A COMPARATIVE MOLECULAR DYNAMICS STUDY OF HLA-B51 AND HLA-B52: IMPLICATIONS FOR THE PATHOGENIC ROLE OF HLA-B51 IN BEHÇET'S DISEASE

by

Deniz Aydın

A Thesis Submitted to the
Graduate School of Engineering
in Partial Fulfillment of the Requirements for
the Degree of

Master of Science

in

Chemical and Biological Engineering

Koç University

September, 2014

Koç University

Graduate School of Sciences and Engineering

This is to certify that I have examined this copy of a master's thesis by

Deniz Aydın

and have found that it is complete and satisfactory in all respects,
and that any and all revisions required by the final
examining committee have been made.

Committee Members:

_____

Prof. Dr. Burak Erman(Advisor)

_____

Asst. Prof. Dr. Mehmet Sayar(Advisor)

_____

Prof. Dr. Alper Demir

_____

Prof. Dr. Ahmet Gül

_____

Asst. Prof. Dr. Alkan Kabakçıoğlu

Date: _____

To my family,

Nesrin, Refhan and Yıldız Aydın

ABSTRACT

Behçet's Disease (BD) is a chronic inflammatory disorder and its aetiology is unknown. Nearly every genetic study related to BD has confirmed the association of HLA-B51 with BD since the time the association was first recognized forty years ago. Despite the recurrent affirmation of HLA-B51 as the strongest genetic risk factor in BD, the exact mechanism of action of HLA-B51 is still unknown. Even though BD is strongly associated with HLA-B51, it is found to be not associated with HLA-B52, which differs from HLA-B51 by only two residues found on the peptide-binding region. In order to understand the effect of these variations on the dynamics of the bound and unbound forms of the two proteins, we performed comparative molecular dynamics simulations on the two MHC class I proteins in the absence and presence of sixteen different peptides. Distance distribution analysis showed that peptide binding results in a shift to a more coherent conformation in HLA-B52, while there is no significant conformational shift in -B51. Atomic fluctuation analysis additionally showed that peptide binding makes HLA-B52 more stable, which is the expected case, whereas it does not have such an effect on -B51. Based on these observations, we speculate that the instability and floppiness of the overall peptide-bound structure of mature HLA-B51 molecules may contribute to its pathogenic role in BD, by causing ER stress induced proinflammatory signaling through unfolded protein response (UPR).

ÖZET

Behçet Hastalığı (BH), kronik inflamatuar bir hastalıktır ve nedeni bilinmemektedir. BH'nın HLA-B51 ile bağlantılı olduğunun bundan kırk yıl önce bulunmasından beri BH ile ilgili yapılan neredeyse her çalışma HLA-B51'in BH ile bağlantılı olduğunu doğrulamıştır. HLA-B51'in BH'nda en güçlü genetik risk faktörü olduğu defalarca doğrulanmasına rağmen, HLA-B51'in etki mekanizması tam olarak bilinmemektedir. BH, HLA-B51 ile bağlantılı olsa bile, HLA-B51 ile arasında yalnızca peptit bağlanma bölgesinde iki amino asit farkı olan HLA-B52 ile bağlantılı bulunmamaktadır. Bu farkın her iki proteinin peptit bağlı ve serbest hallerinin dinamik özelliklerine olan etkisini anlamak için bu iki MHC sınıf I proteininin on altı farklı peptit ile bağlı ve serbest hallerinin karşılaştırmalı moleküler dinamik simülasyonlarını gerçekleştirdik. Mesafe dağılım analizi peptidin bağlanmasının HLA-B52'nin daha tutarlı bir konformasyona kaymasını sağladığını, bunun yanında HLA-B51 için böyle bir kaymanın sağlanmadığını gösterdi. Atomik dalgalanma analizi buna ek olarak peptidin bağlanmasının HLA-B52'yi beklenen şekilde daha stabil bir yapıya ulaştırabildiğini, fakat -B51 üzerinde bu şekilde bir etkisinin olmadığını gösterdi. Bu gözlemlere bağlı olarak, peptit-bağlı, olgun HLA-B51 moleküllerinin instabil ve gevşek yapılarının, tamamlanmamış protein yanıtı ile endoplasmik retikulum (ER) kaynaklı proinflamatuar sinyal oluşumuna neden olarak HLA-B51'in BH'ndaki patojenik rolüne katkıda bulunduğu sonucuna ulaşılabilinir.

ACKNOWLEDGMENTS

First and foremost, I offer my deepest gratitude to my thesis advisor, Prof. Burak Erman, who has supported me throughout these two years with his patience and knowledge whilst allowing me the room to work in my own way. I am truly honoured to have worked with him. Second, I would like to express my deep appreciation to my co-advisor, Assistant Prof. Mehmet Sayar for his useful comments, remarks and his engagement in my learning process during these two years.

I would like to extend my gratitude to my thesis committee members, Prof. Alper Demir, Prof. Ahmet Gül and Assistant Prof. Alkan Kabakçıoğlu for their participation in my thesis jury and for critical reading of my thesis.

I want to acknowledge the support of Koç University, which made the completion of this thesis possible.

A very special word of thanks goes for my parents, Nesrin and Ahmet Refhan Aydın, and my sister, Yıldız Aydın, who have made innumerable sacrifices for my education so that I could learn without limits. They have supported me in every step I took and I hope that my achievements will continue to make them proud.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

Chapter 1

INTRODUCTION

The immune system protects our body from harmful stimuli mainly by recognizing and responding to antigens. Antigens are found on the surface of living cells, viruses, fungi, or bacteria; whereas nonliving substances such as drugs, toxins, chemicals, pollens and other particles that are foreign to the body can also act as antigens. In the cell, protein molecules of the host's own phenotype or of a foreign invader are continually synthesized and degraded. Cells present these antigenic peptides to the immune system via a major histocompatibility complex molecule (MHC), known as the human leukocyte antigen (HLA) in humans. Each HLA molecule on the cell surface displays an antigenic peptide.

The HLA genes that code for these proteins are one the most polymorphic loci in the human genome. Hundreds of alleles of HLA genes are known to exist in the human population, each of which is represented by a letter and a number set, such as HLA-B51. The existence of many different alleles in the human population allows each person's immune system to react to a wide range of foreign invaders. The repertoire of HLA proteins expressed by a person varies from individual to individual as a result of genetic differences.

Each individual's immune system is tuned to a unique set of HLA and self proteins produced by that individual. HLA types are inherited, and some of them are associated with certain autoimmune disorders. Individuals expressing certain HLA

antigens are more susceptible to develop certain autoimmune diseases, such as type I diabetes and ankylosing spondylitis. In a related manner, the HLA-B51 allele is determined to be the strongest susceptibility allele for the development of Behçet's Disease (BD) [16, 17]. On the other hand, a very closely related allele to HLA-B51, HLA-B52 is found to be not associated with BD [17, 54]. BD is a chronic, multi-systemic, inflammatory disorder characterized by prolonged and recurrent aphtha or ulcers and tissue damage. Even though the clinical features and manifestations of BD are well understood [8], the pathogenesis and the aetiology of the disease still remains unclear. HLA-B51 is identified as the strongest genetic risk factor in BD, but the exact mechanism of action of HLA-B51 is also still unknown. Thus, the present study aims to investigate the differences in the dynamics of the two HLA proteins, HLA-B51 and -B52, in order to provide a better understanding towards the role of HLA-B51 on the pathogenesis of BD.

This thesis is organized as follows. In Chapter 2, a detailed literature review is presented, which starts with a review of Behçet's Disease, MHC class I molecules, the peptide presentation mechanism and the unfolded protein response (UPR) mechanism in the cell. It also includes a detailed review of the structural properties HLA-B51 and -B52. In Chapter 3, the methodologies used in this study are explained. The details of the models used in the study and the molecular dynamics (MD) simulations are presented. In Chapter 4, first, hypotheses considering the role of HLA-B51 in BD, which are based on the concepts introduced in the literature review, are presented. Then, the different computational calculations employed to investigate the differences in the dynamics of the two HLA proteins and the results of these calculations are presented. In Chapter 5, the conclusions drawn from this study are highlighted.

Chapter 2

LITERATURE REVIEW

2.1 Behçet's Disease

Behçet's disease (BD) is a systemic inflammatory disorder characterized by recurrent aphthous ulcers and ocular inflammation. The disease is a rare form of vasculitis (inflammation of the blood vessels), which can occasionally involve large vessels, joints, lungs, kidneys, the gastrointestinal tract and the central nervous system [5]. The ocular inflammation can lead to blindness in severely affected patients [6]. Since it involves blood vessels, practically any organ or tissue in the body can be affected. The disease bears similarities to an illness found in Hippocrates writings in 450 BC in ancient Greece and also the clinical signs reported in the writings of the Chinese physician Zhong-Jing Zhang in 200 AD. Later, the collection of clinical features was thought to be associated with tuberculosis or syphilis until Hulusi Behçet, a Turkish dermatologist proposed in 1937 that the symptoms, which he observed in 3 patients, might constitute a specific disease [6].

2.1.1 Diagnosis

The lack of a universally accepted pathognomonic test for BD has led to the development of an international criteria for its diagnosis [7]. The latest proposed diagnostic criteria is the International Criteria for Behçet's Disease (ICBD), which was created with the participation of 27 countries in 2006 [8]. The ICBD has recently been compared to existing diagnostic criteria and found to have better sensitivity

and specificity. It is therefore proposed to be adopted as the diagnosis/classification criteria for BD [8]. As shown in Table 2.1, ICBD uses a scoring system, in which 3 major symptoms (ocular lesions, genital aphtosis and oral aphtosis) are given 2 points and 3 minor symptoms (skin lesions, neurological manifestations and vascular manifestations) are given 1 point. Positive pathergy test is an optional criterion and is also assigned 1 point. A patient scoring 4 points or above is classified as having BD according to this scoring system [8].

Table 2.1: International Criteria for Behçet's Disease's Diagnosis.

Symptoms	Points
Ocular lesions	2
Genital aphthosis	2
Oral aphthosis	2
Skin lesions	1
Neurological manifestations	1
Vascular manifestations	1
Positive pathergy test	1

2.1.2 Epidemiology

BD is distributed worldwide, but is more prevalent in countries along the Silk Road, which has lead to its alternative name, the Silk Road Disease. The geographic distribution spans the countries of the Mediterranean to the Far East [9]. Table 2.2 presents the reported data for BD prevalence. The highest incidence is seen in Istanbul, Turkey with 420 per 100,000 population [10]. High prevalence rates are also reported from northern Turkey as 380 per 100.000 and from Ankara as 110 per 100,000 population

[11, 12]. Israel and Northern China follow Turkey in reported disease prevalence [9]. On the other hand, BD is rarely encountered in the United States [13]. The geographical distribution therefore suggests that BD is dependent on ethnic origin and related genetic factors rather than environmental factors [14], which also forms the basis for the aim of this thesis.

Table 2.2: The Prevalence of Behçet's Disease.

Country		Disease prevalence per 100,000 population
Turkey	Istanbul	420
	Northern Turkey(rural)	380
	Ankara	110
Israel		∼146
Northern China		110
Iran		80
Korea		∼30
Japan		22
USA		∼ 6

2.1.3 Aetiology and Pathogenesis

Although the cause and the development of BD is still unclear, the current understanding about the onset of the disease suggests that it is initiated by external triggers that affect the immune system in genetically susceptible hosts. It is therefore not considered as a simple hereditary disease, but rather a combinatorial disease that involves both extrinsic factors that result in abnormal immune responses and intrinsic factors [15]. The important points on the actual knowledge about the pathogenesis of the

disease can be summarized as follows:

Summary of Key Points on Actual Knowledge on the Pathogenesis of Behçet's Disease.

Genetic susceptibilities

MHC complex genes: Human leukocyte antigen(HLA)-B*51 is the most strongly associated genetic factor [16]. Genomewide association study has confirmed its role in BD susceptibility [17]

Non MHC complex genes: Interleukin-10 (IL10) and IL-23 receptor-IL-12 receptor 2 (IL23R-IL12RB2) [18], endoplasmic reticulum aminopeptidase-1(ERAP1) [19], chemokine receptor type 1 (CCR1) [19], Signal transducer and activator of transcription 4 (STAT4) [20], Killer cell lectin-like receptor subfamily C, member 4 (KLRC4) [19], GIMAP (GTPase of immunity-associated protein family). [21]

Infectious agents

Streptococci (S. sanguis) in the oral flora [22] and skin hypersensitivity to streptococcal antigens in BD patients [23].

Herpes Simplex Virus (HSV)-1 detected in peripheral blood mononuclear cells, oral ulcers and genital ulcers taken from BD patients [24].

Abnormal antigen presentation and Molecular mimicry

60-kDa heat shock protein (HSP-60) expression found to be higher in oral ulcerations

of BD patients [25]. The bacterial HSP-derived epitopes may serve as local antigens and result in inflammatory reactions [26].

Immune response against retinal S antigen seen in BD uveitis. The homology between some epitopes of retinal S antigen and HLA-B51 is thought to cause

molecular mimicry in BD [27].

Immune dysregulation and inflammatory mediators

Autoimmune responses of Type 1 and type 2 helper T cells (Th1 and Th2) [26], high predominance of natural killer (NK) cells, certain T lymphocytes, neutrophils and macrophage cells in BD lesions [27].

Oxidative stress

Increased activated neutrophil function, increased production of hydrogen peroxide induced hydroxyl radical and reactive oxygen species reported in BD patients [26].

2.2 Association of HLA-B51 with Behçet's Disease

2.2.1 MHC class I molecules

Studies identified HLA-B51, one of the split antigens of HLA-B5, as the most strongly associated genetic marker with BD [16, 17]. HLA-B51 is a Major Histocompatibility Complex (MHC) Class I molecule. MHC is known as the human leukocyte antigen (HLA) in humans, and it is a large genomic region found in all vertebrates that encodes MHC proteins. MHC is considered as the most polymorphic protein in higher vertebrates; the July 2014 release of the IMGT/HLA database lists more than 12000 sequences of Class I and Class II MHC molecules [28]. MHC molecules have an important role in the immune system in the sense that they mediate the interactions between immune and body cells. They present on the cell surface the antigenic determinant of a protein, called epitope, to immune cells. These epitopes can either be antigens originating from the organism itself (self), or from the external environment (non-self). During antigen presentation, immune cells distinguish between self and non-self epitopes, and under normal conditions, the recognition of a non-self epitope results

in the activation of several types of immune responses. Conversely, self-antigens also lead to immune responses, which is known as autoimmunity. In autoimmunity, immune cells fail to differentiate between self and non-self epitopes. Abnormally strong immune responses induced towards self-antigens are responsible for tissue destruction in autoimmune diseases [29] and organ rejections after transplantations [30].

MHC class I molecules are composed of three subunits, as shown in Figure 2.1: a

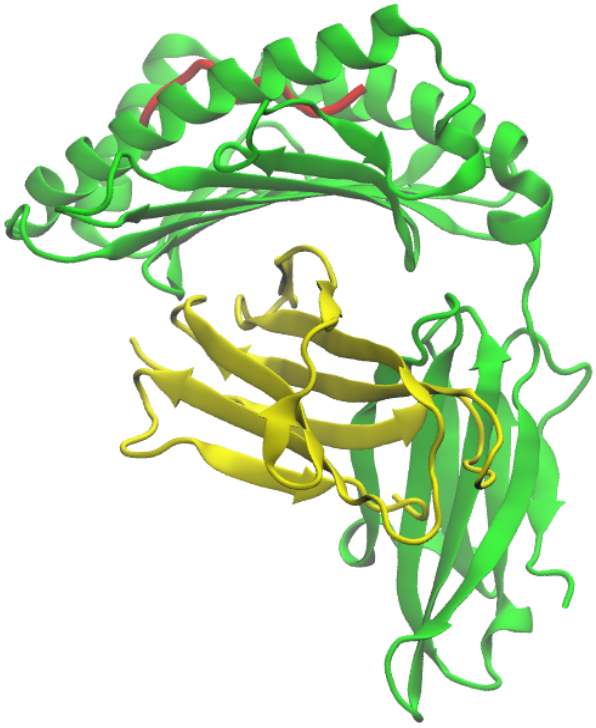


Figure 2.1: Structure of HLA-B51, an MHC class I molecule (PDB ID: 1E27). It is composed of three subunits: the MHC heavy chain (green), the β2-microglobulin (β2m) subunit (yellow) and a bound peptide ligand (red).

45 kDa membrane-integrated glycoprotein heavy chain (HC), which is the polymorphic product of an MHC-linked gene, a 12 kDa soluble protein subunit called beta-2 microglobulin (β2m), and a short peptide ligand [31, 32]. They can bind peptides with 8-11 amino acids, while 9-mer peptides are the most common [33]. β2m subunit is an obligate component of the MHC class I molecules; the expression of MHC class I molecules on the cell surface does not take place without it [34]. Its absence results in a misfolded heavy chain, and the quality control mechanism of the ER ensures that it is translocated into the cytoplasm and degraded by the proteasome [35, 36].

The stability and function of MHC class I proteins depend on their multiple-step assembly, which includes initial folding and β2m association, as well as the peptide loading and the binding affinity of these peptides. It is therefore important to understand the processes which a nascent heavy chain undergoes to be assembled into an MHC-I-peptide complex enabling peptide presentation.

Figure 2.2 shows how peptide loading and presentation happens in the cell. Proteins are continuously synthesized in the cytosol and degraded through the action of large proteolytic complexes called proteasomes. The resulting peptides are transported to the endoplasmic reticulum (ER), by the transporter associated with antigen processing (TAP) [32]. The imported peptides can further be trimmed from their N-terminal by the ERAAP (ER aminopeptidase associated with antigen processing, or ERAP1) enzyme in order to optimize the length of the peptide before loading [37]. Peptides that match the binding criteria of the MHC class I binding pocket are loaded on the MHC molecule. This process requires the interaction of additional chaperone-like components such as calnexin, calreticulin and tapasin, the ER-membrane protein Bap31, and the protein disulfide isomerase ERp57, which, together with the MHC-I heavy chain-β2m and TAP, form the peptide loading complex (PLC) [32]. Unloaded or suboptimally loaded MHC-I molecules stay within the PLC, retained in the ER. Once they are loaded with peptides with optimal affinity, MHC-I heavy chain-β2m-

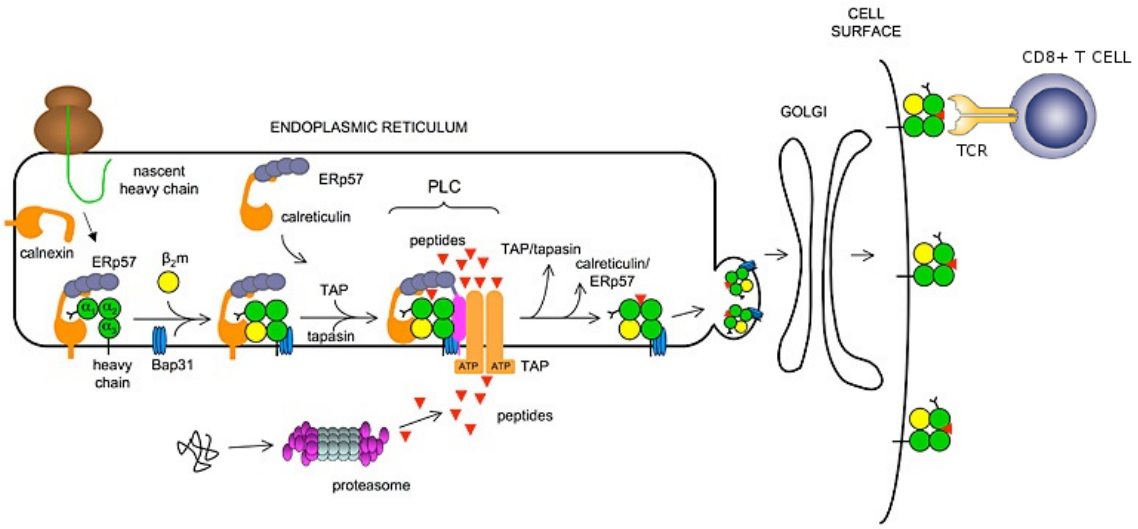


Figure 2.2: Overview of the peptide loading and presentation mechanism in the cell.

peptide complexes (pMHC) detach from the PLC into the Golgi complex and are exported to the cell surface [38]. At any given time, a single cell displays between 1000 to 10000 different peptides [39]. At last, after peptide presentation, the peptide and the β2m subunit dissociates and the free heavy chain is translocated to lysosomes and degraded [40].

Protein degradation is an important step in the peptide loading and presentation process, because it must result in peptides that have the right length for binding to TAP during their transportation to the ER and their future binding to MHC class I molecules. Additionally, protein degradation must take place in a proper speed, since fast replicating pathogens must be detected rapidly in the cell. In case protein degradation does not take place, MHC-I molecules are deprived of peptides and they

stay in the ER for longer periods [31].

Tapasin (TAP-associated glycoprotein) is a transmembrane glycoprotein which is one of the most important components of the PLC. It mediates the interaction between newly-assembled MHC class I molecules and the TAP transporter. Tapasin initially permits the binding of a diverse set of peptides to MHC-I molecules by stabilizing a peptide-receptive conformation [62]. It edits the repertoire of bound peptides [63] by excluding the peptides which cannot meet a certain minimum affinity threshold to conformationally disengage tapasin [62]. Additionally, it governs the success of the TAP-MHC interaction; weak MHC class Itapasin interactions are known to be the cause of inefficient interaction of of some HLA alleles with TAP [64, 65].

Human natural killer (NK) cells play a vital role in the innate immune system, providing fast responses to virus-infected cells and to tumor formation. They express several polymorphic receptors, one of which is the killer cell immunoglobulin-like receptor (KIR) family, which recognize and interact with HLA class I molecules. Each family member interacts with distinct HLA class I allotypes [69]. A member of this group of receptors found on NK cells, KIR3DL1, recognizes HLA-B alleles that contain the Bw4 serological epitope composed of residues 77-83 found on the α2-1 helix, which are also shared by the HLA-B51 and -B52 alleles [70, 71].

MHC class I molecules are found on antigen-presenting cells (APCs) and they present peptides to CD8-expressing cytotoxic T lymphocytes (CTLs), which are white blood cells that destroy infected cells (Figure 2.2). CTLs express T-cell receptors (TCRs) on their surface that help them recognize specific antigens; however, TCR-pMHC interaction alone is not enough for the induction of T cells. The enhancement of T cell signalling requires the participation of the TCR-coreceptor, CD8, which is a transmembrane glycoprotein expressed on CTLs. This results in the ternary complex of TCR-pMHC-CD8, where the CD8-pMHC interaction ensures the T cell's recognition specificity towards pMHC ligands [41].

The Unfolded Protein Response

In eukaryotic cells, proteins enter the ER as unfolded polypeptide chains and they fold and mature in the lumen of the ER. In order for proteins to fold properly, a balance between the ER protein load and the folding capacity must be established. The disruption of this balance in the ER results in an accumulation of misfolded and unfolded proteins, a condition known as ER stress [42].

The flux of proteins into the ER is variable and cells adjust the protein-folding capacity of the ER according to their requirements to deal with this dynamic situation. Such homeostatic control is attained through the action of signal transduction pathways [43]. Kozutsumi and colleagues have observed that the accumulation of unfolded ER proteins activates the expression of genes encoding ER-resident chaperones that assist in protein folding, which was the first clue to the existence of intracellular homeostatic signalling events [44]. The intracellular signalling pathway that mediates this regulation was named the unfolded protein response (UPR) [43].

UPR therefore can be defined as a response to reduce ER stress and to restore homeostasis. It enables eukaryotic cells to respond to changing conditions in the ER by regulating the synthesis of ER-resident proteins. Three main responses are observed in the UPR network. The first action is the reduction of the protein load that enters the ER, which is achieved by reducing protein synthesis and translocation to the ER. The second action is the increase in the folding capacity of the ER, which is achieved by the transcriptional activation of UPR target genes. If these two actions fail to re-establish the homeostasis, cell death is triggered in order to prevent the production of rogue cells that express misfolded proteins [43]. UPR can therefore be considered as a binary switch between the life and death of ER stressed cells [42].

ER-stress mediated cell dysfunction and death is involved in the pathogenesis of a variety of human diseases, including diabetes [45], inflammation [46], neurodegen-

erative disorders [47], such as Alzheimer's disease, Parkinson's disease and bipolar disorder, which are collectively known as 'conformational diseases' [48]. Studies conducted on rats transgenic for HLA-B27, an MHC class I molecule, have determined ER stress and UPR as possible pathogenic mechanisms associated with HLA-B27 in ankylosing spondylitis, another MHC class I-associated inflammatory disease [49, 50]. ER stress and UPR might also have a role in a similar manner in the pathogenesis of HLA-B51-associated BD, even though there exists no evidence supporting such association so far.

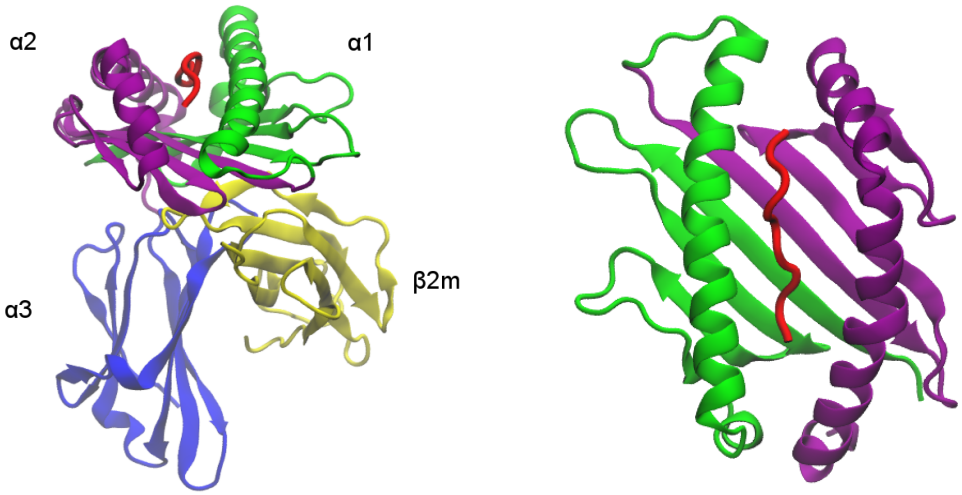
2.2.2 HLA-B51 and -B52

The HLA region is found on chromosome 6, which constitutes of three genetic loci that encode HLA class I subtypes HLA-A, HLA-B and HLA-C [51]. Among the subtypes, HLA-A and HLA-B are the most polymorphic proteins [52]. The polymorphic variance is necessary in order for HLA-A and HLA-B to bind to a large number of peptides. Amino acid changes associated with different alleles are usually located around the antigen binding groove, which provides unique spatial and chemical characteristics to the groove and increases the peptide repertoire. Each individual expresses a small number of MHC molecules; however, the diverse repertoire of peptides and the capability of each peptide to bind to more than one HLA molecule leads to a wide variety of peptide-MHC complexes, enabling the immune system to respond to numerous triggers [51].

The role of HLA proteins is therefore to mediate protective immunity by presenting pathogen-derived peptides to killer T cells. However, instead of being associated with protective immunity in infectious diseases, certain HLA proteins are found to be much more strongly associated with inflammatory, autoimmune-like diseases [53]. HLA-B*5101, the most frequent allele of HLA-B5, was identified as the most strongly

associated genetic marker with BD [16, 17]. HLA-B*52, another split antigen of HLA-B5, differs by only two amino acids from HLA-B51. While BD is strongly associated with HLA-B51, it is interesting that it is nearly found to be not associated at all with HLA-B52, with the few exceptional cases reported [17, 54]. The HLA-B sequence encoding HLA-B*5101 has no variation specific for BD, therefore HLA-B51's pathogenic role in the disease should be related to its structural and functional properties [55, 56].

Structural properties of HLA-B51 and -B52



(a) (b)

Figure 2.3: (a)The four domains of HLA-B51. The heavy chain consists of three subunits: $\alpha 1$ (green), $\alpha 2$ (purple) and $\alpha 3$ (blue).(b)The $\alpha 1$ and $\alpha 2$ domains fit together and form the peptide binding groove.

There are currently two available crystal structures of HLA-B*5101 in the Protein

Data Bank (PDB). The pdb entry 1E27 is HLA-B51 complexed with a 9-mer, LPPVVAKEI, peptide from HIV-1 and 1E28 is, similarly, HLA-B51 complexed with an 8-mer, TAFTIPSI, peptide from HIV-1. Both structures are determined with X-ray crystallography [57]. These extracellular structures comprises three chains: the HLA-I heavy chain, the β2m light chain, and the peptide; which are referred to as chains A, B and C, respectively (Figure 2.1). The heavy chain, which consists of 276 residues, is divided into three domains called α1, α2 and α3 (Figure 2.3a). The β2m light chain consists of 99 residues and it supports the α1 and α2 domains. While α3 interacts only with β2m, β2m interacts with all 3 domains of the heavy chain, which explains the earlier statement that β2m is an obligate component of the MHC-I proteins. The binding site for antigenic peptides is found in the α1-α2 domains of the protein, in the deep cleft located between the two alpha helical regions. α1 and α2 domains each consist of a four-stranded antiparallel β-pleated sheet, which is the N-terminal region of the domain, followed by an α-helical region. The two domains fit together to form a binding pocket where the two antiparallel α-helices sit on top of the platform formed by the eight-stranded β-pleated sheet (Figure 2.3b)[58].

Figure 2.4 and Table 2.3 give more detailed information on the secondary structure elements of the α1 and α2 domains. While the α1 domain has a long curved α-helix (H2, green), the α2 domain has a kinked helix (H2a (pink) and H2b (purple)). Both helices are preceded by a short, nearly vertical helix (H1 for both domains, white). These short helices rise from the lowest corners of the β-sheet platform and connect with the second helices after a sharp turn, following down the twist of the β-sheet on opposite sides. The α2 domain ends with a one-turn helix (H3, blue), right before the start of the α3 domain [2].

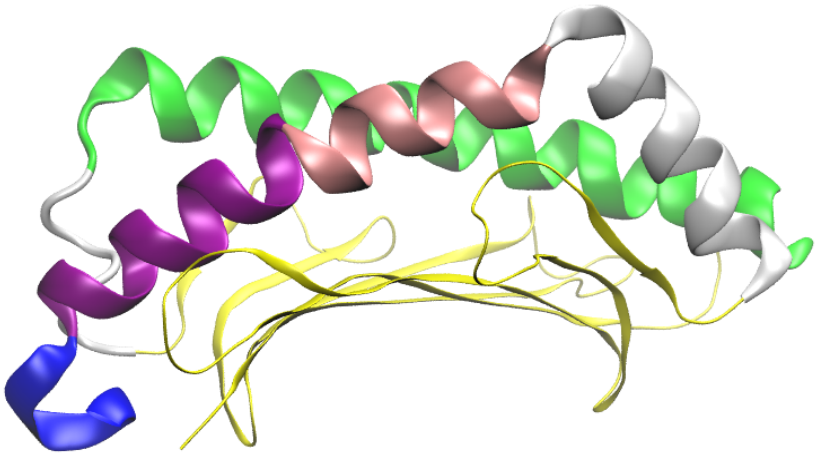


Figure 2.4: The α1 and α2 domains on the peptide-binding groove of HLA-B51. Short vertical helices (H1 for both, white) rise from the opposite lowest corners of the β-sheet platform. The second helices start after a sharp turn. In the α1 domain, the second helix is a long α-helix (H2, green), whereas for the α2 domain it is a kinked helix (H2a (pink) and H2b (purple)). The α2 domain ends with a one-turn helix (H3, blue), right before the start of the α3 domain.

The peptide-binding groove

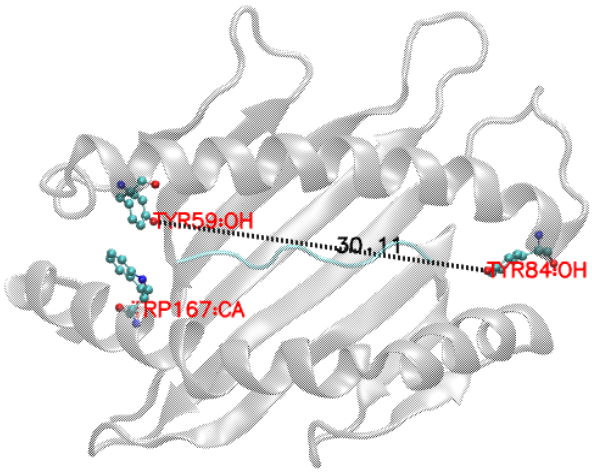
The α-helical regions from α1 (residues 59-84) and α2 (residues 143-171) form the walls of the peptide binding groove. Although the site is open at both ends, the α-helices are curved and thus approach each other at the two ends, setting limits for the peptide binding cleft. The cleft is approximately 30 Å long (Figure 2.5a). As seen in Figure 2.5a, on the left side, the cleft ends where the Trp167 indole ring contacts

Table 2.3: The secondary structure elements of the α1 and α2 domains of HLA-B51. (Intervals are inclusive)

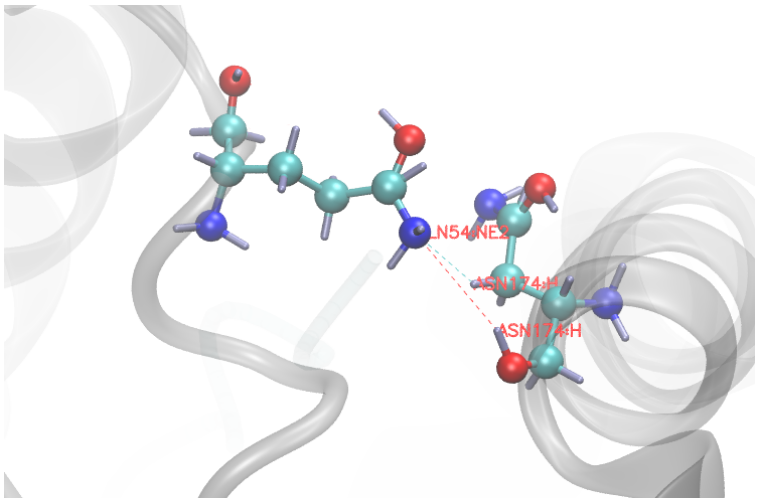
Strand	Residues	Strand	Residues
α1		α2	
S1	3-12	S1	94-103
S2	21-28	S2	109-118
S3	31-37	S3	121-126
S4	45-47	S4	133-135
H1	50-53	H1	138-150
H2	57-85	H2a	152-161
		H2b	163-174
		H3	176-180

Tyr59. Farther on the left of this, the salt bridge between Glu55 and Arg170 and the hydrogen bonds between Gln54 and Asn174 further keep the two helices together (Figure 2.5b). [2].

A pocket in the peptide binding groove is defined as the unit with an individual affinity to the corresponding side chain (or residue) of a peptide. Peptides are anchored onto pockets found in the peptide-binding cleft of HLA proteins only at a few positions of the peptides. These positions are called "anchors", and are usually at P2 or P3 and PC (C-terminus). Other side chains of the bound peptide also make contact with the groove, but not apparently constrained by any specific pocket. A deep pocket is one that has residues predominantly accessible to a 1.4 Å probe but not to a probe >3.5 Å. A shallow pocket has residues mainly accessible to a 3.5 Å probe but not to a 5.0 Å probe [2].



(a)



(b)

Figure 2.5: (a)The α-helices form the walls of the peptide binding groove of HLA-B51. The cleft is approximately 30 Å long. On the left side of the cleft, the cleft ends where the Trp167 indole ring contacts Tyr59 [2].(b)The hydrogen bonds between Gln54 and Asn174 further help to keep the two helices together [2].

There are six specificity pockets on HLA-B51, which are denoted with letters A-F. Four of these pockets (A, B, C and F) are deep pockets and two of them (D and E) are shallow [2].

The side chains at P1, P2, P3, P6, P7, and P9 fit into pockets A, B, D, C, E, and F, respectively [2]. Pocket A accommodates the N-terminus of the peptide. The residues that form up this pocket form hydrogen bonds with the amide group and carbonyl oxygen of the first peptide residue. Pocket B is the P2 binding site and it is found beneath the α1 helix and it is formed by the residues 9, 45, 63, 66, 67, 70 and 99. Pocket C forms the mid-cleft region. Pocket F holds the C-terminus of the peptide, where the carboxyl group of the last residue (PC) and the carbonyl oxygen of the next-to-last residue make hydrogen bonding with side chains of this pocket. It consists of residues 77, 80, 81 and 116. Residue 116 (Tyr for HLA-B51) is a critical residue of the F pocket, as it is found at the bottom of the cleft, with its side chain pointing up at the PC of the bound peptide [59]. The presence of a tyrosine at this location affects the nature of the PC and selects for the presence of a hydrophobic residue at the C-terminus of the peptide, as explained in more detail below.

Peptide motifs

Asparagine (Asn, N) and phenylalanine (Phe, F) in HLA-B51 at positions 63 and 67 are changed to glutamic acid (Glu, E) and serine (Ser, S), respectively, in HLA-B52. The two residues are found in the α1 helix, and are located at the B pocket of the peptide binding groove. Gül et al. suggested that changes in the B pocket may affect the motif of peptides that are able to bind to the pocket, and presentation of certain Behçet's disease-related peptides with their distinctive B- and F-pocket features may be one of the HLA-B51- associated pathogenic mechanisms in BD [60].

Lemmel et al. performed a mass spectrometric sequencing analysis of 22 peptides

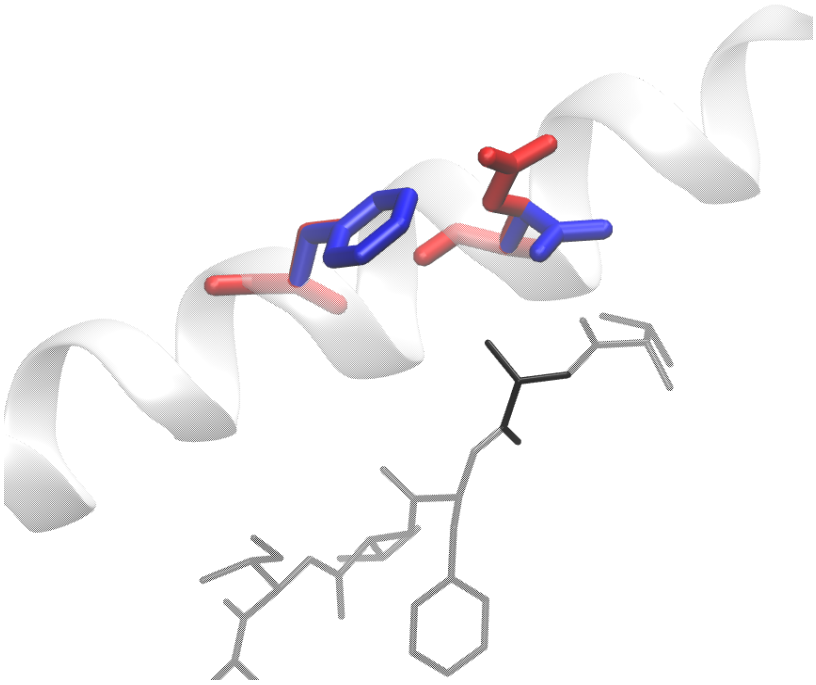


Figure 2.6: Comparison of the B pocket side chains of HLA-B51 and -B52. Asn63 (right, blue) and Phe67 (left, blue) of the α1 helix create steric hindrance in the B pocket of HLA-B51. They are exchanged with Glu63 (right, red) and Ser67 (left, red) in HLA-B52. The B pocket of HLA-B51 is less spacious and thus is able to accommodate only smaller residues.

eluted from HLA-B*5101 molecules from cells expressing this antigen. As a result of the analysis of these peptides, it was found that the amino acids at positions 2 and 9 are the anchor residues of the HLA-B51 derived peptides. The motif they obtained suggests that position 2 (P2) can be occupied by the small amino acids Alanine (A) and Proline (P) and less preferably by Glycine (G) , and position 9 (P9, or PC) can be occupied by the small hydrophobic residues Isoleucine (I) and Valine (V). These residues are critical in binding to the HLA-B51 molecule at the B and F pockets of

antigen binding groove, respectively [55].

Figure 2.6 shows the effect of the two amino acid differences between HLA-B51 and -B52 on the B pockets of the proteins. The B pocket of HLA-B51 is moderately blocked by the bulky aromatic amino acid phenylalanine at position 67, which, along with the kinked side chain of asparagine at position 63 occludes the pocket, making it suitable only for compact, nonpolar amino acids at the P2. On the other hand, for HLA-B52, the small, hydrophilic amino acid serine at position 67 results in a spacious hydrophilic B pocket. HLA-B51 should therefore present a different set of peptides due to the differences in the nature of the B pockets, which might be one of the key points in explaining the association of HLA-B51 with the disease [61].

Chapter 3

METHODS

Most theoretical descriptions of natural phenomena rely heavily on analytical and numerical approximations. Computer simulation methods are used instead of analytical approximations to explore the conformational energy landscape of a molecular system in equilibrium. Simulations illustrate how a molecular system changes from one conguration to another. A simulation be considered as a virtual laboratory in which the behavior of a system can be visualized and predicted. In this respect, computer simulation is a link between experiment and theory. Simulations are able to provide much detail about individual atomic motions that they can be used to find answers to specific questions about the properties of a model system more easily than experiments done on the actual system.

Molecular dynamics (MD) is a computational method where the dynamics of a number of particles is simulated by numerical integration of Newton's equation of motion (Eq. 3.1) for discrete timesteps Δt on the femtosecond timescale, in order to obtain a trajectory of positions $r_i(t)$ and velocities $\dot{r}_i(t)$. Each point in the phase space represents a unique set of positions and velocities. Initial velocities are usually randomly assigned from a Maxwell distribution for the desired temperature, and initial positions are assigned randomly or in a lattice. Given the initial set of conditions, the net force $F_i(t)$ acting on a particle is then calculated by summing up the forces that particles exert on each other. These forces arise from the interaction of particles, and hence the changes in the potential energy as a function of the separation distances of particles. Once the net force is calculated for each particle, its position and velocities

are updated for a small increment in time and the procedure is repeated.

$$\ddot{r}_i = \frac{F_i(t)}{m_i} \quad \text{where} \quad F_i(t) = \sum_j -\frac{\mathrm{d}V_j}{\mathrm{d}r}(r_i) \tag{3.1}$$

MD simulations provide detailed information about the fluctuations and conformational changes of proteins and nucleic acids and therefore, they are a useful tool to investigate the structure, function and dynamics of biological molecules. Molecular Dynamics (MD) simulations have extensively been used to study MHC-peptide complexes, as a complementary tool to experimental methods [73, 52, 74, 75]. MD simulations provide an accurate dynamic view of molecular behavior, serving as a link between microscopic intermolecular interactions and experimental macroscopic quantities.

In this work, we aimed to point out the differences in dynamics between two HLA-B alleles by using computational approaches. While Behçet's Disease (BD) is strongly associated with HLA-B51, it is found to be not associated with HLA-B52 [17, 54]. The only difference between the two proteins is two residues. Asn63 and Phe67 in HLA-B51 are changed to Glu63 and Ser67 in HLA-B52. The two residues are found in the α1 helix, located at the peptide binding groove. The HLA-B sequence encoding HLA-B51 has no variation specific for BD, therefore HLA-B51's pathogenic role in the disease should be related to its structural and functional properties [55, 56]. MD simulations of HLA-B51 and -B52 in water were performed and the fluctuations of residue positions were compared. The structural differences induced by the mutation of two residues are responsible for the behavioural and functional differences of HLA-B51 and -B52 molecules.

The initial structures used in MD simulations were the two available crystal structures of HLA-B*5101 in the Protein Data Bank (PDB). The PDB entry 1E27 is HLA-B51 complexed with a 9-mer, LPPVVAKEI, peptide from HIV-1 and 1E28

is, similarly, HLA-B51 complexed with an 8-mer, TAFTIPSI, peptide from HIV-1. Both structures are determined with X-ray crystallography [57]. These PDB crystal structures served as starting structures for the simulations with and without bound peptides. 16 peptides were chosen by taking into consideration the peptide motif of HLA-B51 (Alanine (A) or Proline (P) at position 2; Isoleucine (I) or Valine (V) at the C-terminal) [55]. 5 of these peptides were 8-mer and 11 were 9-mer peptides. To model the 8-mer peptides complexed with HLA-B51, the PDB structure of HLA-B51 complexed with an 8-mer peptide, 1E28 was taken as the initial structure and the peptide was mutated to obtain the desired one by using Discovery Studio (Accelrys Inc., San Diego, CA, USA). Similarly, to model the 9-mer peptides complexed with HLA-B51, the PDB structure of HLA-B51 complexed with a 9-mer peptide, 1E27, was used as the initial structure and the peptide was mutated. To obtain the same complexes with HLA-B52, the obtained structures were mutated at positions 63 and 67. Finally, to obtain the unbound form of HLA-B51 the peptide was removed from 1E27, and was mutated at positions 63 and 67 to obtain the unbound form of HLA-B52.

All Molecular Dynamics simulations were performed by using the 4.5.5 version of the Gromacs MD simulation package [76] with the CHARMM27 [77] forcefield, which is a forcefield suitable for protein systems. The proteins were soaked in a triclinic box of TIP3P [78] water molecules and simulated using periodic boundary conditions in all three directions. The distance between the edges of the box and the closest atom in the protein was 0.8 nm. Water molecules were substituted by the necessary number of Na^+ and Cl^- ions in order to neutralize the overall system charge and to represent a more typical biological environment. In all, the solvated HLA complex system contained around 48500 atoms.

Unfavorable interactions within the structures were relieved with steepest descent energy minimizations. Two minimizations, first after solvating the protein and second after adding ions, were carried out for each protein/protein-peptide complex to op-

Table 3.1: 8 and 9-mer epitopes used in MD analysis.

No	Peptide
1	YAYDGKDYI
2	TAFTIPSI
3	DAFKIWVI
4	YPDRVPVI
5	YPFKPPKI
6	IPYQDLPHL
7	LPRSTVINI
8	ISWPFVLLI
9	FPRCIFSAI
10	LPSIPVHPI
11	LPSPACQLV
12	MPNQAQMRI
13	MASSPTSI
14	MAWERGPAL
15	FPHTELANL
16	SPASFFSSW

timize each initial solution structure. Both minimizations were set to continue until converging to a smaller value than a defined maximum force value (emtol=100 kJ mol^{-1} nm^{-1}).

Productive MD simulations were performed in the isobaric-isothermal (NPT) ensemble at physiological temperature of 310 K, controlled by the velocity-rescaling thermostat [79] with a temperature coupling time constant of 0.5 ps. Pressure was

kept constant at 1 bar by the Berendsen barostat [80] and the time-constant for pressure coupling was set to 1 ps. The equations of motion were integrated with a 2 fs time step. All bond distances were constrained using linear constraint solver (LINCS) algorithm [81]. Electrostatic interactions were calculated using the Particle-mesh Ewald (PME) [82] summation scheme. The cutoff distance for Coulomb interactions was set to 1 nm and for van der Waals interactions it was set to 1.4 nm. Conformations and energies were stored every 10 ps. Productive MD simulation length was 200 ns for each HLA complex. 34 simulations were carried out in total with this length scale and all calculations that will be explained further on were performed for all 34 simulations.

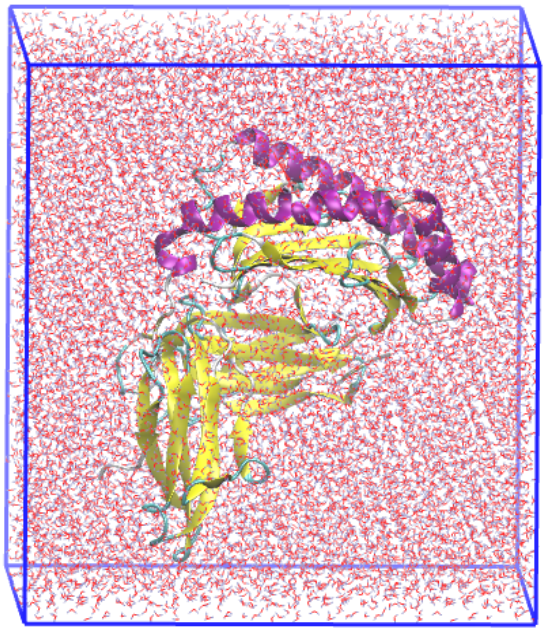


Figure 3.1: The triclinic box containing the solvated HLA complex.

In the post-production phase, for each trajectory, the system was centered in the box with respect to the HLA heavy chain and periodicity effects were removed. The number of frames in the output was decreased by taking a frame each 100 ps. The structure was fitted to the first frame structure of this processed trajectory to eliminate rotational and translation degrees of freedom. Only the carbon alphas were used for the least squares fit. The first frame of the second processed trajectory was then taken and initial atomic coordinates in future calculations were obtained from this structure.

Chapter 4

RESULTS

Behçet's Disease (BD) is a chronic inflammatory disorder and its aetiology is unknown. It has long been noted that it is initiated by external triggers, such as microbial pathogens, that affect the immune system in genetically susceptible hosts. Nearly every genetic study related to BD has confirmed the association of HLA-B51 with BD since the time the association was first recognized forty years ago [83]. Even though studies have expanded the list of possible genetic risk factors for BD [18, 19, 84, 85], HLA-B51 still remains as the strongest genetic association identified so far. The study conducted recently by Ombrello and colleagues [86] with a very large collection of Turkish BD patients has once again affirmed this association.

Despite the recurrent affirmation of HLA-B51 as the strongest genetic risk factor in BD, the exact mechanism of action of HLA-B51 is still unknown. Different hypotheses have been proposed and considered to explain the role of HLA-B51 in the pathogenesis of BD.

The first set of hypotheses is antigen-dependent; they are based on the peptide repertoire of HLA-B51 and peptide presentation to $CD8^+$ T cells. Conformational changes in the antigen binding groove determines the sequence of peptides, which could be bound and presented to cytotoxic T cells. Gul and Ohno [87] have suggested that the presentation of certain Behçet's disease-related HLA-B51 peptides with their distinctive B- and F-pocket features may be one of the causative factors for Behçet's disease. Gebreselassie et al. [88] have shown HLA-B51 as capable of binding a large pool of peptides with lower affinity and have suggested that the looser global asso-

ciation of HLA-B51 with peptides may lead to a promiscuous behavior in peptide presentation to $CD8^+$ T cells. Yasuoka et al. [89] have reported an increased proportion of $CD8^+$ T cells in patients with active BD, exclusively in those with HLA-B51, compared to patients with inactive BD or healthy controls. However, no *in vivo* BD-specific pathogenic immune response or HLA-B51 restricted $CD8^+$ T-cell activation has been shown so far in BD [87].

The second set of hypotheses is antigen-independent; they take into consideration the structural and folding properties of HLA-B51 itself. Gül and Ohno [87] have suggested that the inflammation may be directly related to the structural and peptide-binding properties of HLA-B51. As discussed in the second Chapter, the Unfolded Protein Response (UPR) is a signalling network response to reduce the cellular stress caused by the accumulation of misfolded and unfolded proteins in the ER. HLA-B51 is classified as a slow-folding protein [90], which is shown to be a result of its intrinsic structural and peptide binding characteristics [57]. Binding of lower affinity peptides may lead to slow assembly and misfolding of HLA-B51 [33], which may lead to the instability of mature HLA-B51 molecules and induce a proinflammatory ER stress response through UPR [87].

Binding of peptides to MHC class I molecules has important implications for the pathophysiology of BD. Peptide loading is an important step in the MHC-I molecule assembly that affects the stability of HLA proteins. Under normal conditions, peptide loading happens in the ER, after the peptide-loading complex (PLC) is formed. Unloaded or suboptimally loaded MHC-I molecules are retained in the ER, which implies that peptide loading also determines the molecule's surface expression versus retention in the ER. If peptide loading fails, the "empty" heavy chains display unstable conformations at physiological temperature. These aberrantly folded heavy chains are not trafficked any further to their intended destination and are degraded [91, 92, 93]. Bouvier and Wiley [94] showed that peptide binding is a process that

stabilizes the structure of the peptide binding domain and that peptide binding to the class I binding cleft is mainly mediated through contacts of the peptide N- and C-termini. These termini are held by hydrogen bond networks and they make the highest contribution to the binding energy of the peptide. In a subsequent study, they demonstrated that empty MHC class I molecules are more susceptible to proteolytic cleavage, in particular at spots close to the peptide-binding cleft [95]. Zacharias et al. [52] performed MD simulations of the empty α1-α2 domain of HLA-A*0201 and showed that the α-helical regions in the peptide binding groove make large amplitude movements, including bending and kinking motions, and undergo partial unfolding. One objective of the study presented in this thesis was thus to investigate the differences in dynamics between the unbound and peptide-bound HLA-B51 and -B52.

The second important event in the completion of MHC-I assembly is the binding of β2-microglobulin (β2m) to the heavy chain. Zijlstra et al. [96] and Koller et al. [97] have reported that mice with disruption of the β2m gene do not express MHC class I molecules and have defects in $CD8^+$ T cell development. Williams et al. [34] and Seong et al. [98] have similarly reported that cell surface expression of MHC class I molecules does not take place in mutant β2m-deficient cell lines. β2m therefore is considered as an obligate component of the MHC class I molecules. When free of β2m, MHC-I heavy chains are thermally unstable and acquire linear epitopes, but they becomes thermally stable and display conformational epitopes once assembled with β2m and peptide [99]. The functionality and stability of mature MHC-I molecules are therefore dependent on the simultaneous presence of β2m and peptide.

Even though BD is strongly associated with HLA-B51, it is found to be not associated with HLA-B52 [17, 54]. The two alleles differ in two residues: 63 (Asn in B51 is changed to Glu in B52) and 67 (Phe in B51 is changed to Ser in B52), which are found on the peptide-binding region. In order to understand the effect of these variations on the dynamics of the bound and unbound forms of the two proteins,

comparative molecular dynamics simulations on the two MHC class I proteins in the absence and presence of different peptides were performed.

When analyzing molecular dynamics simulations, it is crucial to monitor the quality of the simulation and the stability of the protein. The root mean square deviation (RMSD) is a numerical measure of how much a conformation deviates from a known initial reference structure (Eq. 4.1), which is usually an X-ray or NMR structure. It reveals the extend to which the system has equilibrated.

$$\mathrm{RMSD} = \left(\frac{1}{N} \sum_{i=1}^{N} (r_i(t) - r_i(0))^2 \right)^{1/2} \tag{4.1}$$

Here, $r_i(t)$ is the position of atom i at time t, $r_i(0)$ is the position of atom i in the reference structure and N is the number of atoms in the system. An RMSD value that is around 0.2 nm for protein backbone atoms or α-carbons is an indication of a stable, equilibrated system and is a quality control measure for protein simulations. All performed simulations showed a stable α-carbon RMSD lying in the range 0.2 - 0.3 nm with respect to the starting structure. None of the peptides, nor the β2m components dissociated spontaneously. Since the structure was fitted to the initial configuration in the post processing operations, no fitting was employed for the RMSD calculations. Figure 4.1 shows a representative RMSD plot that belongs to HLA-B51 and HLA-B52 complexed with the peptide MASSPTSI. Through RMSD calculations it was ensured that the protein systems are equilibrated in the first stage of the analysis of the trajectories.

The root mean square deviation from the reference structure over time can be referred to as the root mean square fluctuation (RMSF). While RMSD is an average that is taken over particles, RMSF is an average taken over time (Eq. 4.2). Root mean square fluctuation calculations were performed for all complexes, where only the α-carbons were taken into consideration. Since the structure was fitted to the

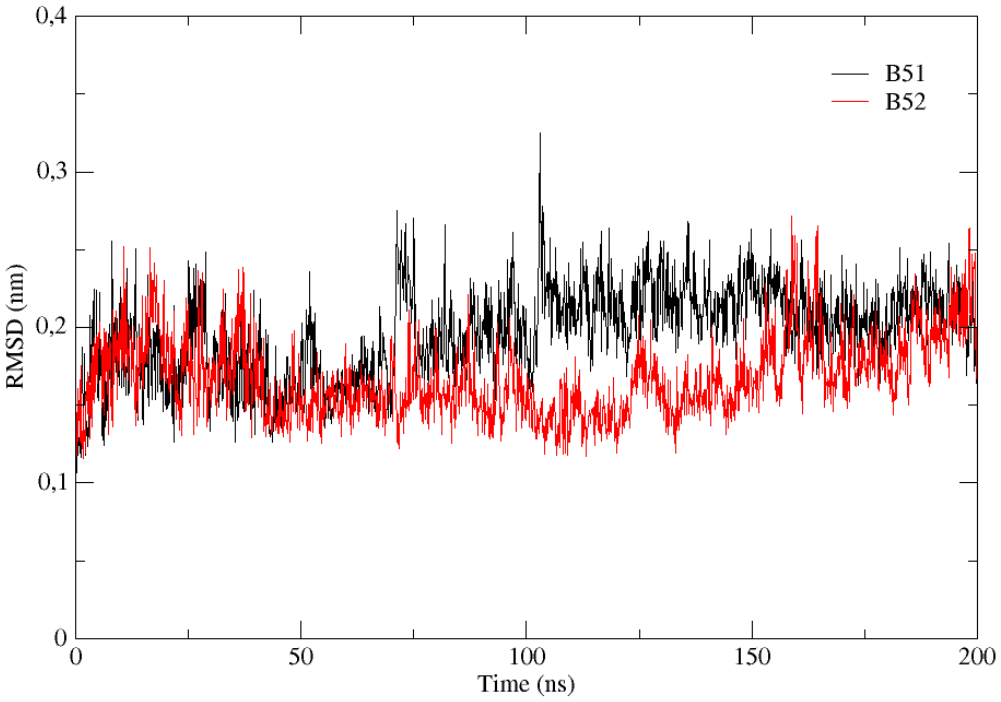


Figure 4.1: An RMSD value that is around 0.2 nm for protein backbone atoms or α-carbons is an indication of a stable, equilibrated system. All performed simulations showed a stable α-carbon RMSD lying in the range 0.2-0.3 nm with respect to the starting structure. This representative RMSD plot shows that the RMSD of α-carbons fluctuates at around 0.2 nm over the 200 ns trajectory. The black line belongs to HLA-B51 complexed with the peptide MASSPTSI and the red line belongs to HLA-B52 complexed with the same peptide.

initial configuration in the post-processing operations, no fitting was employed for the RMSF calculations. Three RMSF calculations were made for each HLA complex: (i) RMSF of the protein (without the peptide, 375 residues), (ii) RMSF of residues belonging to the secondary structure elements in the α1-α2 domains (Table 2.3) and

(iii) RMSF of the peptide (only for the peptide-bound proteins).

$$\text{RMSF} = \left(\frac{1}{T} \sum_{t_j=1}^{T} (r_i(t_j) - r_i(0))^2 \right)^{1/2} \tag{4.2}$$

The comparison of RMSF plots for HLA-B51 and -B52 unbound and bound to different peptides reveals that peptide-bound HLA-B51 fluctuates the most, regardless of the peptide sequence. A general trend observed is that while peptide binding decreases the fluctuation of HLA-B52, it increases the fluctuation of HLA-B51. Especially with certain bound peptides, the difference of the two proteins upon peptide binding is remarkable. Figure 4.2 and 4.3 show a conspicuous example to this difference. Figure 4.2 is the RMSF plot for HLA-B51 and -B52 unbound and bound to the 9-mer peptide MPNQAQMRI. The solid black and red lines belong to the peptide-bound -B51 and -B52, and dashed black and red lines belong to the unbound -B51 and -B52 respectively. It can be seen that peptide-bound HLA-B51 fluctuates the most and peptide-bound HLA-B52 fluctuates the least. As previously mentioned in this Chapter, peptide binding is a process that stabilizes the structure of an HLA protein, especially its peptide binding domain. Figure 4.3a and 4.3b show the effect of peptide binding on the secondary structure elements of the peptide binding domains of HLA-B51 and -B52 (Table 2.3), respectively, and interestingly, they verify this finding only for HLA-B52. Peptide binding increases the overall mobility and peptide-binding region floppiness of HLA-B51, while it results in a lower fluctuation and constrains the mobility of the peptide binding region in HLA-B52.

Figure 4.4 provides a better visualization for this observation. The black and red lines show the difference between the RMSF of the unbound and the peptide-bound -B51 and -B52, respectively, following the equation $\Delta\text{RMSF}=\text{RMSF}_{\text{HLA}-\text{B}}(\text{unbound})-\text{RMSF}_{\text{HLA}-\text{B}}(\text{bound})$. The red line fluctuating above zero indicates that the RMSF of the unbound -B52 is higher than the bound -B52 nearly for all residues. The black

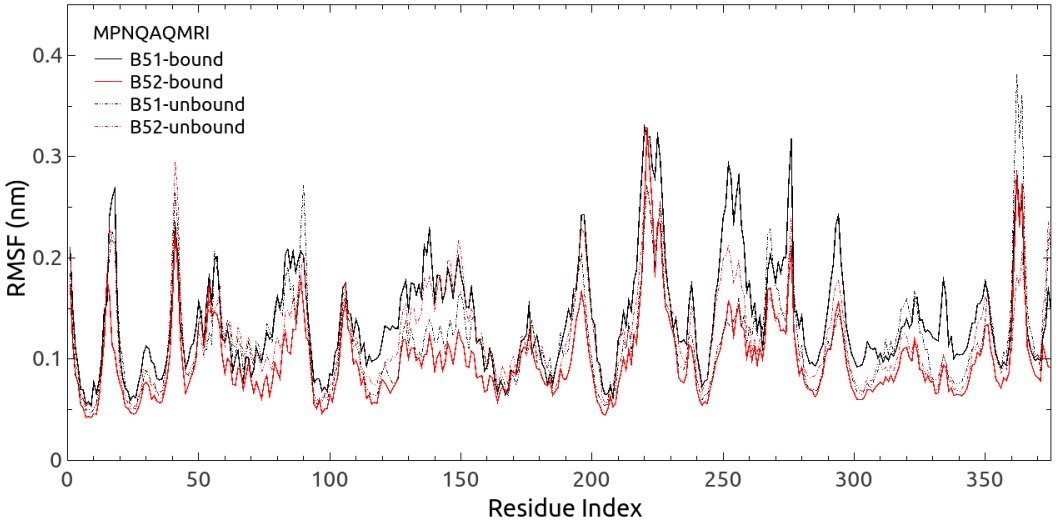
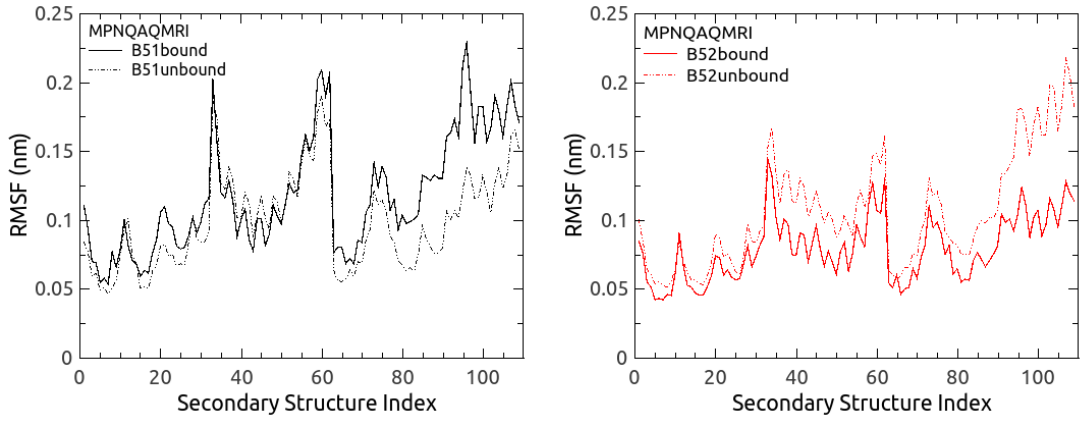


Figure 4.2: RMSF plot for HLA-B51 and -B52 unbound and bound to the 9-mer peptide MPNQAQMRI. The solid black and red lines belong to the peptide-bound -B51 and -B52, and dashed black and red lines belong to the unbound -B51 and -B52 respectively. It can be seen that peptide-bound HLA-B51 fluctuates the most and peptide-bound HLA-B52 fluctuates the least.

line, on the other hand, fluctuates below zero, and indicates that the RMSF of the bound -B51 is higher than the unbound -B51. RMSF difference plots for other peptide sequences are provided in the Appendix. Even though the difference is not as striking as this for all peptides, overall it can be concluded that peptide binding destabilizes the three-dimensional structure of HLA-B51, while it has a stabilizing effect on -B52.

To gain insight into the differences in the mobilities of peptides when bound to HLA-B51 and -B52, we compared the RMSF of peptides. Figure 4.5 shows that for a set of peptides, the vibrational fluctuations of the peptide are larger when bound to HLA-B51. As previously mentioned in this Chapter, HLA-B51 binds lower affinity



(a) (b)

Figure 4.3: (a)RMSF plot for the secondary structure elements of the peptide binding domain of HLA-B51 (residues indicated in Table 2.3) Peptide binding increases the peptide-binding region floppiness of HLA-B51 for this peptide.(b)RMSF plot for the secondary structure elements of the peptide binding domain of HLA-B52 (residues indicated in Table 2.3) Peptide binding constrains the mobility of the peptide binding region in HLA-B52 for this peptide.

peptides, and this set of peptides indeed seem to be more loosely bound and more mobile in the binding pocket of HLA-B51 compared to HLA-B52.

A second set of peptides exhibit a different behavior; although one might say that overall they are more mobile when bound to HLA-B51 as in the first set of peptides, not all their residues actually have higher mobilities individually. Figure 4.6 shows that for this set of peptides, the N-terminal of the peptide is more mobile when bound to HLA-B52, while the C-terminal is more mobile when bound to HLA-B51. As mentioned in the previous Chapters, HLA-B51 and -B52 differ in their B pockets. While the B pocket of HLA-B51 is moderately blocked, that of HLA-B52 is more spacious. The steric hindrances in the B pocket of HLA-B51 may result in a lower

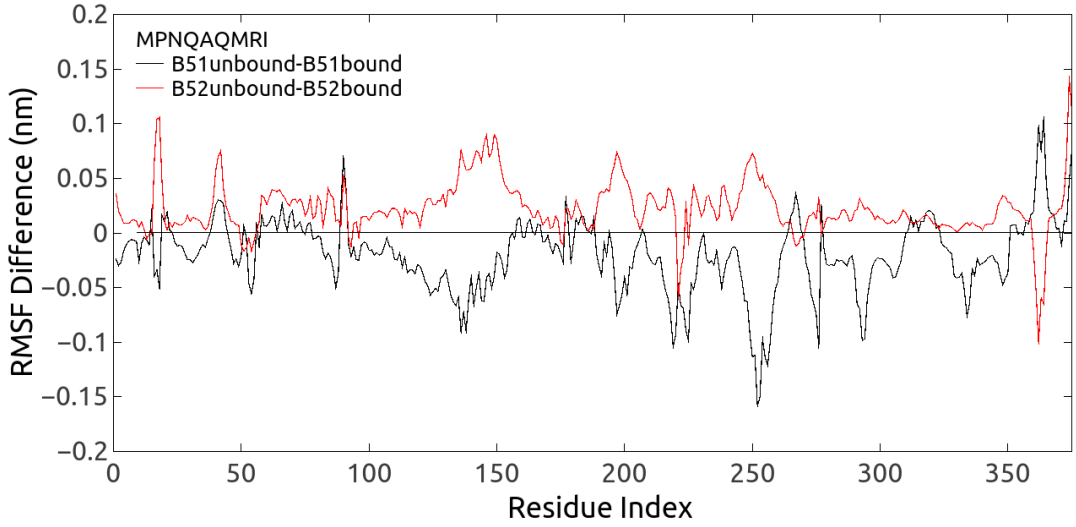


Figure 4.4: The black and red lines show the RMSF difference, following the equation ΔRMSF=RMSF$_{\mathrm{HLA-B}}$(unbound) − RMSF$_{\mathrm{HLA-B}}$(bound). The red line fluctuating above zero indicates that the RMSF of the unbound -B52 is higher than the bound -B52 nearly for all residues. The black line, however, fluctuates below zero, and indicates that the RMSF of the bound -B51 is higher than the unbound -B51.

mobility in the N-terminal, specifically in the second residue (P2) of several peptides.

To obtain a more collective, peptide-independent idea about the mobilities of the peptides when bound to HLA-B51 and -B52, we calculated a defined RMSF difference for each peptide residue, averaged over 16 peptides (Eq. 4.3) Since not all peptides are 9-mers, the RMSF difference of the 9th residue was averaged over 11 peptides. A positive RMSF difference for the ith residue means that the fluctuation of that residue is higher when bound to HLA-B51. As seen in Figure 4.7, all 9 residues have a positive RMSF difference in average. While the difference is lower in the C-terminal of the peptide, a significant difference can be seen in the N-terminal. This result indicates

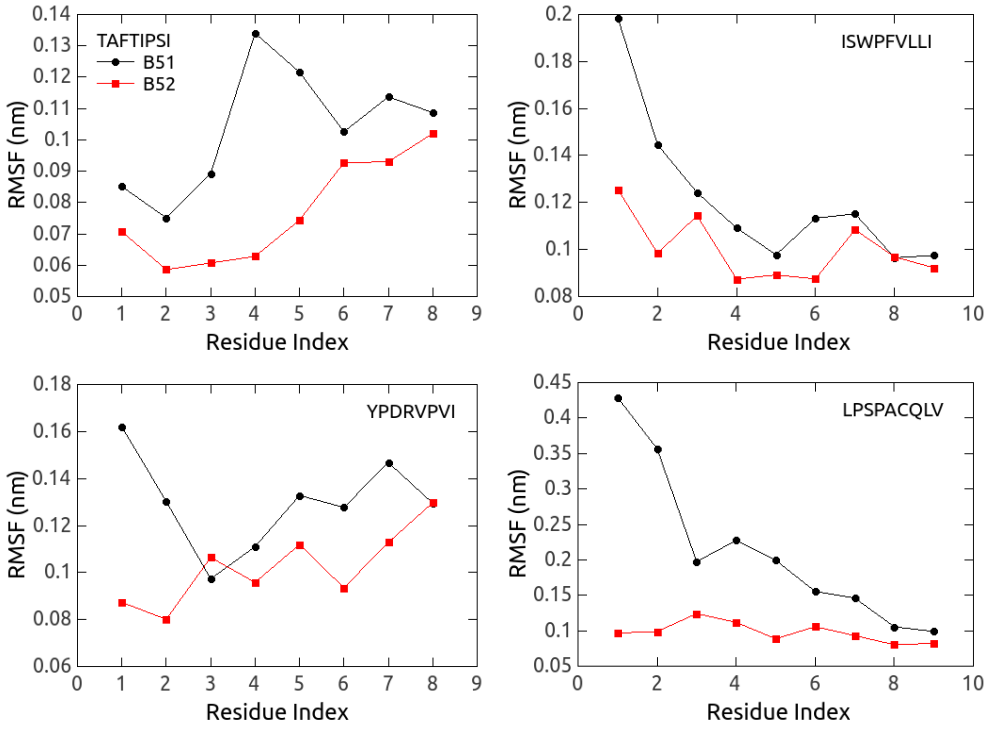


Figure 4.5: RMSF values for four peptides show that peptides have higher fluctuation when bound to HLA-B51. Black lines with circles indicate the RMSF values of the peptide bound to HLA-B51, whereas red lines with squares indicate the RMSF values of the peptide bound to HLA-B52. Peptide sequences are indicated on each panel.

that the affinity of peptides to the HLA-B51 molecule is weaker.

$$\Delta\mathrm{RMSF}(i) = \frac{\sum_{k=1}^{16} \dfrac{\mathrm{RMSF}_{B51}(i) - \mathrm{RMSF}_{B52}(i)}{\mathrm{RMSF}_{B52}(i)}}{16} \tag{4.3}$$

Since both peptide and β2m are crucial components for the stability of HLA-B51, we determined that peptide binding should affect the association of the heavy chain with β2m. The distance between the heavy chain and the β2m chain can be regarded

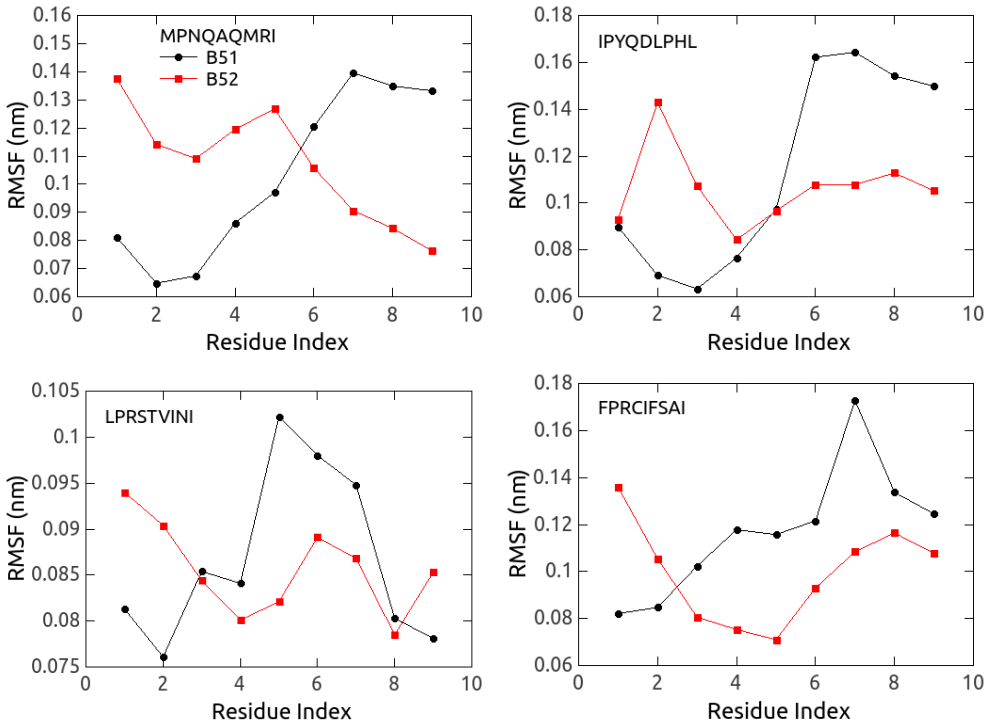


Figure 4.6: RMSF values for four peptides. Black lines with circles indicate the RMSF values of the peptide bound to HLA-B51, whereas red lines with squares indicate the RMSF values of the peptide bound to HLA-B52. Peptide sequences are indicated on each panel. For this second set of peptides, the N-terminal of the peptide is more mobile when bound to HLA-B52, while the C-terminal is more mobile when bound to HLA-B51.

as a measure of the overall stability of the protein, because if the heavy chain starts dissociating and retains a more open conformation, it will start disengaging from β2m, which will increase the distance between the two. Figures 4.9 to 4.13 show the normalized histograms of the distance between the two chains for 16 different peptides. Residue 230 on the heavy chain and residue 8 on the β2m chain were

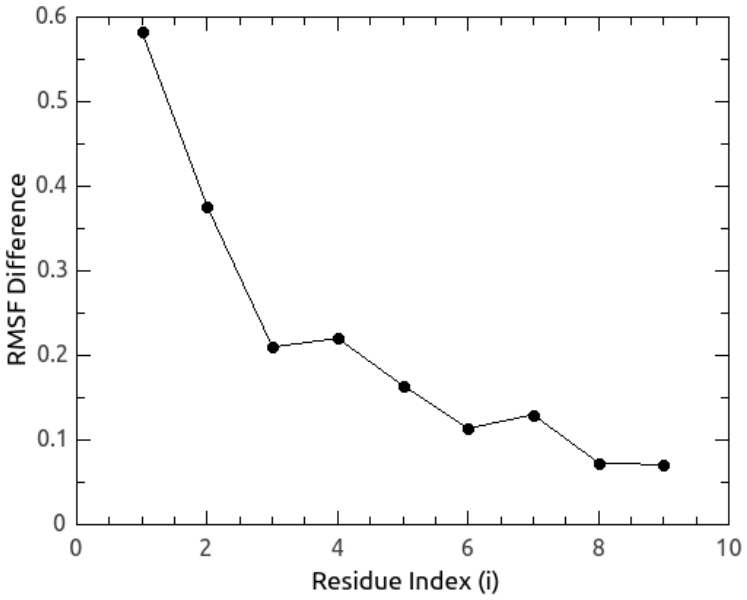


Figure 4.7: RMSF difference averaged over 16 peptides. A positive RMSF difference is seen for all peptide residues, which means that all 9 residues exhibit higher fluctuation in average when bound to HLA-B51.

chosen for the distance calculations (Figure 4.8). For all of these figures, the plots on the left-side show the distance distribution of HLA-B51, for the peptide sequence shown on the right of each panel, while the plots on the right-side show the distance distribution of HLA-B52 for the same peptide sequence. Black lines with circles show the distance distribution for the peptide-bound states of HLA-B51 and -B52, while red lines with squares show the distance distribution for the unbound states. A general trend observed in these plots is that peptide binding causes a decrease in the distance and results in a shift to a more stable conformation in HLA-B52, while there is no significant shift in -B51. Peptide binding makes HLA-B52 more coherent, which is the expected case, while it does not have such an effect on -B51. For peptides shown

in Figure 4.13, the distance distribution of HLA-B52 is bimodal. Here, a peak exists in the area where we normally would expect it to appear in case of a shift to a more stable conformation (12-14 Å). A second peak coincides with the peak observed for the unbound state of HLA-B52 (14-16 Å). Thus, this set of peptides are not fully able to result in a shift to a more stable conformation for HLA-B52.

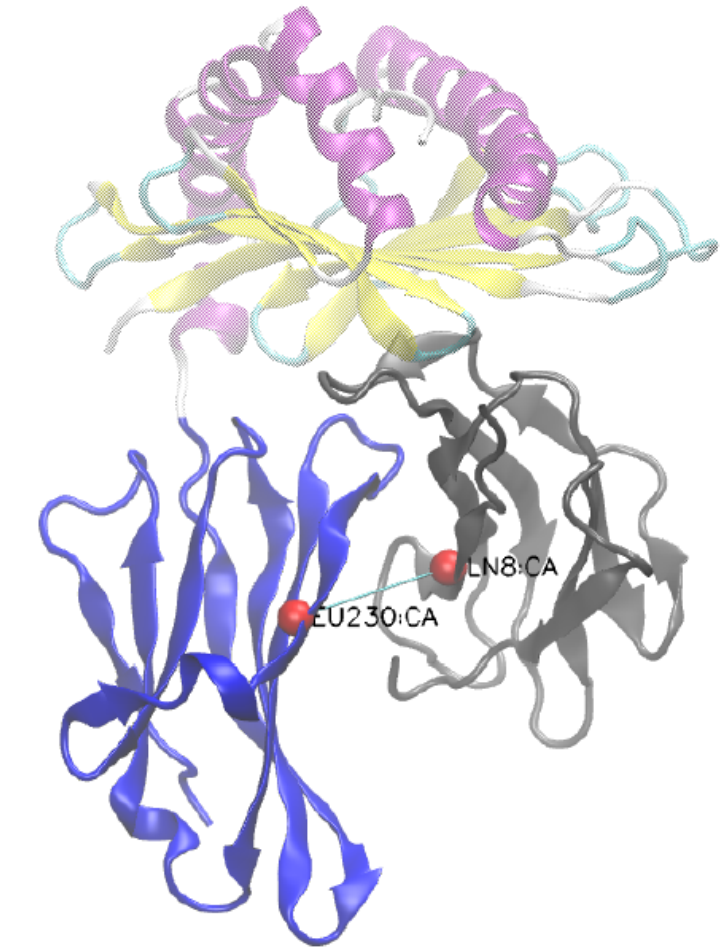


Figure 4.8: The distance between the heavy chain and the β2m chain can be regarded as a measure of the overall stability of the protein, because if the heavy chain starts dissociating and retains a more open conformation, it will start disengaging from β2m, which will increase the distance between the two. Residue 230 on the heavy chain and residue 8 on the β2m chain were chosen for the distance calculations.

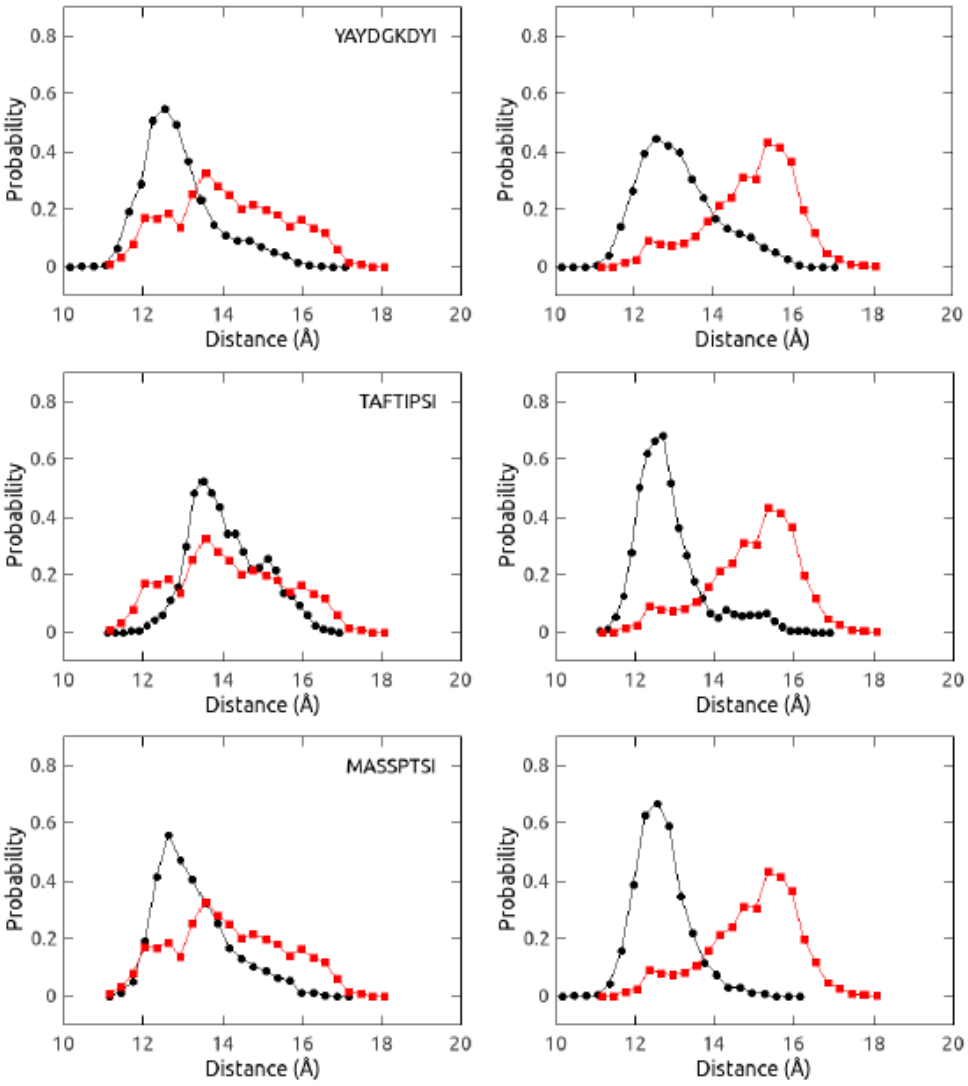


Figure 4.9: The normalized histograms of the distance between the heavy chain and the β2m chain for the three peptides: YAYDGKDYI, TAFTIPSI, MASSPTSI. Left-side plots show the distance distribution of HLA-B51 for the peptide sequence shown on the right of the panel, while the right-side plots show the distance distribution of HLA-B52 for the same peptide sequence. Black lines with circles show the distance distribution for the peptide-bound states of HLA-B51 and -B52, while red lines with squares show the distance distribution for the unbound states.

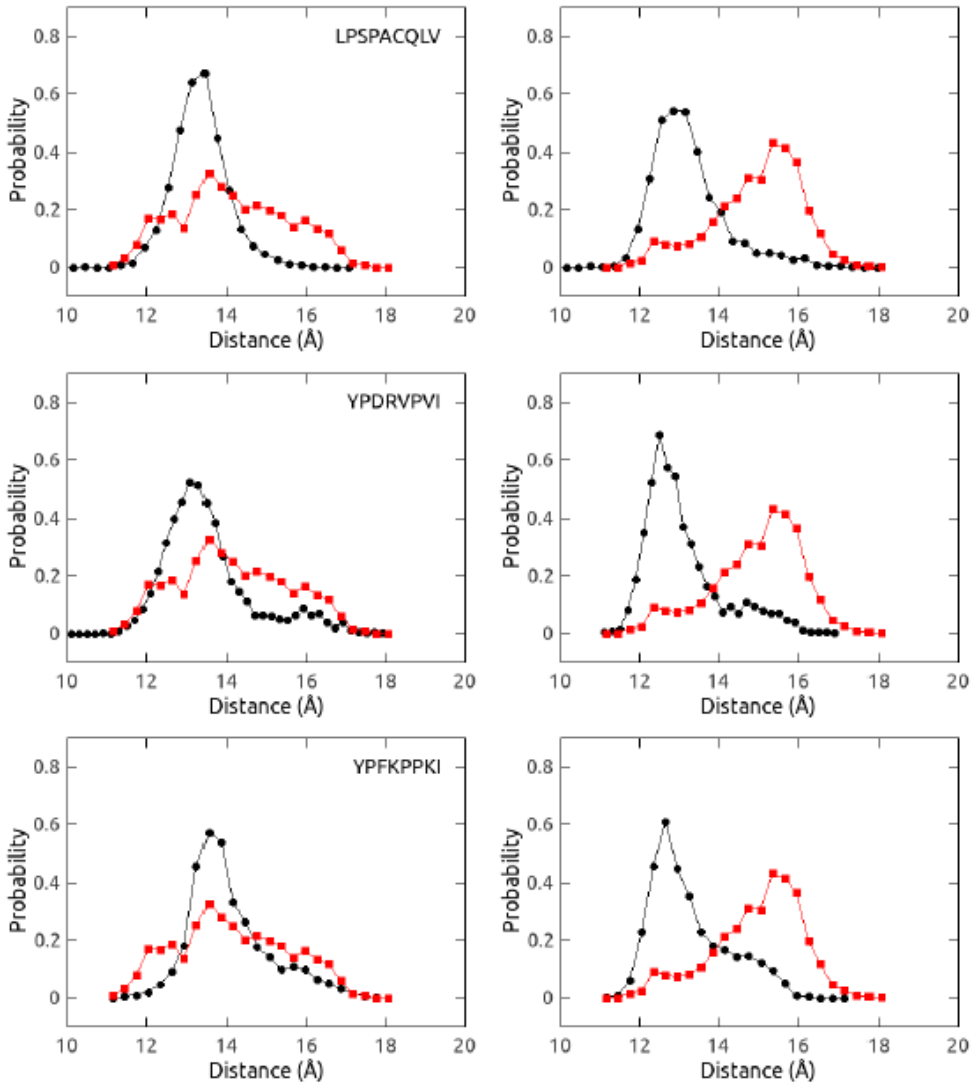


Figure 4.10: The normalized histograms of the distance between the heavy chain and the β2m chain for the three peptides: LPSPACQLV, YPDRVPVI, YPFKPPKI. Left-side plots show the distance distribution of HLA-B51 for the peptide sequence shown on the right of the panel, while the right-side plots show the distance distribution of HLA-B52 for the same peptide sequence. Black lines with circles show the distance distribution for the peptide-bound states of HLA-B51 and -B52, while red lines with squares show the distance distribution for the unbound states.

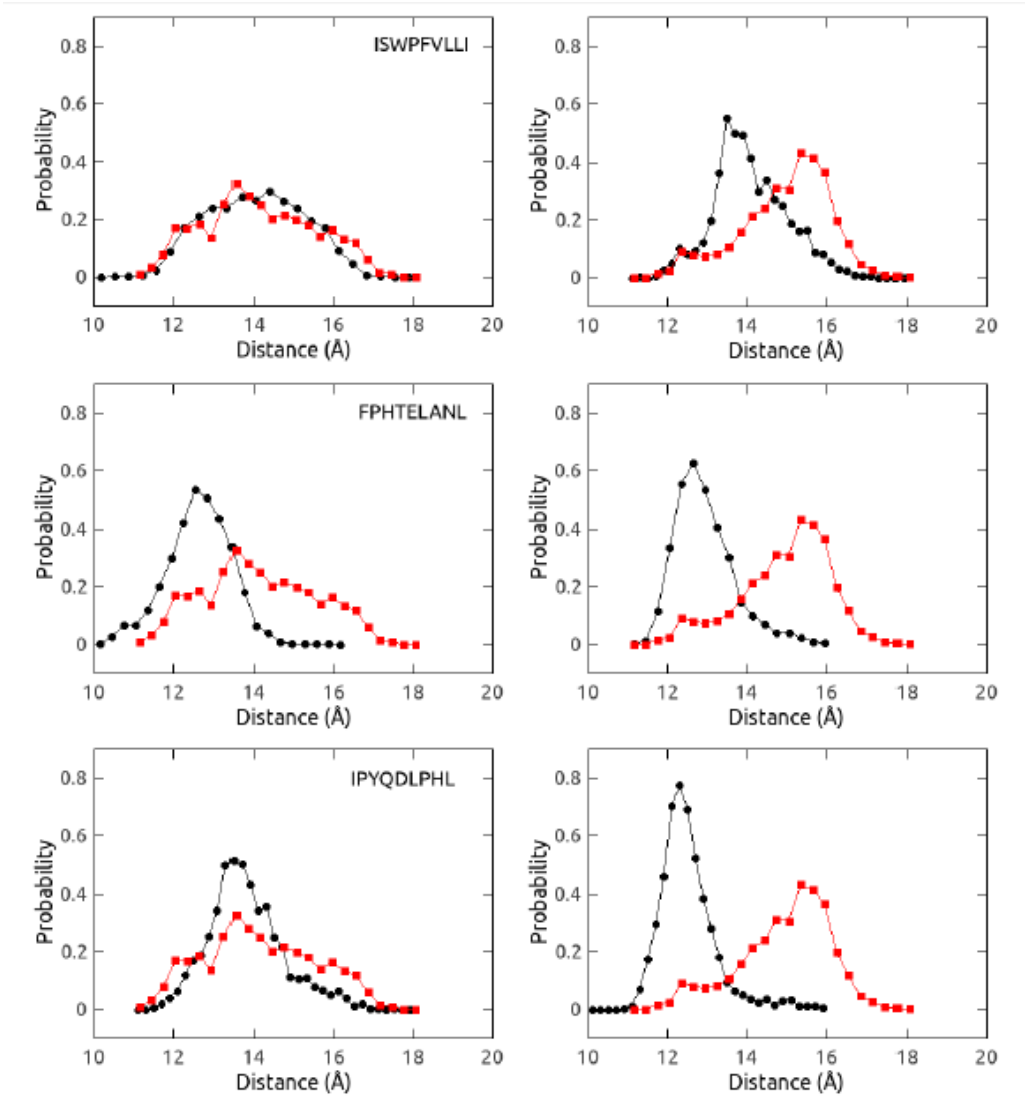


Figure 4.11: The normalized histograms of the distance between the heavy chain and the β2m chain for the three peptides: ISWPFVLLI, FPHTELANL, IPYQDLPHL. Left-side plots show the distance distribution of HLA-B51 for the peptide sequence shown on the right of the panel, while the right-side plots show the distance distribution of HLA-B52 for the same peptide sequence. Black lines with circles show the distance distribution for the peptide-bound states of HLA-B51 and -B52, while red lines with squares show the distance distribution for the unbound states.

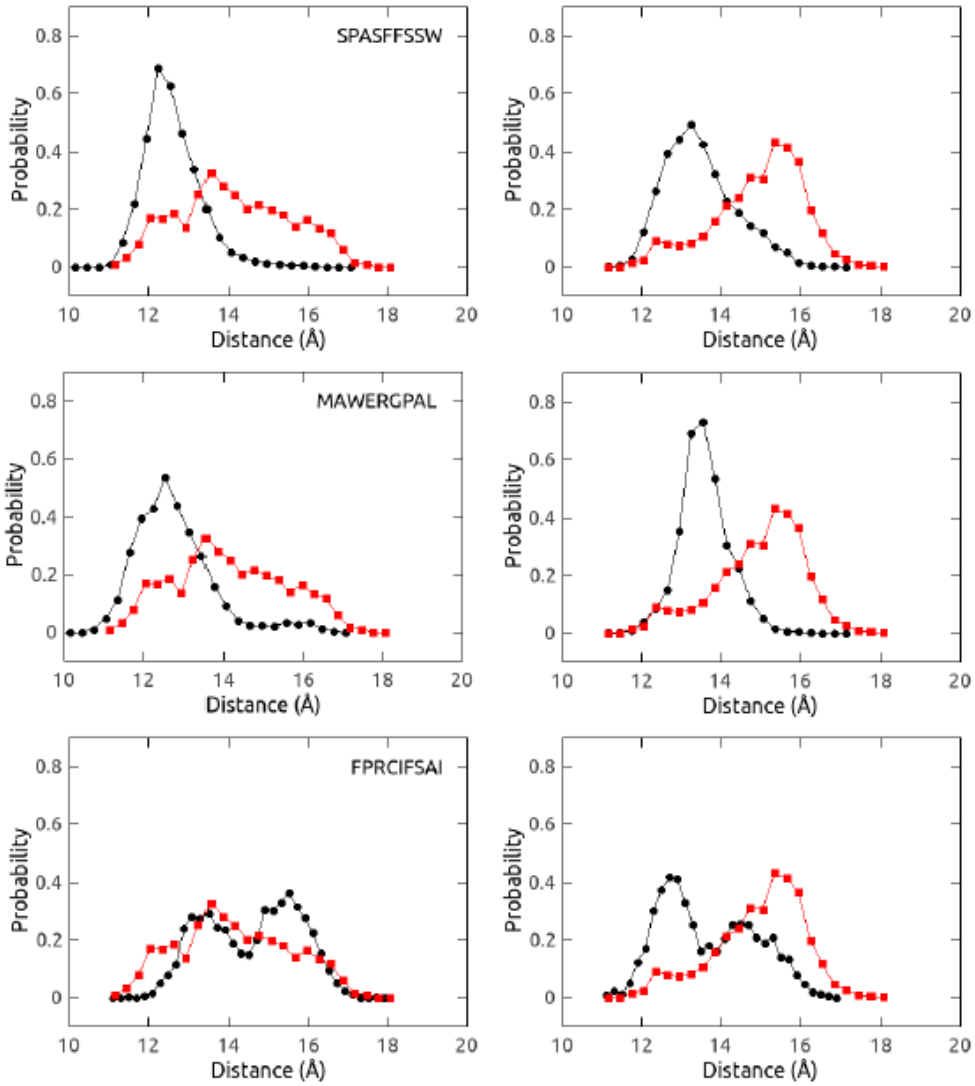


Figure 4.12: The normalized histograms of the distance between the heavy chain and the β2m chain for the three peptides: SPASFFSSW, MAWERGPAL, FPRCIFSAI. Left-side plots show the distance distribution of HLA-B51 for the peptide sequence shown on the right of the panel, while the right-side plots show the distance distribution of HLA-B52 for the same peptide sequence. Black lines with circles show the distance distribution for the peptide-bound states of HLA-B51 and -B52, while red lines with squares show the distance distribution for the unbound states.

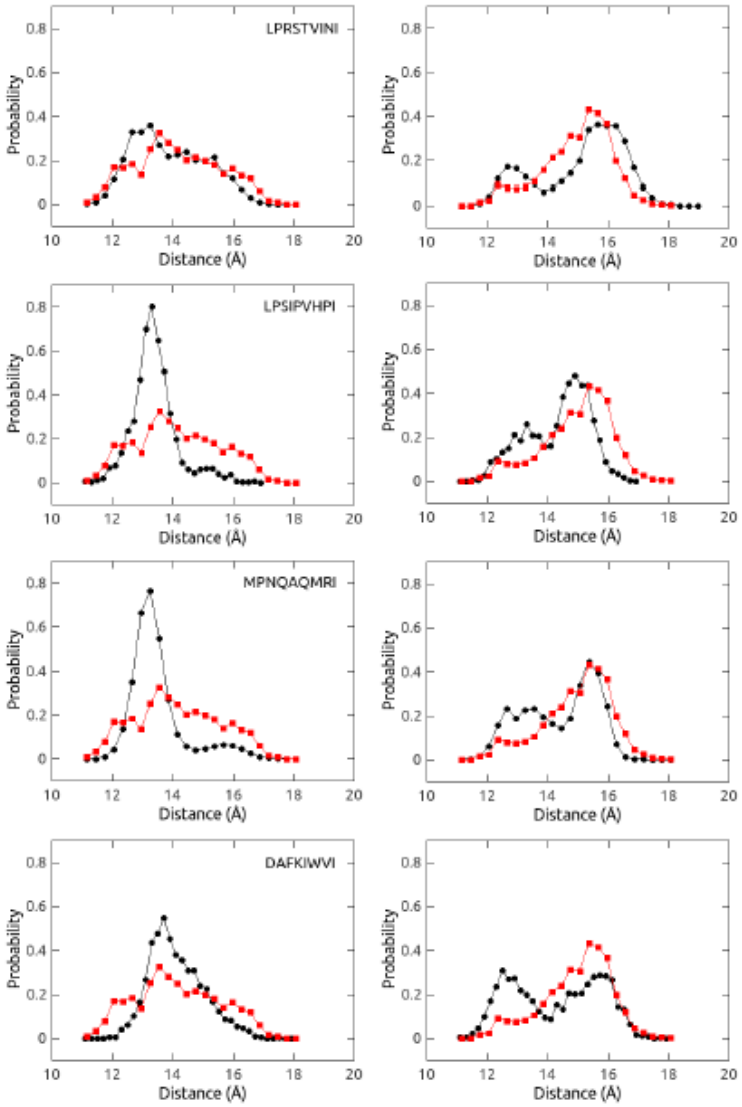


Figure 4.13: The normalized histograms of the distance between the heavy chain and the β2m chain for the four peptides: LPRSTVINI, LPSPVHPI, MPNQAQMRI, DAFKIWVI. Left-side plots show the distance distribution of HLA-B51 for the peptide sequence shown on the right of the panel, while the right-side plots show the distance distribution of HLA-B52 for the same peptide sequence. Black lines with circles show the distance distribution for the peptide-bound states of HLA-B51 and -B52, while red lines with squares show the distance distribution for the unbound states.

Chapter 5

CONCLUSION

In this thesis, we have performed comparative molecular dynamics simulations of HLA-B51 and HLA-B52. The aim of this study was to understand how the two residue difference between the two proteins affect their structural and dynamic properties such that this difference makes HLA-B51 the susceptibility allele for Behçet's Disease. Even though the two residue mutations are found on the peptide-binding pocket, we have observed that they have important global effects.

Through the analysis of MD simulations, we have identified differences in how peptide binding affects the two proteins. RMSF calculations revealed that while peptide binding increases the fluctuation and decreases the stability of HLA-B51, it decreases the fluctuation of HLA-B52 and makes it more coherent and stable. Distance distribution calculations showed that upon peptide binding, the distance between the heavy chain and the β2m chain decreases and leads to a shift to a more stable conformation in HLA-B52. However, it does not induce such a shift in HLA-B51. Therefore, upon peptide binding, HLA-B51 both becomes unstable with a higher RMSF, and its conformation does not shift to a more coherent conformation. Additionally, RMSF calculations for bound peptides reveal that the fluctuation of all 9 residues of the peptide are higher when bound to HLA-B51, which implies that peptides are more loosely bound to the binding pocket of HLA-B51.

These results have important implications for the pathogenic role of HLA-B51 in Behçet's Disease. HLA-B51 binds peptides with lower affinity and this may result in slow assembly and misfolding of HLA-B51. Peptide binding does not make HLA-B51

more stable and coherent, and this may induce a proinflammatory ER stress response in the cell through unfolded protein response. According to the results obtained in this work, the inflammation seems to be antigen-independent and directly linked to the structural properties of HLA-B51.

REFERENCES

[1] A Sette and J Sidney. Nine major HLA class I supertypes account for the vast preponderance of HLA-A and -B polymorphism. *Immunogenetics*, 50(3-4):201–12, November 1999.

[2] M A Saper, P J Bjorkman, and D C Wiley. Refined structure of the human histocompatibility antigen HLA-A2 at 2.6 Åresolution. *Journal of molecular biology*, 219:277–319, 1991.

[3] J W Lewis, A Neisig, J Neefjes, and T Elliott. Point mutations in the alpha 2 domain of HLA-A2.1 define a functionally relevant interaction with TAP. *Current biology : CB*, 6(7):873–83, July 1996.

[4] W K Suh, M a Derby, M F Cohen-Doyle, G J Schoenhals, K Früh, J a Berzofsky, and D B Williams. Interaction of murine MHC class I molecules with tapasin and TAP enhances peptide loading and involves the heavy chain alpha3 domain. *Journal of immunology (Baltimore, Md. : 1950)*, 162(3):1530–40, February 1999.

[5] N Mizuki, M Ota, K Yabuki, Y Katsuyama, H Ando, G D Palimeris, E Kaklamani, M Accorinti, P Pivetti-Pezzi, S Ohno, and H Inoko. Localization of the pathogenic gene of Behçet's disease by microsatellite analysis of three different populations. *Investigative ophthalmology & visual science*, 41(12):3702–8, November 2000.

[6] DH Verity, JE Marr, and S Ohno. Behçet's disease, the Silk Road and HLAB51: historical and geographical perspectives. *Tissue ...*, 54(3):213–220, 1999.

[7] International Study Group for Behçet's Disease. Criteria for diagnosis of Behçet's disease. *Lancet*, 335(8697):1078–80, May 1990.

[8] F Davatchi, K T Calamia, J E Crook, M Schirmer, A Altenburg, E Arromdee, M Baltaci, M Bastos, S Benamour, I Ben Ghorbel, and A Boyvat. The International Criteria for Behçet's Disease (ICBD): a collaborative study of 27 countries on the sensitivity and specificity of the new criteria. *Journal of the European Academy of Dermatology and Venereology : JEADV*, 28(3):338–47, March 2014.

[9] Sung Bin Cho, Suhyun Cho, and Dongsik Bang. New insights in the clinical understanding of Behçet's disease. *Yonsei medical journal*, 53(1):35–42, January 2012.

[10] Gülsevim Azizlerli, Afet Akda Köse, Rfkiye Sarca, Ahmet Gül, lknur Tual Tutkun, Mustafa Kulaç, Recep Tunç, Meri Urgancolu, and Rian Diçi. Prevalence of Behçet's disease in Istanbul, Turkey. *International Journal of Dermatology*, 42(10):803–806, October 2003.

[11] S Yurdakul, I Günaydin, Y Tüzün, N Tankurt, H Pazarli, Y Ozyazgan, and H Yazici. The prevalence of Behçet's syndrome in a rural area in northern Turkey. *The Journal of rheumatology*, 15(5):820–2, January 1988.

[12] A Idil, A Gurler, A Boyvat, D Caliskan, O Ozdemir, A Isik, A Tuncbilek, P Kocyigit, and E Calikoglu. The prevalence of Behçet's disease above the age of 10 years. The results of a pilot study conducted at the Park Primary Health Care Center in Ankara, Turkey. *Ophthalmic epidemiology*, 9(5):325–31, December 2002.

[13] Kenneth T Calamia, Floranne C Wilson, Murat Icen, Cynthia S Crowson, Sherine E Gabriel, and Hilal Maradit Kremers. Epidemiology and clinical characteristics of Behçet's disease in the US: a population-based study. *Arthritis and rheumatism*, 61(5):600–4, May 2009.

[14] Matteo Piga and Alessandro Mathieu. Genetic susceptibility to Behcet's disease: role of genes belonging to the MHC region. *Rheumatology (Oxford, England)*, 50(2):299–310, February 2011.

[15] Sam R Dalvi, Resit Yildirim, and Yusuf Yazici. Behcets Syndrome. 72(17):2223–2241, 2012.

[16] S Ohno, M Ohguchi, S Hirose, H Matsuda, A Wakisaka, and M Aizawa. Close association of HLA-Bw51 with Behçet's disease. *Archives of ophthalmology*, 100(9):1455–8, September 1982.

[17] Ahmet Gul. Genetics of Behçet's disease: lessons learned from genomewide association studies. *Current opinion in rheumatology*, 26(1):56–63, January 2014.

[18] Elaine F Remmers, Fulya Cosan, Yohei Kirino, Michael J Ombrello, Neslihan Abaci, Colleen Satorius, Julie M Le, Barbara Yang, Benjamin D Korman, Aris Cakiris, Oznur Aglar, Zeliha Emrence, Hulya Azakli, Duran Ustek, Ilknur Tugal-Tutkun, Gulsen Akman-Demir, Wei Chen, Christopher I Amos, Michael B Dizon, Afet Akdag Kose, Gulsevim Azizlerli, Burak Erer, Oliver J Brand, Virginia G Kaklamani, Phaedon Kaklamanis, Eldad Ben-Chetrit, Miles Stanford, Farida Fortune, Marwen Ghabra, William E R Ollier, Young-Hun Cho, Dongsik Bang, John O'Shea, Graham R Wallace, Massimo Gadina, Daniel L Kastner, and Ahmet Gül. Genome-wide association study identifies variants in the MHC class

I, IL10, and IL23R-IL12RB2 regions associated with Behçet's disease. *Nature genetics*, 42(8):698–702, August 2010.

[19] Yohei Kirino, George Bertsias, Yoshiaki Ishigatsubo, Nobuhisa Mizuki, Ilknur Tugal-Tutkun, Emire Seyahi, Yilmaz Ozyazgan, F Sevgi Sacli, Burak Erer, Hidetoshi Inoko, Zeliha Emrence, Atilla Cakar, Neslihan Abaci, Duran Ustek, Colleen Satorius, Atsuhisa Ueda, Mitsuhiro Takeno, Yoonhee Kim, Geryl M Wood, Michael J Ombrello, Akira Meguro, Ahmet Gül, Elaine F Remmers, and Daniel L Kastner. Genome-wide association analysis identifies new susceptibility loci for Behçet's disease and epistasis between HLA-B*51 and ERAP1. *Nature genetics*, 45(2):202–7, February 2013.

[20] Shengping Hou, Zhenglin Yang, Liping Du, Zhengxuan Jiang, Qinmeng Shu, Yuanyuan Chen, Fuzhen Li, Qingyun Zhou, Shigeaki Ohno, Rui Chen, Aize Kijlstra, James T Rosenbaum, and Peizeng Yang. Identification of a susceptibility locus in STAT4 for Behçet's disease in Han Chinese in a genome-wide association study. *Arthritis and rheumatism*, 64(12):4104–13, December 2012.

[21] Yun Jong Lee, Yukihiro Horie, Graham R Wallace, Yong Seok Choi, Ji Ah Park, Ji Yong Choi, Ran Song, Young-Mo Kang, Seong Wook Kang, Han Joo Baek, Nobuyoshi Kitaichi, Akira Meguro, Nobuhisa Mizuki, Kenichi Namba, Susumu Ishida, Jinhyun Kim, Edyta Niemczyk, Eun Young Lee, Yeong Wook Song, Shigeaki Ohno, and Eun Bong Lee. Genome-wide association study identifies GIMAP as a novel susceptibility locus for Behcet's disease. *Annals of the rheumatic diseases*, 72(9):1510–6, September 2013.

[22] K Yoshikawa, S Kotake, Y Sasamoto, S Ohno, and H Matsuda. [Close association

of Streptococcus sanguis and Behçet's disease]. *Nippon Ganka Gakkai zasshi*, 95(12):1261–7, December 1991.

[23] The Behcet's Disease Research Committee of Japan. Skin hypersensitivity to streptococcal antigens and the induction of systemic symptoms by the antigens in Behçet's disease–a multicenter study. The Behcet's Disease Research Committee of Japan. *The Journal of rheumatology*, 16(4):506–11, April 1989.

[24] R Eglin. Detection of RNA complementary to herpes-simplex virus in mononuclear cells from patients with Behcet's syndrome and recurrent oral ulcers. *The Lancet*, 320(8312):1356–1361, December 1982.

[25] Ediz Deniz, Ulker Guc, Nesimi Buyukbabani, and Ahmet Gul. HSP 60 expression in recurrent oral ulcerations of Behçet's disease. *Oral surgery, oral medicine, oral pathology, oral radiology, and endodontics*, 110(2):196–200, August 2010.

[26] Cem Evereklioglu. Current concepts in the etiology and treatment of Behçet disease. *Survey of ophthalmology*, 50(4):297–350, 2005.

[27] Marc Pineton de Chambrun, Bertrand Wechsler, Guillaume Geri, Patrice Cacoub, and David Saadoun. New insights into the pathogenesis of Behçet's disease. *Autoimmunity reviews*, 11(10):687–98, August 2012.

[28] James Robinson, Jason a Halliwell, Hamish McWilliam, Rodrigo Lopez, Peter Parham, and Steven G E Marsh. The IMGT/HLA database. *Nucleic acids research*, 41(Database issue):D1222–7, January 2013.

[29] Csaba Vizler, Nadege Bercovici, Anne Cornet, Christophe Cambouris, and Roland S. Liblau. Role of autoreactive CD8+ T cells in organ-specific autoimmune diseases: Insight from transgenic mouse models, 1999.

[30] Paulo N. Rocha, Troy J. Plumb, Steven D. Crowley, and Thomas M. Coffman. Effector mechanisms in transplant rejection, 2003.

[31] Eric Pamer and Peter Cresswell. MECHANISMS OF MHC CLASS I RESTRICTED ANTIGEN. 1998.

[32] Daniel C Chapman and David B Williams. ER quality control in the biogenesis of MHC class I molecules. *Seminars in cell & developmental biology*, 21(5):512–9, July 2010.

[33] T Sakaguchi, M Ibe, K Miwa, Y Kaneko, S Yokota, K Tanaka, and M Takiguchi. Binding of 8-mer to 11-mer peptides carrying the anchor residues to slow assembling HLA class I molecules (HLA-B*5101). *Immunogenetics*, 45(4):259–65, January 1997.

[34] D B Williams, B H Barber, R A Flavell, and H Allen. Role of beta 2-microglobulin in the intracellular transport and surface expression of murine class I histocompatibility molecules. *Journal of immunology (Baltimore, Md. : 1950)*, 142(8):2796–806, April 1989.

[35] E A Hughes, C Hammond, and P Cresswell. Misfolded major histocompatibility complex class I heavy chains are translocated into the cytoplasm and degraded by the proteasome. *Proceedings of the National Academy of Sciences of the United States of America*, 94(5):1896–901, March 1997.

[36] C Hammond and A Helenius. Quality control in the secretory pathway. *Current opinion in cell biology*, 7(4):523–9, August 1995.

[37] Nigil Haroon and Robert D Inman. Endoplasmic reticulum aminopeptidases:

Biology and pathogenic potential. *Nature reviews. Rheumatology*, 6(8):461–7, August 2010.

[38] K. M. Paulsson, M. Jevon, J. W. Wang, S. Li, and P. Wang. The Double Lysine Motif of Tapasin Is a Retrieval Signal for Retention of Unstable MHC Class I Molecules in the Endoplasmic Reticulum. *The Journal of Immunology*, 176(12):7482–7488, June 2006.

[39] V H Engelhard. Structure of peptides associated with class I and class II MHC molecules. *Annual review of immunology*, 12:181–207, 1994.

[40] Hana Mahmutefendić, Gordana Blagojević, Maja Ilić Tomaš, Natalia Kučić, and Pero Lučin. Segregation of open Major Histocompatibility Class I conformers at the plasma membrane and during endosomal trafficking reveals conformation-based sorting in the endosomal system. *The international journal of biochemistry & cell biology*, 43(4):504–15, April 2011.

[41] Yili Li, Yiyuan Yin, and Roy a Mariuzza. Structural and biophysical insights into the role of CD4 and CD8 in T cell activation. *Frontiers in immunology*, 4(July):206, January 2013.

[42] Christine M Oslowski and Fumihiko Urano. Measuring ER stress and the unfolded protein response using mammalian tissue culture system. *Methods in enzymology*, 490:71–92, January 2011.

[43] David Ron and Peter Walter. Signal integration in the endoplasmic reticulum unfolded protein response. *Nature reviews. Molecular cell biology*, 8(7):519–29, July 2007.

[44] Y Kozutsumi, M Segal, K Normington, M J Gething, and J Sambrook. The presence of malfolded proteins in the endoplasmic reticulum signals the induction of glucose-regulated proteins. *Nature*, 332(6163):462–4, March 1988.

[45] Heather P Harding and David Ron. Endoplasmic reticulum stress and the development of diabetes: a review. *Diabetes*, 51 Suppl 3:S455–61, December 2002.

[46] Wensheng Lin, Heather P Harding, David Ron, and Brian Popko. Endoplasmic reticulum stress modulates the response of myelinating oligodendrocytes to the immune cytokine interferon-gamma. *The Journal of cell biology*, 169(4):603–12, May 2005.

[47] Wulf Paschen and Thorsten Mengesdorf. Endoplasmic reticulum stress response and neurodegeneration. *Cell calcium*, 38(3-4):409–15, January 2005.

[48] Hiderou Yoshida. ER stress and diseases. *The FEBS journal*, 274(3):630–58, February 2007.

[49] Matthew J Turner, Monica L Delay, Shuzhen Bai, Erin Klenk, and Robert A Colbert. HLA-B27 up-regulation causes accumulation of misfolded heavy chains and correlates with the magnitude of the unfolded protein response in transgenic rats: Implications for the pathogenesis of spondylarthritis-like disease. *Arthritis and rheumatism*, 56(1):215–23, January 2007.

[50] Robert A Colbert, Monica L DeLay, Erin I Klenk, and Gerlinde Layh-Schmitt. From HLA-B27 to spondyloarthritis: a journey through the ER. *Immunological reviews*, 233(1):181–202, January 2010.

[51] Joseph D Comber and Ramila Philip. MHC class I antigen presentation and

implications for developing a new generation of therapeutic vaccines. *Therapeutic advances in vaccines*, 2(3):77–89, May 2014.

[52] Martin Zacharias and Sebastian Springer. Conformational flexibility of the MHC class I alpha1-alpha2 domain in peptide bound and free states: a molecular dynamics simulation study. *Biophysical journal*, 87(4):2203–14, October 2004.

[53] J. Margulies, D.H., Natarajan, K., Rossjohn, J., and McCluskey. Major Histocompatibility Complex (MHC) Molecules: Structure, Function, and Genetics. In *Fundamental Immunology*, pages 570–613. 2008.

[54] N Arber, T Klein, Z Meiner, E Pras, and A Weinberger. Close association of HLA-B51 and B52 in Israeli patients with Behçet's syndrome. *Annals of the rheumatic diseases*, 50(6):351–3, June 1991.

[55] C Lemmel, H G Rammensee, and S Stevanovic. Peptide motif of HLA-B*5101 and the linkage to Behcets Disease. In *Immunology of Behcet's Disease*, page 127. 2003.

[56] Y Takemoto, T Naruse, K Namba, N Kitaichi, M Ota, Y Shindo, N Mizuki, A Gul, W Madanat, H Chams, F Davatchi, H Inoko, S Ohno, and A Kimura. Re-evaluation of heterogeneity in HLA-B*510101 associated with Behçet's disease. *Tissue antigens*, 72(4):347–53, October 2008.

[57] K Maenaka, T Maenaka, H Tomiyama, M Takiguchi, D I Stuart, and E Y Jones. Nonstandard peptide binding revealed by crystal structures of HLA-B*5101 complexed with HIV immunodominant epitopes. *Journal of immunology (Baltimore, Md. : 1950)*, 165(6):3260–7, September 2000.

[58] P J Bjorkman and P Parham. *Structure, function, and diversity of class I major histocompatibility complex molecules.*, volume 59. January 1990.

[59] A C M Young, S G Nathenson, and J C Sacchettini. Structural studies of class I major histocompatibility complex proteins: insights into antigen presentation. *FASEB*, 9(5):26–36, 1995.

[60] Ahmet Gul and Shigeaki Ohno. Genetics of Behcet's Disease. In *Behcet's Syndrome*, pages 265–276. Springer New York, 1 edition, 2010.

[61] Kirsten Falk, Olaf Rotzschke, Masafumi Takiguchi, Volker Gnau, Stefan Stevanovic, Gunther Jung, and Hans-georg Rammensee. Peptide motifs of HLA-B51 , -B52 and -B78 molecules, and implications for Behcet's disease. (2):223–228, 1995.

[62] Mingnan Chen and Marlene Bouvier. Analysis of interactions in a tapasin/class I complex provides a mechanism for peptide selection. *The EMBO journal*, 26(6):1681–90, March 2007.

[63] Mark Howarth, Anthony Williams, Anne B Tolstrup, and Tim Elliott. Tapasin enhances MHC class I peptide presentation according to peptide half-life. *Proceedings of the National Academy of Sciences of the United States of America*, 101(32):11737–42, August 2004.

[64] A Neisig, R Wubbolts, X Zang, C Melief, and J Neefjes. Allele-specific differences in the interaction of MHC class I molecules with transporters associated with antigen processing. *Journal of immunology (Baltimore, Md. : 1950)*, 156:3196–3206, 1996.

[65] Chen A. Peh, Scott R. Burrows, Megan Barnden, Rajiv Khanna, Peter Cresswell, Denis J. Moss, and James McCluskey. HLA-B27-restricted antigen presentation in the absence of tapasin reveals polymorphism in mechanisms of HLA class I peptide loading. *Immunity*, 8:531–542, 1998.

[66] H R Turnquist, H J Thomas, K R Prilliman, C T Lutz, W H Hildebrand, and J C Solheim. HLA-B polymorphism affects interactions with multiple endoplasmic reticulum proteins. *European journal of immunology*, 30(10):3021–8, October 2000.

[67] B. Park, S. Lee, E. Kim, and K. Ahn. A Single Polymorphic Residue Within the Peptide-Binding Cleft of MHC Class I Molecules Determines Spectrum of Tapasin Dependence. *The Journal of Immunology*, 170(2):961–968, January 2003.

[68] Florian Sieker, Tjerk P Straatsma, Sebastian Springer, and Martin Zacharias. Differential tapasin dependence of MHC class I molecules correlates with conformational changes upon peptide dissociation: a molecular dynamics simulation study. *Molecular immunology*, 45(14):3714–22, August 2008.

[69] A G Brooks, J C Boyington, and P D Sun. Natural killer cell recognition of HLA class I molecules. *Reviews in immunogenetics*, 2:433–448, 2000.

[70] G Saruhan-Direskeneli, F a Uyar, A Cefle, S C Onder, E Eksioglu-Demiralp, S Kamali, M Inanç, L Ocal, and A Gül. Expression of KIR and C-type lectin receptors in Behcet's disease. *Rheumatology (Oxford, England)*, 43(4):423–7, April 2004.

[71] C E Voorter, S van der Vlies, M Kik, and E M van den Berg-Loonen. Unexpected

Bw4 and Bw6 reactivity patterns in new alleles. *Tissue antigens*, 56(4):363–70, October 2000.

[72] Julian P Vivian, Renee C Duncan, Richard Berry, Geraldine M O'Connor, Hugh H Reid, Travis Beddoe, Stephanie Gras, Philippa M Saunders, Maya a Olshina, Jacqueline M L Widjaja, Christopher M Harpur, Jie Lin, Sebastien M Maloveste, David a Price, Bernard a P Lafont, Daniel W McVicar, Craig S Clements, Andrew G Brooks, and Jamie Rossjohn. Killer cell immunoglobulin-like receptor 3DL1-mediated recognition of human leukocyte antigen B. *Nature*, 479(7373):401–5, December 2011.

[73] Shunzhou Wan, Darren R Flower, and Peter V Coveney. Toward an atomistic understanding of the immune synapse: large-scale molecular dynamics simulation of a membrane-embedded TCR-pMHC-CD4 complex. *Molecular immunology*, 45(5):1221–30, March 2008.

[74] Shunzhou Wan, Peter Coveney, and Darren R Flower. Large-scale molecular dynamics simulations of HLA-A*0201 complexed with a tumor-specific antigenic peptide: can the alpha3 and beta2m domains be neglected? *Journal of computational chemistry*, 25(15):1803–13, December 2004.

[75] Ulrich Omasits, Bernhard Knapp, Martin Neumann, Othmar Steinhauser, Hannes Stockinger, Rene Kobler, and Wolfgang Schreiner. Analysis of key parameters for molecular dynamics of pMHC molecules. *Molecular Simulation*, 34(8):781–793, July 2008.

[76] David Van Der Spoel, Erik Lindahl, Berk Hess, Gerrit Groenhof, Alan E Mark, and Herman J C Berendsen. GROMACS: fast, flexible, and free. *Journal of computational chemistry*, 26(16):1701–18, December 2005.

[77] Nicolas Foloppe and Alexander D Mackerell. All-Atom Empirical Force Field for Nucleic Acids : I . Parameter Optimization Based on Small Molecule and Condensed Phase Macromolecular Target Data. 21(2):86–104, 2000.

[78] William L. Jorgensen, Jayaraman Chandrasekhar, Jeffry D. Madura, Roger W. Impey, and Michael L. Klein. Comparison of simple potential functions for simulating liquid water. *The Journal of Chemical Physics*, 79(2):926, 1983.

[79] Giovanni Bussi, Davide Donadio, and Michele Parrinello. Canonical sampling through velocity rescaling. *The Journal of chemical physics*, 126(1):014101, January 2007.

[80] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, a. DiNola, and J. R. Haak. Molecular dynamics with coupling to an external bath. *The Journal of Chemical Physics*, 81(8):3684, 1984.

[81] Berk Hess, Henk Bekker, Herman J. C. Berendsen, and Johannes G. E. M. Fraaije. LINCS: A linear constraint solver for molecular simulations. *Journal of Computational Chemistry*, 18(12):1463–1472, September 1997.

[82] Tom Darden, Darrin York, and Lee Pedersen. Particle mesh Ewald: An Nlog(N) method for Ewald sums in large systems. *The Journal of Chemical Physics*, 98(12):10089, 1993.

[83] Shigeaki Ohno, Koki Aoki, Seiji Sugiura, Eiichi Nakayama, Katsuaki Itakura, and Miki Aizawa. HL-A5 AND BEHCET'S DISEASE. *The Lancet*, 302(7842):1383–1384, December 1973.

[84] Yohei Kirino, Qing Zhou, Yoshiaki Ishigatsubo, Nobuhisa Mizuki, Ilknur Tugal-Tutkun, Emire Seyahi, Yilmaz Özyazgan, Serdal Ugurlu, Burak Erer, Neslihan

Abaci, Duran Ustek, Akira Meguro, Atsuhisa Ueda, Mitsuhiro Takeno, Hidetoshi Inoko, Michael J Ombrello, Colleen L Satorius, Baishali Maskeri, James C Mullikin, Hong-Wei Sun, Gustavo Gutierrez-Cruz, Yoonhee Kim, Alexander F Wilson, Daniel L Kastner, Ahmet Gül, and Elaine F Remmers. Targeted resequencing implicates the familial Mediterranean fever gene MEFV and the toll-like receptor 4 gene TLR4 in Behçet disease. *Proceedings of the National Academy of Sciences of the United States of America*, 110(20):8134–9, May 2013.

[85] A Gul, F A Uyar, M Inanç, L Ocal, J H Barrett, O Aral, M Koniçe, and G Saruhan-Direskeneli. A weak association of HLA-B*2702 with Behçet's disease. *Genes and immunity*, 3(6):368–72, September 2002.

[86] Michael J Ombrello, Yohei Kirino, Paul I W de Bakker, Ahmet Gül, Daniel L Kastner, and Elaine F Remmers. Behçet disease-associated MHC class I residues implicate antigen binding and regulation of cell-mediated cytotoxicity. *Proceedings of the National Academy of Sciences of the United States of America*, 111(24):8867–72, June 2014.

[87] Ahmet Gul and Shigeaki Ohno. HLA-B*51 and Behçet Disease. *Ocular immunology and inflammation*, 20(1):37–43, February 2012.

[88] Daniel Gebreselassie, Hans Spiegel, and Stanislav Vukmanovic. Sampling of major histocompatibility complex class I-associated peptidome suggests relatively looser global association of HLA-B*5101 with peptides. *Human immunology*, 67(11):894–906, November 2006.

[89] H Yasuoka, Y Yamaguchi, N Mizuki, T Nishida, Y Kawakami, and M Kuwana. Preferential activation of circulating CD8+ and gammadelta T cells in patients

with active Behçet's disease and HLA-B51. *Clinical and experimental rheumatology*, 26(4 Suppl 50):S59–63, 2008.

[90] A Hill, M Takiguchi, and A McMichael. Different rates of HLA class I molecule assembly which are determined by amino acid sequence in the alpha 2 domain. *Immunogenetics*, 37(2):95–101, January 1993.

[91] H G Ljunggren, N J Stam, C Ohlén, J J Neefjes, P Höglund, M T Heemels, J Bastin, T N Schumacher, A Townsend, and K Kärre. Empty MHC class I molecules come out in the cold. *Nature*, 346(6283):476–80, August 1990.

[92] M L Wei and P Cresswell. HLA-A2 molecules in an antigen-processing mutant cell containn signal sequence-derived peptides. *Nature*, 356:443–446, 1992.

[93] E Z Wolpert, M Petersson, B J Chambers, J K Sandberg, R Kiessling, H G Ljunggren, and K Kärre. Generation of CD8+ T cells specific for transporter associated with antigen processing deficient cells. *Proceedings of the National Academy of Sciences of the United States of America*, 94(21):11496–501, October 1997.

[94] M Bouvier and D C Wiley. Importance of peptide amino and carboxyl termini to the stability of MHC class I molecules. *Science (New York, N.Y.)*, 265:398–402, 1994.

[95] M Bouvier and D C Wiley. Structural characterization of a soluble and partially folded class I major histocompatibility heavy chain/beta 2m heterodimer. *Nature structural biology*, 5:377–384, 1998.

[96] M Zijlstra, M Bix, N E Simister, J M Loring, D H Raulet, and R Jaenisch.

Beta 2-microglobulin deficient mice lack CD4-8+ cytolytic T cells. *Nature*, 344(6268):742–6, April 1990.

[97] B. Koller, P Marrack, J. Kappler, and O Smithies. Normal development of mice deficient in beta 2M, MHC class I proteins, and CD8+ T cells. *Science*, 248(4960):1227–1230, June 1990.

[98] R. H. Seong. Rescue of Daudi cell HLA expression by transfection of the mouse beta 2- microglobulin gene. *Journal of Experimental Medicine*, 167(2):288–299, February 1988.

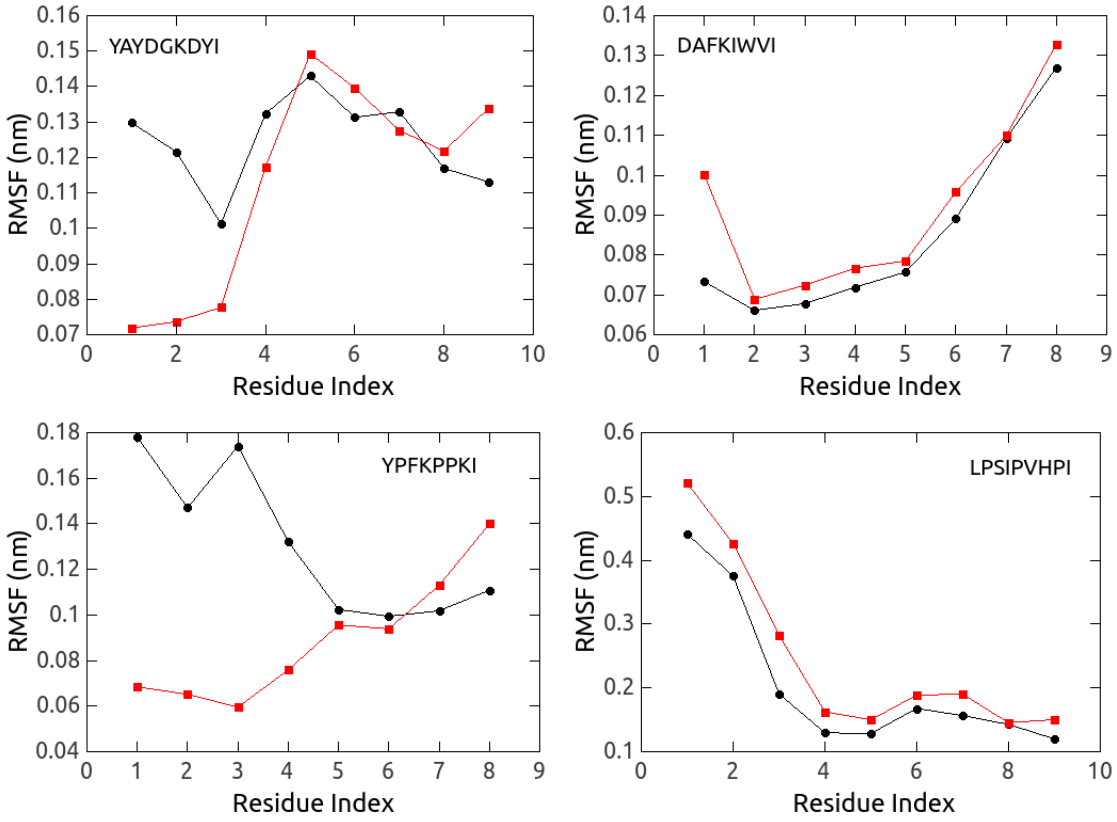
[99] Aline Martayan, Leonardo Sibilio, Elisa Tremante, Elisa Lo Monaco, Arend Mulder, Doriana Fruci, Agata Cova, Licia Rivoltini, and Patrizio Giacomini. Class I HLA folding and antigen presentation in beta 2-microglobulin-defective Daudi cells. *Journal of immunology (Baltimore, Md. : 1950)*, 182(6):3609–17, March 2009.

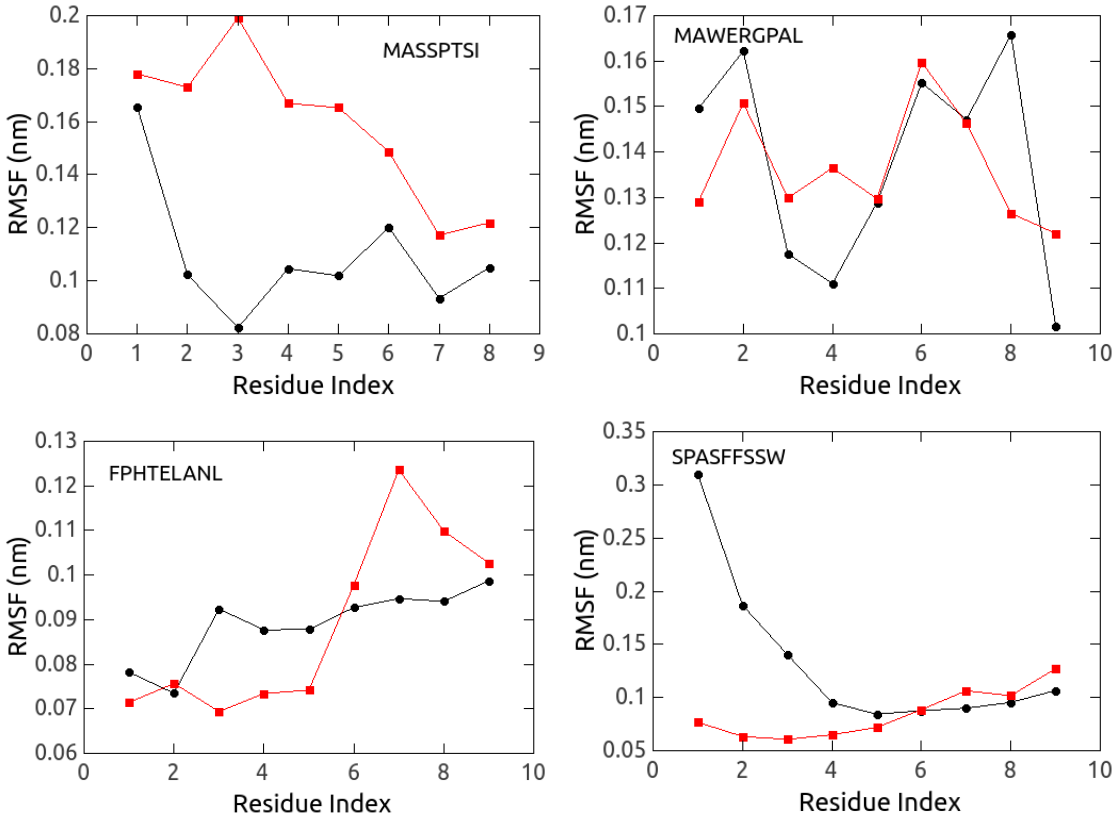
Appendices

Appendix 1

RMSF OF PEPTIDES

Black lines with circles indicate the RMSF values of the peptide bound to HLA-B51, whereas red lines with squares indicate the RMSF values of the peptide bound to HLA-B52. Peptide sequences are indicated on each panel.

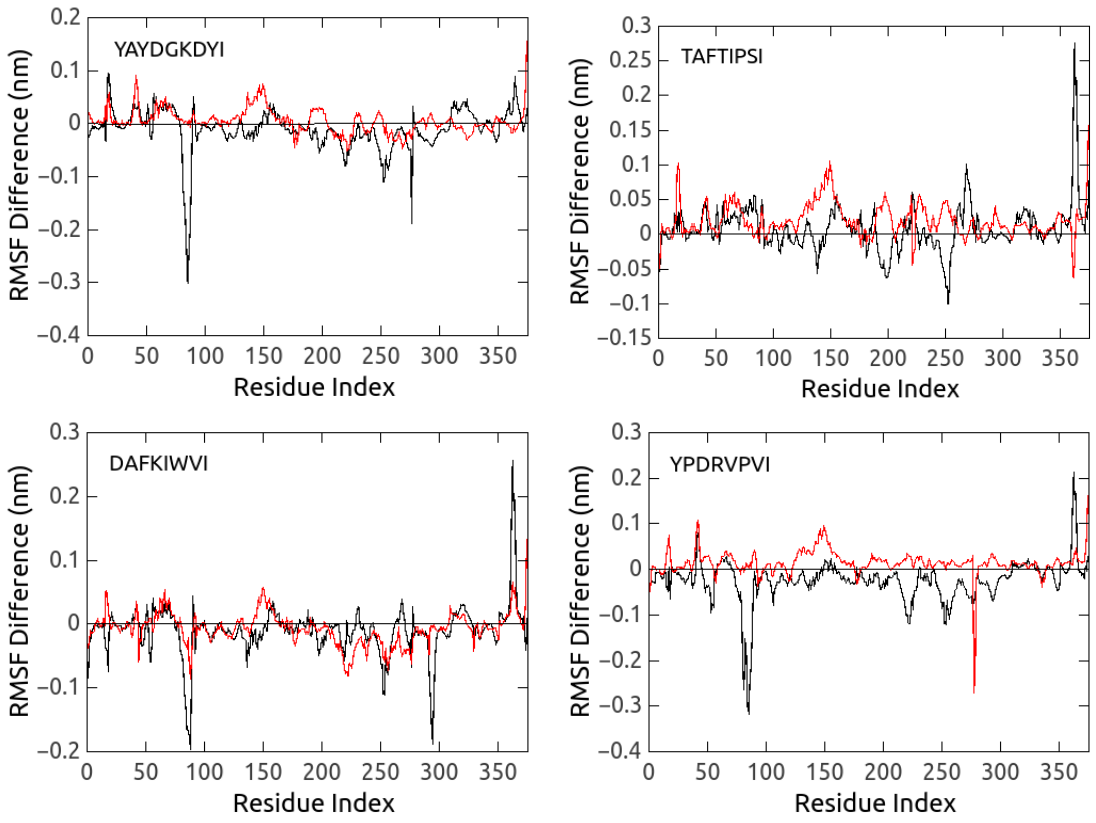


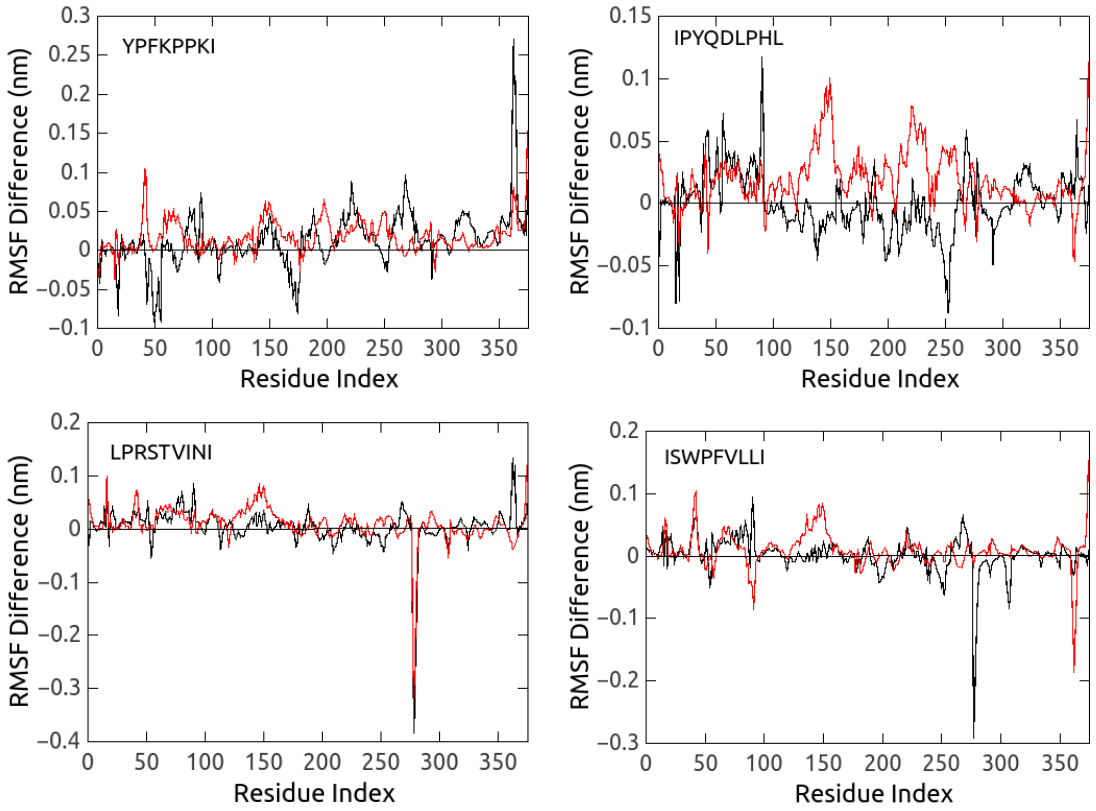

MASSPTSI
RMSF (nm)
Residue Index
MAWERGPAL
RMSF (nm)
Residue Index
FPHTELANL
RMSF (nm)
Residue Index
SPASFFSSW
RMSF (nm)
Residue Index

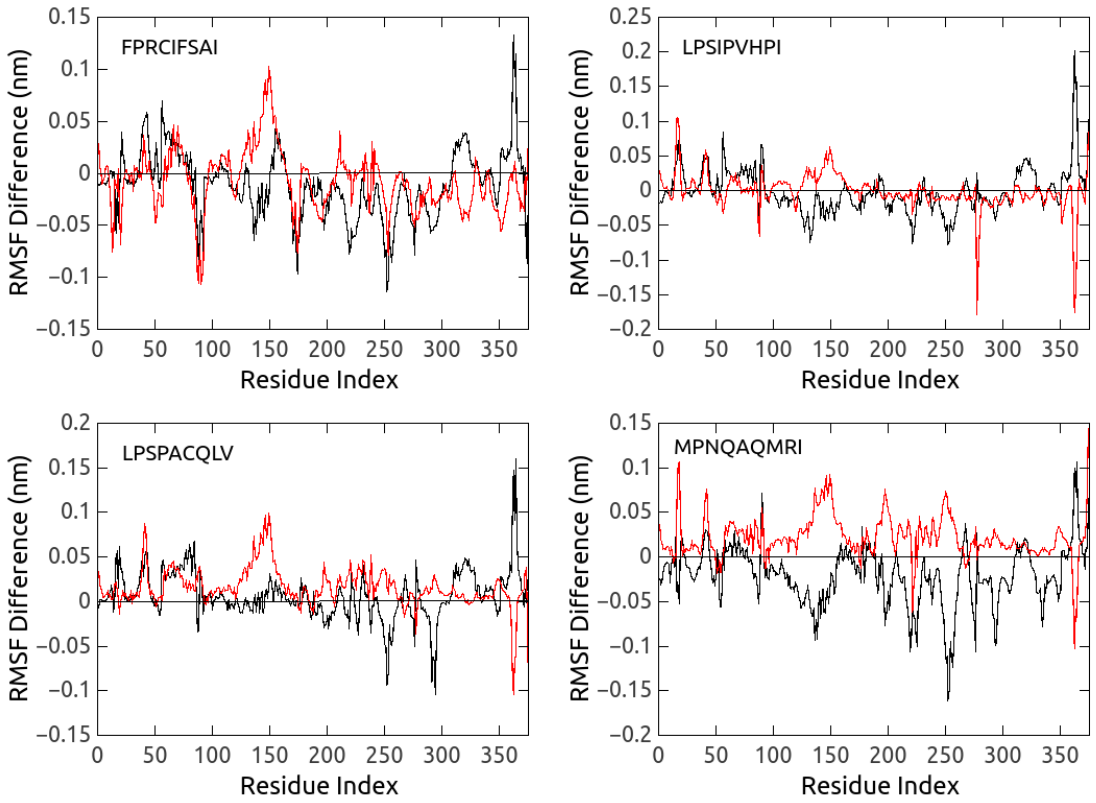
Appendix 2

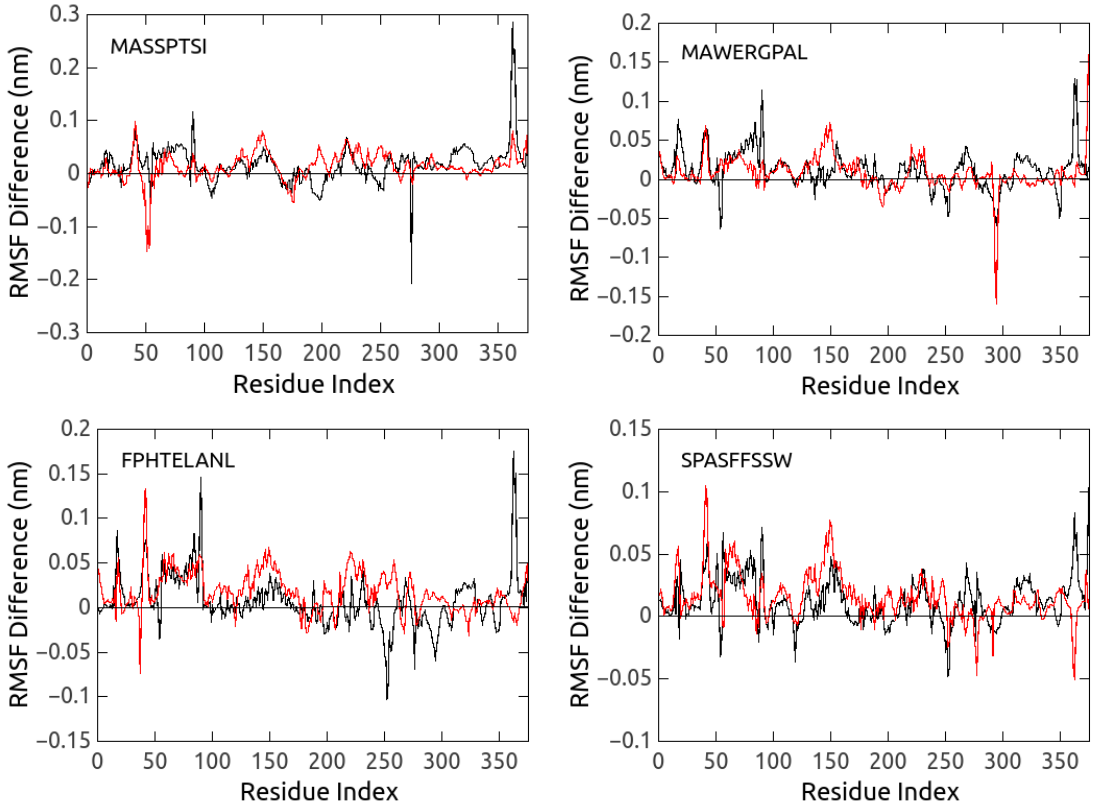
RMSF DIFFERENCES

The difference between the RMSF of the unbound and the peptide-bound -B51 and -B52, following the equation $\Delta\text{RMSF}=\text{RMSF}_{\text{HLA-B}}(\text{unbound})-\text{RMSF}_{\text{HLA-B}}(\text{bound})$. The black line indicates the RMSF difference for HLA-B51, where the red line indicates the RMSF difference of HLA-B52. Peptide sequences are indicated on each panel.




YPFKPPKI
IPYQDLPHL
LPRSTVINI
ISWPFVLLI
RMSF Difference (nm)
Residue Index


FPRCIFSAI
LPSIPVHPI
LPSPACQLV
MPNQAQMRI
RMSF Difference (nm)
Residue Index


MASSPTSI
MAWERGPAL
FPHTELANL
SPASFFSSW
RMSF Difference (nm)
Residue Index