May;68(5):1057–9.
37. Ziegler A, Loll B, Misselwitz R, Uchanska-Ziegler B. Implications of structural and thermodynamic studies of HLA-B27 subtypes exhibiting differential association with ankylosing spondylitis. Adv Exp Med Biol. 2009;649:177–95.
38. Allen RL, O'Callaghan CA, McMichael AJ, Bowness P. Cutting Edge: HLA-B27 Can Form a Novel β2-Microglobulin-Free Heavy Chain Homodimer Structure. J Immunol. 1999 May 1;162(9):5045–8.
39. Cheng X, Mei Y, Ji X, Xue Q, Chen D. Molecular mechanism of the susceptibility difference between HLA-B*27:02/04/05 and HLA-B*27:06/09 to ankylosing spondylitis: substitution analysis, MD simulation, QSAR modelling, and in vitro assay. SAR QSAR Environ Res. 2016 May 3;27(5):409–25.
40. Fabian H, Huser H, Narzi D, Misselwitz R, Loll B, Ziegler A, et al. HLA-B27 Subtypes Differentially Associated with Disease Exhibit Conformational Differences in Solution. J Mol Biol. 2008 Feb 22;376(3):798–810.
41. Hein Z, Borchert B, Abualrous ET, Springer S. Distinct mechanisms survey the structural integrity of HLA-B*27:05 intracellularly and at the surface. PLOS ONE. 2018 Aug 2;13(8):e0200811.
42. Narzi D, Becker CM, Fiorillo MT, Uchanska-Ziegler B, Ziegler A, Böckmann RA. Dynamical Characterization of Two Differentially Disease Associated MHC Class I Proteins in Complex with Viral and Self-Peptides. J Mol Biol. 2012 Jan 13;415(2):429–42.

43. Pöhlmann T, Böckmann RA, Grubmüller H, Uchanska-Ziegler B, Ziegler A, Alexiev U. Differential Peptide Dynamics Is Linked to Major Histocompatibility Complex Polymorphism. J Biol Chem. 2004 Feb 7;279(27):28197–201.
44. Khan MA. Polymorphism of HLA-B27: 105 Subtypes Currently Known. Curr Rheumatol Rep. 2013 Aug 29;15(10):362.
45. Nasution AR, Mardjuadi A, Kunmartini S, Suryadhana NG, Setyohadi B, Sudarsono D, et al. HLA-B27 subtypes positively and negatively associated with spondyloarthropathy. J Rheumatol. 1997 Jun;24(6):1111–4.
46. Van GAALEN FA. Does HLA-B*2706 protect against ankylosing spondylitis? A meta-analysis. Int J Rheum Dis. 2012;15(1):8–12.
47. Paladini F, Taccari E, Fiorillo MT, Cauli A, Passiu G, Mathieu A, et al. Distribution of HLA–B27 subtypes in Sardinia and continental Italy and their association with spondylarthropathies. Arthritis Rheum. 2005;52(10):3319–21.
48. Liu X, Hu LH, Li YR, Chen FH, Ning Y, Yao QF. The association of HLA-B*27 subtypes with ankylosing spondylitis in Wuhan population of China. Rheumatol Int. 2010 Mar 1;30(5):587–90.
49. Liu Y, Jiang L, Cai Q, Danoy P, Barnardo MCNM, Brown MA, et al. Predominant association of HLA-B*2704 with ankylosing spondylitis in Chinese Han patients. Tissue Antigens. 2010 Jan;75(1):61–4.
50. Birinci A, Bilgici A, Kuru O, Durupinar B. HLA-B27 polymorphism in Turkish patients with ankylosing spondylitis. Rheumatol Int. 2006 Feb 1;26(4):285–7.
51. Acar M, Cora T, Tunc R, Acar H. HLA-B27 subtypes in Turkish patients with ankylosing spondylitis and healthy controls. Rheumatol Int. 2012 Oct 1;32(10):3103–5.
52. IPD-IMGT/HLA Database [Internet]. [cited 2022 May 12]. Available from: https://www.ebi.ac.uk/ipd/imgt/hla/alleles/
53. Dominguez C, Boelens R, Bonvin AMJJ. HADDOCK: A Protein−Protein Docking Approach Based on Biochemical or Biophysical Information. J Am Chem Soc. 2003 Feb 19;125(7):1731–7.
54. Morris GM, Lim-Wilby M. Molecular Docking. In: Kukol A, editor. Molecular Modeling of Proteins [Internet]. Totowa, NJ: Humana Press; 2008 [cited 2022 May 16]. p. 365–82. (Methods Molecular Biology™). Available from: https://doi.org/10.1007/978-1-59745-177-2_19
55. Vangone A, Rodrigues JPGLM, Xue LC, van Zundert GCP, Geng C, Kurkcuoglu Z, et al. Sense and simplicity in HADDOCK scoring: Lessons from CASP-CAPRI round 1. Proteins. 2017 Mar;85(3):417–23.
56. Leach AR. Molecular Modelling: Principles and Applications [Internet]. Prentice Hall; 2001. Available from: https://books.google.com.tr/books?id=kB7jsbV-uhkC
57. Hospital A, Goñi JR, Orozco M, Gelpi JL. Molecular dynamics simulations: advances and applications. Adv Appl Bioinforma Chem. 2015 Nov 19;8:37–47.

58. Konteatis ZD, Klon AE, Zou J, Meshkat S. Chapter Fourteen - Computational Approach to De Novo Discovery of Fragment Binding for Novel Protein States. In: Kuo LC, editor. Methods in Enzymology [Internet]. Academic Press; 2011 [cited 2021 Dec 21]. p. 357–80. (Fragment-Based Drug Design; vol. 493). Available from: https://www.sciencedirect.com/science/article/pii/B9780123812742000145
59. Swope WC, Andersen HC, Berens PH, Wilson KR. A computer simulation method for the calculation of equilibrium constants for the formation of physical clusters of molecules: Application to small water clusters. J Chem Phys. 1982 Jan;76(1):637–49.
60. Waidyasooriya HM, Hariyama M, Kasahara K. An FPGA Accelerator for Molecular Dynamics Simulation Using OpenCL. Int J Networked Distrib Comput. 2017 Jan 2;5(1):52–61.
61. Cygan RT. Molecular Modeling in Mineralogy and Geochemistry. Rev Mineral Geochem. 2001 Jan 1;42(1):1–35.
62. Binder K, Horbach J, Kob W, Paul W, Varnik F. Molecular dynamics simulations. J Phys Condens Matter. 2004 Jan;16(5):S429–53.
63. de Vries SJ, van Dijk M, Bonvin AMJJ. The HADDOCK web server for data-driven biomolecular docking. Nat Protoc. 2010 May;5(5):883–97.
64. Phillips JC, Braun R, Wang W, Gumbart J, Tajkhorshid E, Villa E, et al. Scalable molecular dynamics with NAMD. J Comput Chem. 2005;26(16):1781–802.
65. Humphrey W, Dalke A, Schulten K. VMD: Visual molecular dynamics. J Mol Graph. 1996 Feb 1;14(1):33–8.
66. PyMOL | pymol.org [Internet]. [cited 2020 Jun 26]. Available from: https://pymol.org/2/
67. R: The R Project for Statistical Computing [Internet]. [cited 2022 May 17]. Available from: https://www.r-project.org/
68. Laskowski RA, Jabłońska J, Pravda L, Vařeková RS, Thornton JM. PDBsum: Structural summaries of PDB entries. Protein Sci. 2018;27(1):129–34.
69. Vries SJ de, Bonvin AMJJ. CPORT: A Consensus Interface Predictor and Its Performance in Prediction-Driven Docking with HADDOCK. PLOS ONE. 2011 Mar 25;6(3):e17695.
70. Cortes A, Pulit SL, Leo PJ, Pointon JJ, Robinson PC, Weisman MH, et al. Major histocompatibility complex associations of ankylosing spondylitis are complex and involve further epistasis with ERAP1. Nat Commun. 2015 May 21;6(1):7146.
71. Lopez de Castro J a., Alvarez I, Marcilla M, Paradela A, Ramos M, Sesma L, et al. HLA-B27: a registry of constitutive peptide ligands. Tissue Antigens. 2004;63(5):424–45.
72. Uchanska-Ziegler B, Ziegler A, Schmieder P. Structural and dynamic features of HLA-B27 subtypes. Curr Opin Rheumatol. 2013 Jul;25(4):411–8.
73. Trellet M, Melquiond ASJ, Bonvin AMJJ. A Unified Conformational Selection and Induced Fit Approach to Protein-Peptide Docking. PLOS ONE. 2013 Mar 13;8(3):e58769.
74. Thévenet P, Shen Y, Maupetit J, Guyon F, Derreumaux P, Tufféry P. PEP-FOLD: an updated de novo structure prediction server for both linear and disulfide bonded cyclic peptides. Nucleic Acids Res. 2012 Jul 1;40(W1):W288–93.

75. Lee J, Cheng X, Swails JM, Yeom MS, Eastman PK, Lemkul JA, et al. CHARMM-GUI Input Generator for NAMD, GROMACS, AMBER, OpenMM, and CHARMM/OpenMM Simulations Using the CHARMM36 Additive Force Field. J Chem Theory Comput. 2016 Jan 12;12(1):405–13.
76. Jo S, Kim T, Iyer VG, Im W. CHARMM-GUI: A web-based graphical user interface for CHARMM. J Comput Chem. 2008;29(11):1859–65.
77. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, de Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. Nat Methods. 2017 Jan;14(1):71–3.
78. Sukhwal A, Sowdhamini R. Oligomerisation status and evolutionary conservation of interfaces of protein structural domain superfamilies. Mol Biosyst. 2013 Jun 4;9(7):1652–61.
79. Wallace AC, Laskowski RA, Thornton JM. LIGPLOT: a program to generate schematic diagrams of protein-ligand interactions. Protein Eng. 1995 Feb;8(2):127–34.
80. Reynisson B, Alvarez B, Paul S, Peters B, Nielsen M. NetMHCpan-4.1 and NetMHCIIpan-4.0: improved predictions of MHC antigen presentation by concurrent motif deconvolution and integration of MS MHC eluted ligand data. Nucleic Acids Res. 2020 Jul 2;48(W1):W449–54.
81. Calis JJA, Maybeno M, Greenbaum JA, Weiskopf D, Silva ADD, Sette A, et al. Properties of MHC Class I Presented Peptides That Enhance Immunogenicity. PLOS Comput Biol. 2013 Oct 24;9(10):e1003266.
82. Zheng L, Huang S, Mu N, Zhang H, Zhang J, Chang Y, et al. RAACBook: a web server of reduced amino acid alphabet for sequence-dependent inference by using Chou's five-step rule. Database. 2019 Jan 1;2019:baz131.
83. Create Elegant Data Visualisations Using the Grammar of Graphics [Internet]. [cited 2022 May 8]. Available from: https://ggplot2.tidyverse.org/index.html
84. Beerbaum M, Ballaschk M, Erdmann N, Schnick C, Diehl A, Uchanska-Ziegler B, et al. NMR spectroscopy reveals unexpected structural variation at the protein-protein interface in MHC class I molecules. J Biomol NMR. 2013 Sep 5;57.
85. RAET1E retinoic acid early transcript 1E [Homo sapiens (human)] - Gene - NCBI [Internet]. [cited 2022 May 30]. Available from: https://www.ncbi.nlm.nih.gov/gene?Db=gene&Cmd=DetailsSearch&Term=135250
86. Yair-Sabag S, Tedeschi V, Vitulano C, Barnea E, Glaser F, Melamed Kadosh D, et al. The Peptide Repertoire of HLA-B27 may include Ligands with Lysine at P2 Anchor Position. PROTEOMICS. 2018;18(9):1700249.
87. Narzi D, Winkler K, Saidowsky J, Misselwitz R, Ziegler A, Böckmann RA, et al. Molecular Determinants of Major Histocompatibility Complex Class I Complex Stability: SHAPING ANTIGENIC FEATURES THROUGH SHORT AND LONG RANGE ELECTROSTATIC INTERACTIONS *. J Biol Chem. 2008 Aug 22;283(34):23093–103.
88. Madden DR, Gorga JC, Strominger JL, Wiley DC. The structure of HLA-B27 reveals nonamer self-peptides bound in an extended conformation. Nature. 1991 Sep;353(6342):321–5.

8 APPENDIX

Appendix 1. HLA-B*27:05 complexes with viral peptides

HLA-B Variants	PDB code	Resolution of 3D Structure (Å)	Sequence of bound peptides	DOI	Ligand
HLA-B2705 (with B2M)	4G9D	1.6	KRWIILGLNK (HIV peptide)	10.1016/j.immuni.2012.11.021	Glycerol (GOL)
HLA-B2705 (with B2M)	4G8G	2.4	KRWIILGLNK (HIV peptide)	10.1016/j.immuni.2012.11.021	-
HLA-B2705 (with B2M)	4G8I	1.6	KRWIIMGLNK (HIV peptide)	10.1016/j.immuni.2012.11.021	Tris Buffer (TRS)
HLA-B2705 (with B2M)	4G9F	1.9	KRWIIMGLNK (HIV peptide)	10.1016/j.immuni.2012.11.021	Sulfate ion (SO4)
HLA-B2705 (with B2M)	2BSS	2	KRWIILGLNK (HIV peptide)	10.1002/eji.200425724	-
HLA-B2705 (with B2M)	2BSR	2.3	RRIYDLIEL (Epstein-Barr Nuclear Antigen-6)	10.1002/eji.200425724	-
HLA-B2705 (with B2M)	2BST	2.1	SRYWAIRTR (INFLUENZA NUCLEOPROTEIN)	10.1002/eji.200425724	-
HLA-B2705 (with B2M)	1UXS	1.55	RRRWRRLTV (Epstein-Barr membrane protein)	10.1074/jbc.M410807200	Glycerol (GOL)

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides.

Potential Salt Bridges									
			0 ns						
HLA Subtype	Peptide Sequence	Simulation Number	Residue 1 (Binding Groove)			Residue 2 (Peptide)			Distance
			Res Num	Res Name	Atom Name	Res Num	Res Name	Atom Name	
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH1	2.78
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.63
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH1	3.17
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH2	2.7
B*27:05	ARGQPGVMG	Simulation 1	63	GLU	OE1	2	ARG	NH2	2.91
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH1	3.2
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.58
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH2	3.26
B*27:05	ARGQPGVMG	Simulation 2	63	GLU	OE1	2	ARG	NH2	2.75
B*27:05	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH1	2.85
B*27:05	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.89
B*27:05	ARGQPGVMG	Simulation 3	45	GLU	OE2	2	ARG	NH2	2.79
B*27:05	ARGQPGVMG	Simulation 3	63	GLU	OE1	2	ARG	NH2	3.27
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH1	2.59
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH2	3.06
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH1	3.13
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH2	2.77
B*27:05	ARGQPGVMG	Simulation 4	63	GLU	OE1	2	ARG	NH2	2.84
B*27:05	DRASFIKNL	Simulation 1	63	GLU	OE1	2	ARG	NH2	2.68
B*27:05	DRASFIKNL	Simulation 1	63	GLU	OE2	2	ARG	NH2	3.38
B*27:05	DRASFIKNL	Simulation 1	74	ASP	OD1	7	LYS	NZ	2.68
B*27:05	DRASFIKNL	Simulation 1	77	ASP	OD2	7	LYS	NZ	2.65
B*27:05	DRASFIKNL	Simulation 1	116	ASP	OD2	7	LYS	NZ	2.72
B*27:05	DRASFIKNL	Simulation 1	163	GLU	OE2	2	ARG	NH1	2.89
B*27:05	DRASFIKNL	Simulation 2	63	GLU	OE1	2	ARG	NH2	3.15
B*27:05	DRASFIKNL	Simulation 2	63	GLU	OE2	2	ARG	NH2	2.86
B*27:05	DRASFIKNL	Simulation 2	77	ASP	OD2	7	LYS	NZ	2.71
B*27:05	DRASFIKNL	Simulation 2	116	ASP	OD1	7	LYS	NZ	3.36
B*27:05	DRASFIKNL	Simulation 2	116	ASP	OD2	7	LYS	NZ	2.86
B*27:05	DRASFIKNL	Simulation 3	63	GLU	OE1	2	ARG	NH2	2.75
B*27:05	DRASFIKNL	Simulation 3	63	GLU	OE2	2	ARG	NH2	3.67
B*27:05	DRASFIKNL	Simulation 3	77	ASP	OD2	7	LYS	NZ	2.59
B*27:05	DRASFIKNL	Simulation 3	116	ASP	OD1	7	LYS	NZ	2.77
B*27:05	DRASFIKNL	Simulation 3	116	ASP	OD2	7	LYS	NZ	3.57
B*27:05	DRASFIKNL	Simulation 3	163	GLU	OE1	2	ARG	NH1	3.98
B*27:05	DRASFIKNL	Simulation 3	163	GLU	OE2	2	ARG	NH1	2.79

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

B*27:05	DRASFIKNL	Simulation 4	63	GLU	OE1	2	ARG	NH2	2.93
B*27:05	DRASFIKNL	Simulation 4	63	GLU	OE2	2	ARG	NH2	3.42
B*27:05	DRASFIKNL	Simulation 4	77	ASP	OD2	7	LYS	NZ	2.6
B*27:05	DRASFIKNL	Simulation 4	116	ASP	OD1	7	LYS	NZ	2.93
B*27:05	DRASFIKNL	Simulation 4	116	ASP	OD2	7	LYS	NZ	2.6
B*27:05	DRASFIKNL	Simulation 4	163	GLU	OE2	2	ARG	NH1	2.79
B*27:05	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH2	2.8
B*27:05	KRWIILGLNK (HIV)	Simulation 1	77	ASP	OD2	10	LYS	NZ	3.93
B*27:05	KRWIILGLNK (HIV)	Simulation 1	116	ASP	OD1	10	LYS	NZ	2.74
B*27:05	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE1	1	LYS	NZ	3.3
B*27:05	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE2	1	LYS	NZ	2.79
B*27:05	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.83
B*27:05	KRWIILGLNK (HIV)	Simulation 2	77	ASP	OD2	10	LYS	NZ	2.53
B*27:05	KRWIILGLNK (HIV)	Simulation 2	116	ASP	OD1	10	LYS	NZ	2.68
B*27:05	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE1	1	LYS	NZ	2.55
B*27:05	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE2	1	LYS	NZ	2.83
B*27:05	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.76
B*27:05	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH2	3.37
B*27:05	KRWIILGLNK (HIV)	Simulation 3	77	ASP	OD2	10	LYS	NZ	2.82
B*27:05	KRWIILGLNK (HIV)	Simulation 3	116	ASP	OD2	10	LYS	NZ	2.56
B*27:05	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE1	1	LYS	NZ	2.73
B*27:05	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE2	1	LYS	NZ	2.87
B*27:05	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE1	2	ARG	NH2	3
B*27:05	KRWIILGLNK (HIV)	Simulation 4	77	ASP	OD2	10	LYS	NZ	3.86
B*27:05	KRWIILGLNK (HIV)	Simulation 4	116	ASP	OD1	10	LYS	NZ	2.6
B*27:05	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE1	1	LYS	NZ	3.41
B*27:05	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE2	1	LYS	NZ	2.65
B*27:09	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH2	2.72
B*27:09	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH2	3.91
B*27:09	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.66
B*27:09	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH2	3.75
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.65
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE2	2	ARG	NH2	3.44
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.75
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.58
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE1	2	ARG	NH2	3.25
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE2	2	ARG	NH1	3.05
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE2	2	ARG	NH2	2.92
B*27:09	DRASFIKNL	Simulation 2	76	GLU	OE1	7	LYS	NZ	2.75
B*27:09	DRASFIKNL	Simulation 2	77	ASP	OD1	7	LYS	NZ	2.99
B*27:09	DRASFIKNL	Simulation 2	163	GLU	OE1	2	ARG	NH1	2.62

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

B*27:09	DRASFIKNL	Simulation 2	163	GLU	OE1	2	ARG	NH2	3.32
B*27:09	DRASFIKNL	Simulation 2	163	GLU	OE2	2	ARG	NH1	3.71
B*27:09	DRASFIKNL	Simulation 2	163	GLU	OE2	2	ARG	NH2	2.91
B*27:09	DRASFIKNL	Simulation 3	76	GLU	OE1	7	LYS	NZ	2.71
B*27:09	DRASFIKNL	Simulation 3	76	GLU	OE2	7	LYS	NZ	3.61
B*27:09	DRASFIKNL	Simulation 4	76	GLU	OE2	7	LYS	NZ	2.61
B*27:09	DRASFIKNL	Simulation 4	77	ASP	OD1	7	LYS	NZ	2.66
B*27:09	DRASFIKNL	Simulation 4	163	GLU	OE2	2	ARG	NH1	3.12
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH1	2.7
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.67
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE2	2	ARG	NH1	3.22
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE2	2	ARG	NH2	2.85
B*27:09	KRWIILGLNK (HIV)	Simulation 1	63	GLU	OE1	2	ARG	NH1	2.66
B*27:09	KRWIILGLNK (HIV)	Simulation 1	77	ASP	OD1	10	LYS	NZ	2.89
B*27:09	KRWIILGLNK (HIV)	Simulation 1	77	ASP	OD2	10	LYS	NZ	3.17
B*27:09	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE1	1	LYS	NZ	2.66
B*27:09	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE2	1	LYS	NZ	2.69
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH1	3.05
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.61
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE2	2	ARG	NH1	2.74
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE2	2	ARG	NH2	3.66
B*27:09	KRWIILGLNK (HIV)	Simulation 2	63	GLU	OE1	2	ARG	NH1	2.81
B*27:09	KRWIILGLNK (HIV)	Simulation 2	77	ASP	OD1	10	LYS	NZ	2.73
B*27:09	KRWIILGLNK (HIV)	Simulation 2	77	ASP	OD2	10	LYS	NZ	2.85
B*27:09	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE2	1	LYS	NZ	2.6
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH1	3.2
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.72
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH1	2.63
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH2	3.59
B*27:09	KRWIILGLNK (HIV)	Simulation 3	63	GLU	OE1	2	ARG	NH1	2.75
B*27:09	KRWIILGLNK (HIV)	Simulation 3	77	ASP	OD1	10	LYS	NZ	2.66
B*27:09	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE1	1	LYS	NZ	3.88
B*27:09	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE2	1	LYS	NZ	2.57
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE1	2	ARG	NH1	3.5
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.72
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH1	2.72
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.19
B*27:09	KRWIILGLNK (HIV)	Simulation 4	63	GLU	OE2	2	ARG	NH1	2.69
B*27:09	KRWIILGLNK (HIV)	Simulation 4	77	ASP	OD1	10	LYS	NZ	2.69
B*27:09	KRWIILGLNK (HIV)	Simulation 4	77	ASP	OD2	10	LYS	NZ	3.94
B*27:09	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE1	1	LYS	NZ	2.65

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

Potential Salt Bridges									
			0 ns						
HLA Subtype	Peptide Sequence	Simulation Number	Residue 1 (Binding Groove)			Residue 2 (Peptide)			Distance
			Res Num	Res Name	Atom Name	Res Num	Res Name	Atom Name	
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH1	2.67
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.76
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH1	3.15
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH2	2.64
B*27:05	ARGQPGVMG	Simulation 1	63	GLU	OE1	2	ARG	NH2	2.79
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH1	2.84
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH2	3.34
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH1	3.76
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH2	2.89
B*27:05	ARGQPGVMG	Simulation 2	63	GLU	OE2	2	ARG	NH2	2.73
B*27:05	ARGQPGVMG	Simulation 3	None Detected						
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH1	2.79
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH2	3.66
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH1	3.29
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH2	2.65
B*27:05	ARGQPGVMG	Simulation 4	63	GLU	OE2	2	ARG	NH2	2.78
B*27:05	DRASFIKNL	Simulation 1	63	GLU	OE1	2	ARG	NH2	3.69
B*27:05	DRASFIKNL	Simulation 1	63	GLU	OE2	2	ARG	NH2	2.78
B*27:05	DRASFIKNL	Simulation 1	74	ASP	OD1	7	LYS	NZ	2.63
B*27:05	DRASFIKNL	Simulation 1	77	ASP	OD2	7	LYS	NZ	2.66
B*27:05	DRASFIKNL	Simulation 1	116	ASP	OD1	7	LYS	NZ	2.73
B*27:05	DRASFIKNL	Simulation 2	77	ASP	OD1	7	LYS	NZ	3.87
B*27:05	DRASFIKNL	Simulation 2	77	ASP	OD2	7	LYS	NZ	2.71
B*27:05	DRASFIKNL	Simulation 2	116	ASP	OD2	7	LYS	NZ	3.92
B*27:05	DRASFIKNL	Simulation 3	63	GLU	OE1	2	ARG	NH2	3
B*27:05	DRASFIKNL	Simulation 3	63	GLU	OE2	2	ARG	NH2	2.91
B*27:05	DRASFIKNL	Simulation 3	74	ASP	OD1	7	LYS	NZ	2.58
B*27:05	DRASFIKNL	Simulation 3	77	ASP	OD2	7	LYS	NZ	2.56
B*27:05	DRASFIKNL	Simulation 4	63	GLU	OE1	2	ARG	NH2	2.89
B*27:05	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE2	2	ARG	NH2	3.88
B*27:05	KRWIILGLNK (HIV)	Simulation 1	116	ASP	OD1	10	LYS	NZ	2.81
B*27:05	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE1	1	LYS	NZ	2.67
B*27:05	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.71
B*27:05	KRWIILGLNK (HIV)	Simulation 2	77	ASP	OD1	10	LYS	NZ	3.99
B*27:05	KRWIILGLNK (HIV)	Simulation 2	77	ASP	OD2	10	LYS	NZ	2.66

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

B*27:05	KRWIILGLNK (HIV)	Simulation 2	116	ASP	OD2	10	LYS	NZ	2.83
B*27:05	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE1	1	LYS	NZ	2.6
B*27:05	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH2	2.74
B*27:05	KRWIILGLNK (HIV)	Simulation 3	116	ASP	OD1	10	LYS	NZ	2.91
B*27:05	KRWIILGLNK (HIV)	Simulation 3	116	ASP	OD2	10	LYS	NZ	3.8
B*27:05	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE1	1	LYS	NZ	3.07
B*27:05	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE2	1	LYS	NZ	2.87
B*27:05	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.66
B*27:05	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.66
B*27:05	KRWIILGLNK (HIV)	Simulation 4	77	ASP	OD2	10	LYS	NZ	4
B*27:05	KRWIILGLNK (HIV)	Simulation 4	116	ASP	OD2	10	LYS	NZ	3.49
B*27:05	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE1	1	LYS	NZ	2.82
B*27:05	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE2	1	LYS	NZ	2.95
B*27:09	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH2	2.98
B*27:09	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH2	2.83
B*27:09	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH2	3.78
B*27:09	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH2	2.72
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH1	3.69
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.66
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE2	2	ARG	NH1	2.73
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE2	2	ARG	NH2	3.18
B*27:09	ARGQPGVMG	Simulation 3	63	GLU	OE2	2	ARG	NH2	2.78
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH1	3.67
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.73
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH1	2.67
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.27
B*27:09	ARGQPGVMG	Simulation 4	63	GLU	OE1	2	ARG	NH2	2.91
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE1	2	ARG	NH1	3.83
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE1	2	ARG	NH2	2.67
B*27:09	DRASFIKNL	Simulation 2	None Detected						
B*27:09	DRASFIKNL	Simulation 3	163	GLU	OE2	2	ARG	NH1	2.72
B*27:09	DRASFIKNL	Simulation 3	163	GLU	OE2	2	ARG	NH2	2.9
B*27:09	DRASFIKNL	Simulation 4	77	ASP	OD1	7	LYS	NZ	3.49
B*27:09	DRASFIKNL	Simulation 4	77	ASP	OD2	7	LYS	NZ	2.65
B*27:09	DRASFIKNL	Simulation 4	163	GLU	OE1	2	ARG	NH1	3.16
B*27:09	DRASFIKNL	Simulation 4	163	GLU	OE1	2	ARG	NH2	3.91
B*27:09	DRASFIKNL	Simulation 4	163	GLU	OE2	2	ARG	NH1	2.8
B*27:09	DRASFIKNL	Simulation 4	163	GLU	OE2	2	ARG	NH2	2.71
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH1	2.97
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.94
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE2	2	ARG	NH1	3.14

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE2	2	ARG	NH2	2.59
B*27:09	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE1	1	LYS	NZ	2.71
B*27:09	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE2	1	LYS	NZ	2.84
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH1	3.35
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE2	2	ARG	NH1	2.89
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE2	2	ARG	NH2	2.65
B*27:09	KRWIILGLNK (HIV)	Simulation 2	76	GLU	OE1	10	LYS	NZ	2.65
B*27:09	KRWIILGLNK (HIV)	Simulation 2	76	GLU	OE2	10	LYS	NZ	3.75
B*27:09	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE1	1	LYS	NZ	2.64
B*27:09	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE2	1	LYS	NZ	3.64
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH1	2.67
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.88
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH1	3.03
B*27:09	KRWIILGLNK (HIV)	Simulation 3	77	ASP	OD1	10	LYS	NZ	2.63
B*27:09	KRWIILGLNK (HIV)	Simulation 3	77	ASP	OD2	10	LYS	NZ	3.53
B*27:09	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE1	1	LYS	NZ	2.58
B*27:09	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE2	1	LYS	NZ	3.06
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH1	2.99
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.42

Potential Salt Bridges

HLA Subtype	Peptide Sequence	Simulation Number	100 ns						
			Residue 1 (Binding Groove)			Residue 2 (Peptide)			Distance
			Res Num	Res Name	Atom Name	Res Num	Res Name	Atom Name	
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH1	3.87
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH2	2.72
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH1	2.7
B*27:05	ARGQPGVMG	Simulation 1	45	GLU	OE2	2	ARG	NH2	3.06
B*27:05	ARGQPGVMG	Simulation 1	63	GLU	OE2	2	ARG	NH2	2.8
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH1	2.88
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH2	3.71
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH1	3.43
B*27:05	ARGQPGVMG	Simulation 2	45	GLU	OE2	2	ARG	NH2	2.79
B*27:05	ARGQPGVMG	Simulation 3	None Detected						
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH1	3.6
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.87
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH1	2.69
B*27:05	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.51
B*27:05	DRASFIKNL	Simulation 1	63	GLU	OE1	2	ARG	NH2	2.74
B*27:05	DRASFIKNL	Simulation 1	63	GLU	OE2	2	ARG	NH2	3.86
B*27:05	DRASFIKNL	Simulation 1	74	ASP	OD1	7	LYS	NZ	3.91

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

B*27:05	DRASFIKNL	Simulation 1	74	ASP	OD2	7	LYS	NZ	2.51
B*27:05	DRASFIKNL	Simulation 1	77	ASP	OD2	7	LYS	NZ	2.55
B*27:05	DRASFIKNL	Simulation 1	116	ASP	OD2	7	LYS	NZ	2.85
B*27:05	DRASFIKNL	Simulation 2	45	GLU	OE1	2	ARG	NH1	3.82
B*27:05	DRASFIKNL	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.75
B*27:05	DRASFIKNL	Simulation 2	45	GLU	OE2	2	ARG	NH1	3.18
B*27:05	DRASFIKNL	Simulation 2	45	GLU	OE2	2	ARG	NH2	3.56
B*27:05	DRASFIKNL	Simulation 2	63	GLU	OE1	2	ARG	NH2	2.82
B*27:05	DRASFIKNL	Simulation 2	74	ASP	OD2	7	LYS	NZ	2.53
B*27:05	DRASFIKNL	Simulation 2	77	ASP	OD1	7	LYS	NZ	2.63
B*27:05	DRASFIKNL	Simulation 3	62	ARG	NH2	1	ASP	OD1	3.16
B*27:05	DRASFIKNL	Simulation 3	63	GLU	OE1	2	ARG	OD1	3.27
B*27:05	DRASFIKNL	Simulation 3	63	GLU	OE2	2	ARG	NH2	2.85
B*27:05	DRASFIKNL	Simulation 3	74	ASP	OD1	7	LYS	NZ	2.62
B*27:05	DRASFIKNL	Simulation 3	77	ASP	OD2	7	LYS	NZ	2.69
B*27:05	DRASFIKNL	Simulation 4	74	ASP	OD2	7	LYS	NZ	2.68
B*27:05	DRASFIKNL	Simulation 4	77	ASP	OD1	7	LYS	NZ	2.68
B*27:05	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.29
B*27:05	KRWIILGLNK (HIV)	Simulation 1	77	ASP	OD2	10	LYS	NZ	3.92
B*27:05	KRWIILGLNK (HIV)	Simulation 1	116	ASP	OD2	10	LYS	NZ	2.72
B*27:05	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE1	1	LYS	NZ	3.76
B*27:05	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE2	1	LYS	NZ	2.68
B*27:05	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH2	3.12
B*27:05	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE2	2	ARG	NH2	2.97
B*27:05	KRWIILGLNK (HIV)	Simulation 2	116	ASP	OD2	10	LYS	NZ	3.76
B*27:05	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE1	1	LYS	NZ	3.6
B*27:05	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE2	1	LYS	NZ	2.78
B*27:05	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH2	2.73
B*27:05	KRWIILGLNK (HIV)	Simulation 3	63	GLU	OE1	1	LYS	NZ	2.98
B*27:05	KRWIILGLNK (HIV)	Simulation 3	63	GLU	OE2	1	LYS	NZ	2.83
B*27:05	KRWIILGLNK (HIV)	Simulation 3	77	ASP	OD2	10	LYS	NZ	3.06
B*27:05	KRWIILGLNK (HIV)	Simulation 3	116	ASP	OD1	10	LYS	NZ	3.83
B*27:05	KRWIILGLNK (HIV)	Simulation 3	116	ASP	OD2	10	LYS	NZ	3.65
B*27:05	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.92
B*27:05	KRWIILGLNK (HIV)	Simulation 4	116	ASP	OD1	10	LYS	NZ	2.53
B*27:05	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE1	1	LYS	NZ	2.62
B*27:05	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE2	1	LYS	NZ	3.71
B*27:09	ARGQPGVMG	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.37
B*27:09	ARGQPGVMG	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.83
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH1	3.5
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE1	2	ARG	NH2	2.76

Appendix 2. Salt bridges results of MD simulations for HLA-B27 subtypes with peptides (cont.)

B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE2	2	ARG	NH1	2.69
B*27:09	ARGQPGVMG	Simulation 3	45	GLU	OE2	2	ARG	NH2	3.48
B*27:09	ARGQPGVMG	Simulation 3	63	GLU	OE2	2	ARG	NH2	3.77
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH1	3.74
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE1	2	ARG	NH2	2.94
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH1	2.73
B*27:09	ARGQPGVMG	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.28
B*27:09	ARGQPGVMG	Simulation 4	63	GLU	OE1	2	ARG	NH2	2.98
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE1	2	ARG	NH2	3.52
B*27:09	DRASFIKNL	Simulation 1	163	GLU	OE2	2	ARG	NH3	3.96
B*27:09	DRASFIKNL	Simulation 2	63	GLU	OE2	2	ARG	NH2	2.87
B*27:09	DRASFIKNL	Simulation 3	163	GLU	OE1	2	ARG	NH1	3.33
B*27:09	DRASFIKNL	Simulation 3	163	GLU	OE1	2	ARG	NH2	2.63
B*27:09	DRASFIKNL	Simulation 3	163	GLU	OE2	2	ARG	NH1	3.14
B*27:09	DRASFIKNL	Simulation 3	163	GLU	OE2	2	ARG	NH2	3.85
B*27:09	DRASFIKNL	Simulation 4	77	ASP	OD1	7	LYS	NZ	2.71
B*27:09	DRASFIKNL	Simulation 4	77	ASP	OD2	7	LYS	NZ	2.69
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH1	2.83
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE1	2	ARG	NH2	3.54
B*27:09	KRWIILGLNK (HIV)	Simulation 1	45	GLU	OE2	2	ARG	NH1	3.14
B*27:09	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE1	1	LYS	NZ	3.61
B*27:09	KRWIILGLNK (HIV)	Simulation 1	163	GLU	OE2	1	LYS	NZ	2.74
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH1	3.23
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE1	2	ARG	NH2	2.73
B*27:09	KRWIILGLNK (HIV)	Simulation 2	45	GLU	OE2	2	ARG	NH1	3.33
B*27:09	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE1	1	LYS	NZ	2.57
B*27:09	KRWIILGLNK (HIV)	Simulation 2	163	GLU	OE2	1	LYS	NZ	3.27
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH1	2.83
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE1	2	ARG	NH2	3.7
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH1	3.29
B*27:09	KRWIILGLNK (HIV)	Simulation 3	45	GLU	OE2	2	ARG	NH2	2.6
B*27:09	KRWIILGLNK (HIV)	Simulation 3	76	GLU	OE1	10	LYS	NZ	3.08
B*27:09	KRWIILGLNK (HIV)	Simulation 3	76	GLU	OE2	10	LYS	NZ	2.86
B*27:09	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE1	1	LYS	NZ	3.37
B*27:09	KRWIILGLNK (HIV)	Simulation 3	163	GLU	OE2	1	LYS	NZ	2.67
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH1	2.67
B*27:09	KRWIILGLNK (HIV)	Simulation 4	45	GLU	OE2	2	ARG	NH2	3.25
B*27:09	KRWIILGLNK (HIV)	Simulation 4	77	ASP	OD1	10	LYS	NZ	2.59
B*27:09	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE1	1	LYS	NZ	2.82
B*27:09	KRWIILGLNK (HIV)	Simulation 4	163	GLU	OE2	1	LYS	NZ	3.11

9 CURRICULUM VITAE