

PAPER DETAILS

TITLE: HLA PROTEINLERİNİN GNM METODU İLE DİNAMİK KARAKTERİZASYONU

AUTHORS: Onur SERÇİNOĞLU, Pemra ÖZBEK

PAGES: 376-399

ORIGINAL PDF URL: <https://dergipark.org.tr/tr/download/article-file/587170>

HLA Proteinlerinin GNM Metodu ile Dinamik Karakterizasyonu

Onur SERÇİNOĞLU¹, Pemra ÖZBEK^{*2},

¹Marmara Üniversitesi, Fen Bilimleri Enstitüsü, Biyomühendislik Anabilim Dalı,
34722, İstanbul (ORCID: 0000-0003-1361-8160)

²Marmara Üniversitesi, Mühendislik Fakültesi, Biyomühendislik Bölümü, 34722,
İstanbul (ORCID: 0000-0002-3043-0015)

(Alınış / Received: 03.03.2017, Kabul / Accepted: 01.02.2018,
Online Yayınlanma / Published Online: 15.05.2018)

Anahtar Kelimeler
Protein dinamiği,
Gaussian ağ yapı
modeli,
HLA,
Hızlı modlar

Özet: Biyolojik sistemler, birbirleri ile sürekli bilgi alışverişinde bulunan ağıyapılar, biyolojik moleküller de bu yapının temel noktaları olarak tanımlanabilirler. Yaşamsal süreklilik açısından sistemin herhangi bir noktasında oluşacak bir sinyalin, hücreler arasında, hücre içinde veya hücre dışından hücre içine iletilmesi büyük önem taşımaktadır. Bu çalışmada, sistematik bir çerçevede HLA (İnsan Lökosit Antijenleri-Human Leukocyte Antigen) molekülleri üzerinde çalışılmıştır. HLAlar hücrelere yapışan yabancı proteinleri tanıyarak kendinden olan ve olmayanı ayırmakta kullanılır, bağışıklık sisteminin temelini oluştururlar. Çalışma kapsamında hesapsal bir yöntem olan Gaussian Ağ yapı modeli uygulanarak HLA sınıf 1 grubuna mensup yapıların yüksek frekanslarda öne çıkan bölgeleri tespit edilmiştir. Tüm moleküllerde hızlı modlarda aynı rezidülerin ön plana çıktığı görülmekte olup bu rezidülerin arasında bir iletişim patikası olduğu düşünülmektedir.

Dynamic Characterization of HLA Molecules by GNM

Keywords
Protein Dynamics,
Gaussian network
model,
HLA,
Fast modes

Abstract: Biological systems communicate within a network and biological molecules are the main building blocks of this network. For the continuation of life, it is necessary for the signals to be transmitted within and among the cells in a living organism. This study is conducted on HLA (Human Leukocyte Antigen) structures in a systematic manner. HLAs having a role in immune system recognize the foreign substances and attack them. By employing Gaussian Network Model (GNM) on HLA class I molecules, the residues that are energetically active at high frequency modes are detected. In all the molecules, same residues are observed to have the highest fluctuating profile, which are thought to form an interaction network.

*Sorumlu yazar: pemra.ozbek@marmara.edu.tr

1. Giriş

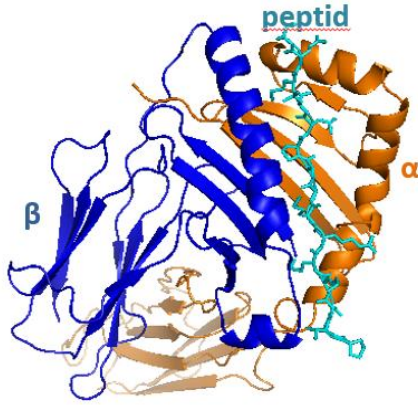
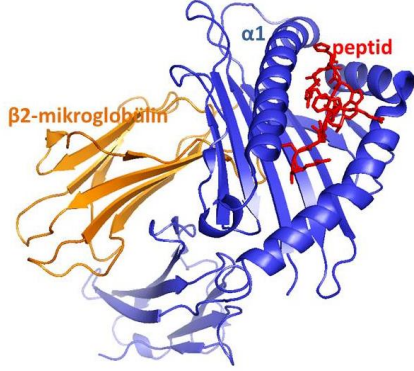
Proteinler/biyolojik makro moleküller hücreyi oluşturan yapı taşlarıdır. Esnek ve dinamik bir yapıya sahip olan proteinler, hücrenin yaşamsal fonksiyonlarını sağlamak için birbirleri ile iletişim kurarlar. Hücreye dışarıdan gelen veya hücre içerisinde oluşan etkilerin ve/veya bu etkilere karşı oluşturulan tepkilerin hücre içinde bir yerden başka bir yere veya hücre zarına iletimi bu moleküller arasındaki etkileşimler sayesinde gerçekleşmektedir. Moleküller kendilerine ulaşan bir sinyale tepki olarak, yapısal birtakım değişikliklere gidebildiği gibi, bazen de yalnızca içsel dinamiğindeki değişikliklerle de kendilerinden sonraki adıma iletirler. Sinyal iletim mekanizmasındaki bozukluklar protein fonksiyonuna etki eden mutasyonlarla ilişkilendirilebilir.

Omurgalılarda bulunan MHC (Major Histocompatibility Complex, Büyük Doku Uygunluk Kompleksi) sistemi insanlarda ilk kez lökositlerde gözlemlenmiş olmasından dolayı İnsan Lökosit Antijenleri-Human Leukocyte Antigen (HLA) olarak adlandırılmıştır. Genetik bilgisi 6 numaralı kromozomun kısa kolunda yerleşik bulunan HLA özellikleri doku özellikleridir. Bağışıklık sistemi fonksiyonları ile ilişkili olup tıpkı kan grupları gibi kalıtsaldır. Bu genlerdeki proteinler antijen olarak da bilinirken bağışıklık sistemi açısından son derece önemlidir. HLAlar hastalıklara karşı savunma mekanizması ve kanserle mücadelede etkin role sahipken otoimmün sistemiyle de ilişkilidir. HLAlar tarafından kodlanan proteinler kişiye özeldir ve bağışıklık sistemi tarafından kendinden olan ve olmayanı ayırmakta kullanılır. Bu sayede hastalık tetikleyici yabancı proteinlere karşı savunma mekanizması oluşturur. Hücrelere yapışan yabancı proteinleri tanı, savunma hücreleri de onlara

saldırır. Bazı durumlarda ise vücudun kendi hücrelerine de tıpkı yabancı hücrelere saldırdığı gibi saldırır ve otoimmün hastalıklarına sebep olur [1]. Bu hastalıkların mekanizması ya da sebebi halen daha çözülememiş konulardan olup araştırmaların çoğu otoimmün hastalıkların hangi genle ilişkili olduğunu tayin edip sonrasında da bu genlerin işleyiş mekanizmasını anlayarak hastalıkların önüne geçmeye odaklanmıştır [2]. HLA proteinlerine dair bilgilerin birçoğu organ bağışi çalışmalarından sağlanmıştır [3]. IMGT/HLA veri bankasında [4] yaklaşık 9000 HLA dizisi bulunmasına rağmen belirli sayıda proteinin 3 boyutlu yapı bilgisi mevcuttur. Organ nakillerinde HLA özelliklerinin (doku özellikleri) uyumu son derece önemlidir, aksi takdirde bağışıklık sistemi tarafından diğer insanların doku antijenleri yabancı bir doku olarak algılanır ve reddetme tepkilerine yol açabilir. Fizyolojik işlevleri protein antijenlerinden kaynaklanan peptidleri antijene özgül T lenfositlerine göstermektedir. MHC lokusu tüm hayvanlarda bulunan bir gen koleksiyonudur. Tüm türlerde MHC lokusu iki set olarak birçok polimorfik gen bulundurur, bunlar sınıf I ve sınıf II MHC genleri olarak anılır. Bu genler T hücrelerine peptid gösteren sınıf I ve sınıf II MHC moleküllerini kodlar.

HLA alfa-büklümlerle(helix) birbirine bağlı beta-sheetlerden oluşan alfa/beta motif oluşturan bir moleküldür (Şekil 1). β 2-mikroglobülin antiparalel beta-sheet döngülerle(loop) birbirine bağlıdır. Ayrıca beşik yapı modeli de göstermektedir. Her bir sınıf I molekülü bir α zinciri ve ona kovalan bağlı olmayan bağlarla bağlanan, β 2-mikroglobülin olarak adlandırılan ve MHC dışında kodlanan bir proteinden oluşur. Sınıf I molekülünün amino ucunun α 1 ve α 2 büklümleri 8-11 amino asit uzunluğunda peptidlerin

yerleşebileceği büyüklükte bir peptid bağlama oluşu oluşturur. Peptid bağlama oluşunun tabanı T



Şekil 1. Sınıf I HLA-B*27:09 aleli olan 1OF2 [5] protein yapısının (üst) ve Sınıf II HLA-DR3 yapısı olan 1A6A [6] proteininin (alt) Pymol ("The PyMol Molecular Graphics System, Version #, Schrodinger, LLC.") gösterimi

lenfositlerine gösterilecek peptidlerin bağlandığı kısım, oluşun yanları ve üst kısmı ise T hücre reseptörü ile temas halinde bulunan kısımdır. Sınıf I moleküllerinin polimorfik rezidüleri, yani farklı bireylerin MHC molekülleri arasında farklılık gösteren amino asitleri, α zincirinin $\alpha 1$ ve $\alpha 2$ alt birimlerinde yer alır. Zincirin $\alpha 3$ birimi değişken değildir ve T hücre eş-reseptörü CD8'in bağlanma noktasını içerir. Her sınıf II MHC molekülü bir α ve bir β zincirinden oluşur. Her iki zincirin amino ucu, $\alpha 1$ ve $\beta 1$ alt birimleri, polimorfik noktaları içerir ve 10-30 amino asit uzunluğunda

peptidlerin yerleşebileceği büyüklükte bir peptid bağlama oluşu oluşturur. Değişken olmayan $\beta 2$ alt birimi T hücre eş-reseptörü CD4'ün bağlanma noktasını içerir.

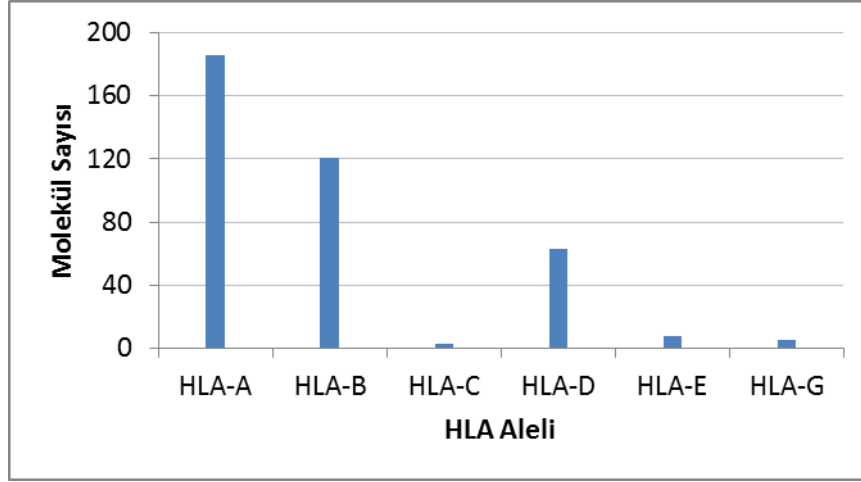
Bu çalışma kapsamında üç boyutlu protein yapısı belli olan HLA proteinleri literatür aracılığı ile sınıflarına göre gruplanarak bir veri seti hazırlanmıştır. Hesapsal yöntem olarak Elastik Ağ modellerinden Gaussian Ağ Yapı modeli [7, 8] kullanılarak HLA sınıf I molekülleri genelinde belirgin bir patikanın varlığı araştırılmıştır.

2. Materyal ve Metot

2.1. Veri Setinin Oluşturulması

PDB [10] veritabanı arama fonksiyonu kullanılarak HLA araması sonucu toplam 475 yapı sonuç olarak listelenmektedir. 475 yapı içerisinde 66 yapının 3D yapısında ribozom kompleksi olduğu tespit edildiğinden bu 66 yapı elenmiştir. Bu bilgiye ek olarak PDB veritabanında HLA sınıfına göre bir gruplama yapılmadığından bu amaç için özel olarak oluşturulmuş olan IEDB [9] veritabanında arama fonksiyonu kullanılarak 3D MHC araması yapılmıştır. Arama sonucunda 322 yapı listelenmiştir. Bu iki liste birbirleriyle karşılaştırılıp birleştirilmiştir. Daha sonra insandan elde edilmemiş olan girdiler elenerek 386 adet yapı içeren son liste oluşturulmuştur. 386 yapıdaki dağılım şu şekildedir:

- 186 tanesi HLA-A
- 121 tanesi HLA-B
- 3 tanesi HLA-C
- 63 tanesi HLA-D
- 8 tanesi HLA-E
- 5 tanesi HLA-G'dir (Şekil 2).



Şekil 1. Verisetinde yer alan HLA alel tiplerinde bulunan molekül sayısı

Her bir alel grubu içerisindeki sınıflandırılma Şekil 3’de verilmiştir. Buradaki yatay kolonlar, her bir alel grubundaki üç boyutlu yapısı belirli olan protein sayısını sembolize etmektedir.

Verisetindeki tüm moleküller X-Ray deney metoduyla elde edilmiştir. PDB [10] ve IEDB [11] veritabanlarından gelen bilgilerde eksik olan peptid sekans bilgisi ve molekül zincir bilgisi PDBsum [12] ve literatür aracılığı ile tamamlanmıştır. IEDB veritabanından gelen alel tipleri verisinde eksik olan kısımlar ve PDB verilerinde karşılaşılabilen eksik/fazla rezidülerin belirlenmesi için ccPDB [13] sunucusu kullanılarak sınıf I moleküllerin HLA zincirleri toplu halde FASTA formatında elde edilmiş, daha sonra FASTA dosya formatında çıktı alınarak filogenetik ağaç şeması oluşturulması için “Phylogeny.fr Robust Phylogenetic Analysis For The Non-Specialist” [14] sunucusunda giriş dosyası olarak kullanılmış ve her bir alel için filogenetik ağaç şemaları elde edilmiştir. Böylelikle farklı moleküllerin HLA zincirleri arasındaki farklılıklar ve benzerlikleri elde edilmiştir. Aynı bölgede gruplanan moleküller genellikle aynı alel tipine ait olmakla birlikte, tek moleküllerin oluşturduğu ayrı gruplar

genellikle PDB verisinde bulunan hataları ve eksiklikleri belirtmektedir (örn. deneysel olarak belirlenemeyen ancak varlığı bilinen rezidüleri).

Tüm protein yapıları için alel tipi, peptid tipi, çözünürlük, varsa PDB verisindeki mutasyon ve eksik rezidüleri bilgileriyle birlikte 386 yapı içeren veriseti tamamlanmıştır. Çalışmada kullanılan veri seti detaylı olarak Ek Tablo 1’de yer almaktadır. Bu çalışma kapsamında burada yer alan sınıf 1 molekülleri üzerinde detaylı olarak çalışılmıştır.

SINIF 1		SINIF 2	
A*68:01	3	B*57:03	6
A*24:02	7	B*57:01	6
A*11:01	4	B*53:01	2
A*03:01	3	B*52:01	1
A*02:07	1	B*51:01	2
A*02:06	1	B*44:05	7
A*02:03	1	B*44:03	5
A*02:01	164	B*44:02	9
A*01:01	2	B*41:04	1
C*04:01	2	B*41:03	1
C*03:04	1	B*35:08	16
E*01:03	2	B*35:01	21
E*01:01	6	B*27:09	10
G*01:01	6	B*27:05	17
		B*15:01	3
		B*14:02	1
		B*08:01	12
		B*07:02	1
		DRB3*03	2
		DRB3*01	1
		DRB1*15	1
		DRB1*04	4
		DRB1*01	5
		DRA*01	17
		DR4	1
		DR3	2
		DR2	1
		DR1	22
		DQA1	3
		DQ8	2
		DQ	1
		DP	1

Şekil 2. Verisetinde yer alan HLA alellerinin kendi içinde sınıflandırılması

2.2. Gaussian Ağ Yapı Modeli'nin Tüm Sınıf I Moleküllerine Toplu Olarak Uygulanması

Gaussian Ağ yapı modeli (GNM) [7, 8] veri setindeki tüm sınıf 1 moleküllerine toplu olarak uygulanmıştır. Bu metot ile tek tek her bir protein yapısının harmonik titreşimsel analizi ile yüksek frekanslı modlardaki hareketleri elde edilmiştir. GNM çıktısı olarak da her bir rezidünün diğer rezidüler ile yapmış olduğu dalgalanmalarının toplamını ifade eden D değeri gösterilmiştir. Bu sayede proteinlerin yüksek frekanslarda ortak ve farklılık gösteren bölgeleri tespit edilmiş olup, fonksiyonlarındaki farklılıklar dalgalanmalarındaki farklılıklar ile açıklanmaya çalışılmış, mutasyonların dalgalanmalara etkisi araştırılmıştır. Bu ve benzeri enerji bazlı hesapsal yaklaşımlar geliştirilerek iletişim patikası üzerinde yer alan rezidülerin tahmini yapılmaktadır [15, 16]. Bu çalışmada da aynı gruba mensup proteinlerde benzer iletişim mekanizmalarının varlığı incelenmiştir.

İncelenen her bir protein yapısı için, PDB dosyasında yer alan her amino asitin sadece alfa karbonlarının koordinat değerleri esas alınmış ve Kirchhoff matrisi (Γ_{ij}), aşağıdaki gibi oluşturulmuştur:

$$\Gamma_{ij} = \begin{cases} -1 & \text{eğer } i \neq j \text{ ve } R_{ij} \leq r_c \\ 0 & \text{eğer } i \neq j \text{ ve } R_{ij} > r_c \\ -\sum_{i,j \neq i} \Gamma_{ij} & \text{eğer } i = j \end{cases} \quad (1)$$

Kirchhoff matrisi, her bir amino asitin birbiri ile belirlenmiş olan bir sınır etkileşim mesafesi (r_c) dahilinde bağlı olup olmadığı bilgisini içeren ve böylelikle protein yapısının topolojik bilgisini kabaca içeren bir matristir. Bu çalışmada sınır etkileşim mesafesi olarak, literatürde de sıklıkla kullanılan 7 Å kullanılmıştır.

Gaussian Ağ Yapı Modeli'nde i ve j rezidüleri arasındaki çapraz korelasyon $\langle \Delta R_i \cdot \Delta R_j \rangle$, Kirchhoff matrisinin tersi $([\Gamma^{-1}]_{ij})$ ile aşağıdaki gibi ilintilidir [8].

$$\langle \Delta R_i \cdot \Delta R_j \rangle = \frac{3k_B T}{\gamma} [\Gamma^{-1}]_{ij} \quad (2)$$

Yukarıdaki eşitlikte k_B Boltzmann katsayısını (J/K), T mutlak sıcaklık değerini (K) ve γ da ağ yapı modelinde yay sabitini ifade etmektedir. Bu değerlerin hepsi birer sabit sayı olduğundan, $\langle \Delta R_i \cdot \Delta R_j \rangle$ ve Γ_{ij} matrisleri birbirileri ile doğrudan ilişkilidirler. Bunun sonucunda da, sistemin tüm dinamiğini $[\Gamma^{-1}]_{ij}$ matrisi belirlemektedir. Bu çalışmada hesaplanan çapraz korelasyon matrisleri, bu ilintiler doğrultusunda doğrudan $[\Gamma^{-1}]_{ij}$ matrisine eşit olarak kabul edilmişlerdir. Dolayısıyla fiziksel herhangi bir birim ile ifade edilmemişlerdir.

$[\Gamma^{-1}]_{ij}$ matrisi, doğrudan Kirchhoff matrisi'ne tekil değer ayrışımı uygulanarak bulunur:

$$\begin{aligned} \langle \Delta R_i \cdot \Delta R_j \rangle &\cong [\Gamma^{-1}]_{ij} \\ &\cong [U \Lambda^{-1} U^T]_{ij} \end{aligned} \quad (3)$$

Yukarıdaki eşitlikte U özvektör matrisini, Λ ise köşegen değerleri özdeğerlerden oluşan matrisi ifade etmektedir. Λ matrisi aynı zamanda karakteristik kök matrisi olarak da isimlendirilir ve sistemin toplam hareketini temsil eden modların frekansını temsil eder. Tekil değer

ayrışımı, aynı zamanda tekil olarak modları ifade eden özvektörlerin ve özdeğerlerin toplamı olarak da ifade edilebilir:

$$\begin{aligned} \langle \Delta R_i \cdot \Delta R_j \rangle &\cong [U \Lambda^{-1} U^T]_{ij} \\ &\cong \sum_k [\lambda_k^{-1} u_k u_k^T]_{ij} \end{aligned} \quad (4)$$

Yukarıdaki eşitlikte λ_k k'inci modun özdeğerini, u_k ise k'inci modun özvektörünü ifade etmektedir. Tekil değer ayrışımını takiben yapılan bu işlem, sistemin eşitlik pozisyonundan yaptığı dalgalanma hareketinin (titreşimlerinin), frekanslarına göre farklı modlara ayrılarak incelenmesine olanak sağlar. N amino asite sahip bir sistemin karakteristik kök matrisinin (Λ) sıfırdan farklı N-1 adet modu bulunur.

λ_k değerleri arasında düşük değere sahip olanlar en yavaş olan kolektif ve global hareketleri, yüksek değerler de en hızlı olan, sınırlı sayıda rezidünün katıldığı yerel hareketleri ifade ederler. Bu çalışmada HLA moleküllerinde moleküler yapıda rezidüler arasındaki olası etkileşim patikalarının molekül topolojisinden bulunması amaçlandığından, frekansı yüksek olan yerel hareketlere odaklanılmıştır. Bu kapsamda en yüksek frekansa sahip olan ilk üç mod hesap kapsamına alınmıştır. Her üç mod için ayrı olarak i ve j rezidüleri arasındaki göreceli dalgalanmayı veren ortalama kareler dalgalanma matrisleri $\langle \Delta R_{ij}^2 \rangle$, aşağıdaki gibi hesaplanmıştır:

$$\begin{aligned} \langle \Delta R_{ij}^2 \rangle &= \langle (\Delta R_i - \Delta R_j)^2 \rangle \\ &= \langle \Delta R_i^2 \rangle + \langle \Delta R_j^2 \rangle - 2 \langle \Delta R_i \Delta R_j \rangle \end{aligned} \quad (5)$$

Ek olarak, her üç modu da temsil eden ve özdeğerlere göre hesaplanmış bir “ağırlıklı ortalama” kareler dalgalanma matrisi de aşağıdaki gibi hesaplanmıştır:

$$\langle \Delta R_{ij}^2 \rangle_{1-2-3} = \frac{\sum_{k=1}^3 \langle \Delta R_{ij}^2 \rangle_k \cdot \lambda_k^{-1}}{\sum_{k=1}^3 \lambda_k^{-1}} \quad (6)$$

Ortalama kareler dalgalanma matrisi, her bir i rezidüsünün, her bir j rezidüsü ile olan göreceli dalgalanmaları arasındaki korelasyonu ifade etmektedir. Ağ yapısında önemli rol oynayan rezidüleri bulabilmek için, simetrik olan $\langle \Delta R_{ij}^2 \rangle$ matrisinde her bir i rezidüsü için i satırında veya kolonunda bulunan değerler toplanarak, gösterge bir değer elde edilebilir:

$$D_i = \sum_j \langle \Delta R_{ij}^2 \rangle \quad (7)$$

Yukarıdaki eşitlikte D_i değeri, i rezidüsünün ağ yapısı içerisinde diğer tüm rezidüler ile etkileşiminin toplamını ifade eder. Yüksek D_i değerine sahip olan rezidülerin, ilgili moddaki harekette “merkezi” bir rol oynadıkları varsayılabilir. Rezidüler arasındaki ortalama kareler dalgalanma matrisinin, sistemin dışarıdan aldığı enerji sonucu oluşan ortalama enerji dalgalanmaları (ΔU) ile de ilintili olduğu varsayılabilir.

$$\langle \Delta U \cdot (\Delta R_{ij})^2 \rangle = k \cdot T \cdot \langle (\Delta R_{ij})^2 \rangle \quad (8)$$

Dolayısı ile D_i değeri aynı zamanda enerji dalgalanmaları ile de aşağıdaki gibi ilişkilidir:

$$D_i = \sum_j \langle \Delta R_{ij}^2 \rangle \quad (9)$$

$$= \frac{1}{k \cdot T} \sum_j \langle \Delta U \cdot (\Delta R_{ij})^2 \rangle$$

Yukarıda detayları verilen GNM metodu, veri setindeki tüm sınıf I moleküllerine toplu olarak uygulanmıştır. Sınıf I moleküllerinde özellikle ilgi konusu olan bölge peptid bağlanma bölgesi olduğundan, metod sadece antijen zinciri (HLA zinciri) temel alınarak uygulanmıştır. Bu zincir, peptidin bağlandığı $\alpha 1$, $\alpha 2$ ve ayrıca $\alpha 3$ alt birimlerinden oluşur ve kristalografide belirlenebilen toplam amino asit sayısı bir kaç istisna dışında 276’dır.

2.3. HLA Yapısında Rezidü Korunurluğu Değerlerinin Elde Edilmesi

HLA yapısında bulunan her bir amino-asit rezidü pozisyonunun farklı HLA dizileri de göz önünde bulundurularak korunurluğunu elde etmek için, ConSurf 2016 sunucusu kullanılmıştır [17]. Sunucu, verilen bir PDB yapısı ile başlayarak, homolog dizileri otomatik olarak UniRef90 veri tabanından bulmakta, bunların çoklu dizi eşleşmelerini gerçekleştirerek diziler arası evrimsel ilişkiyi ifade eden filogenetik ağaç oluşturmaktadır. Bu veri, dizideki her bir pozisyonun evrimsel değişim hızını tahmin etmek amacıyla kullanılmaktadır.

Veri setinde bulunan her bir Sınıf I HLA aleli için, yapı çözünürlük değeri (Angstrom cinsinden) en düşük olan yapılar seçilmiştir. Bu seçilen yapıların her biri için ConSurf sunucusu kullanılarak rezidü pozisyonlarının korunurluğunu ifade eden ve 0 ile 9 arasında değişen “skor” değerleri elde edilmiştir. HLA yapısındaki her bir pozisyon için tek bir skor değeri elde etmek için, veri setindeki her bir yapıda, ilgili pozisyonun skor değerinin aritmetik ortalaması alınmıştır.

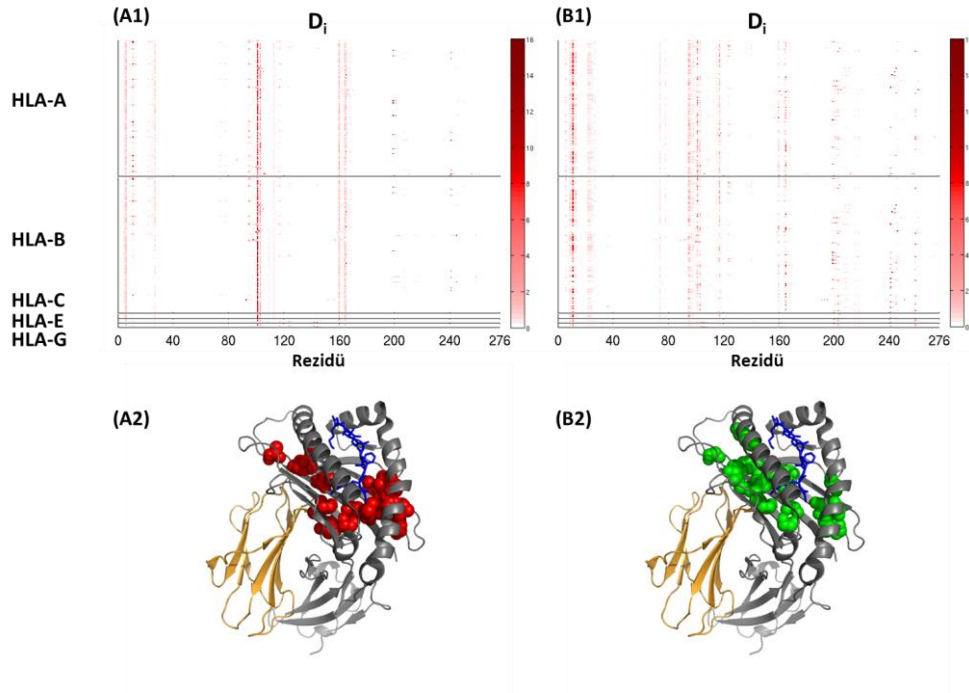
2.4. Ortalama Kare-Kök Sapma Hesaplaması

Oluşturulan veri setinde bulunan HLA yapıları arasındaki Ortalama Kare-Kök Sapma (RMSD) hesabı, veri setinde bulunan her bir Sınıf I HLA yapısının bir değerine üstüne çakıştırılarak, yapılar arasında alfa karbonların aralarındaki ortalama mesafelerinin ortalaması hesaplanarak gerçekleştirilmiştir. Buna göre, iki yapı arasındaki RMSD değeri ne kadar yüksek ise, aralarındaki benzerlik o kadar düşüktür. Dolayısıyla, RMSD, iki yapı arasındaki benzerliğin/ bir göstergesi olarak işlev görür.

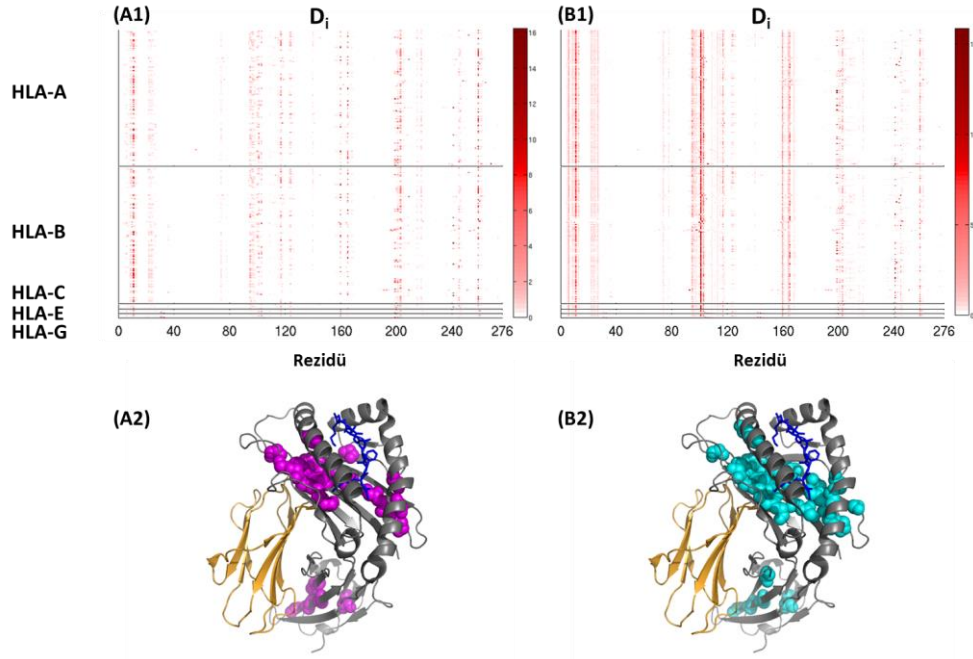
3. Bulgular

Molekül sayısı oldukça fazla olduğundan, sonuçlar belirtilen modların ve ağırlıklı ortalama modunun değerlerinin topluca gösterimi olarak Şekil 4 ve Şekil 5’de A1 ve B1 bölmelerinde verilmiştir. Her moddan elde edilen baskın (görelî yüksek) D

değerlerinin molekül üzerinde hangi rezidülere karşılık geldiği de sırasıyla A2 ve B2 bölmelerinde gösterilmiştir. Buna göre Şekil 4’de A1 ve A2 en hızlı mod olan mod 1’e, B1 ve B2 en hızlı ikinci mod olan mod 2’ye, Şekil 5’de ise A1 ve A2 en hızlı üçüncü mod olan mod 3’e, B1 ve B2 ise bu modların ağırlıklı ortalama modu olan mod 1-2-3’e denk gelmektedir. A1 ve B1 bölmelerinde verilen grafiklerde X eksenî antijen zincirinin sıralı rezidü sayısını ifade ederken, Y eksenî ise genlere göre gruplandırılmış molekülleri ifade etmektedir. Y ekseninde gen seviyesinde gruplandırmanın ötesinde herhangi bir gruplandırma yoktur ve molekülün herhangi bir şekilde artan veya azalan kantitatif bir özelliğini ifade etmemektedir. A2 ve B2 bölmelerinde baskın D değerine sahip olan rezidüler van der Waals kürecikleri şeklinde gösterilmişlerdir.



Şekil 4. Sınıf I için (A1&A2)GNM mod 1 D_i değerleri ve üç boyutlu yapı üzerinde etkileşim patikalarının gösterimi (kırmızı küreler) (B1&B2) GNM mod 2 ve D_i değerleri üç boyutlu yapı üzerinde etkileşim patikalarının gösterimi (yeşil küreler)



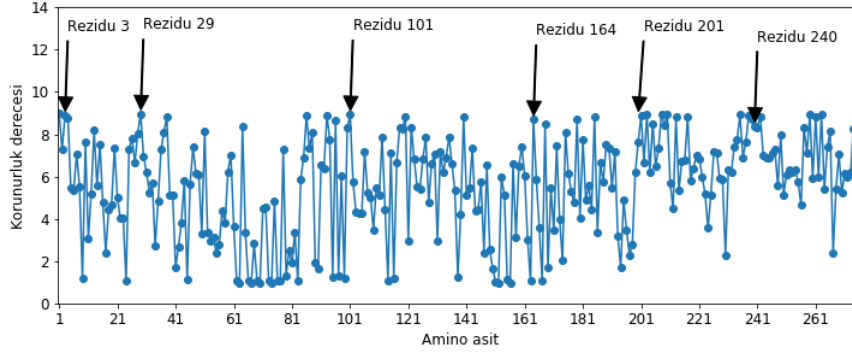
Şekil 5. Sınıf I için (A1&A2)GNM mod 3 D_i değerleri ve üç boyutlu yapı üzerinde etkileşim patikalarının gösterimi (pembe küreler) (B1&B2) GNM mod 1-2-3 D_i değerleri ve üç boyutlu yapı üzerinde etkileşim patikalarının gösterimi (mavi küreler)

Şekil 4 A1 bölümünde de görülebileceği üzere, en hızlı mod olan mod 1’de Sınıf I moleküllerinin tamamının büyük oranda aynı rezidüler etrafında bir iletişim patikasına sahip olduğu gözlenmektedir. Özellikle Cys-101 numaralı rezidü, iletişim patikasında merkezi bir rol oynamakta, patikada bunun dışında Arg-6, Tyr-27 ve Leu-160 numaralı rezidüler öne çıkmaktadır. Bu rezidülerin HLA molekülünün alfa-1 bölgesinin Arg-6 numaralı rezidüsünden alfa-2 bölgesinin Leu-160 numaralı rezidüsüne uzanan bir yapıda birbirleri ile etkileşim halinde olduğu gözlenmiştir.

En hızlı ikinci mod olan mod 2, benzer bir rezidü etkileşim profili vermektedir (Şekil 4 B1 bölümü). Mod 1 ile arasındaki farklılık, Ser-13 numaralı rezidünün tüm Sınıf I moleküllerinde daha merkezi bir rol oynamasıdır.

En hızlı üçüncü mod olan mod 3, yine aynı profile, mod 2’ye benzer sonuç vermektedir (Şekil 5 A1 bölümü). Diğer modlar ile arasındaki fark, bu modda $\alpha 3$ alt biriminde yer alan Thr-200 ve Thr-240 numaralı rezidüler etrafında kümelenen baskın rezidülerin bu modda daha belirgin olarak gözlenmesidir.

Şekil 4 ve Şekil 5’te bahsedilen profilden sapmalar da gözlenmektedir. GNM yönteminde, bu şekilde yapısal olarak farklı bölgeler içerisinde görülen yerel titreşimlerin farklı moleküler yapılarında frekanslar arasında kaymalar yapabildiği bilinmektedir.



Şekil 6. Sınıf I HLA yapılarında her bir amino asit pozisyonunun ConSurf 2016 sunucusundan elde edilen korunurluk dereceleri.

Dolayısıyla, bu sapmaların belirgin olarak farklı iletişim patikalarını ifade etmedikleri, bu sebeple de farklı aleller arasındaki mutasyonlardan kaynaklanan farklı profilleri ifade etmedikleri değerlendirilmiştir. Bu tez, mod 1, 2 ve 3'ten daha önce açıklandığı şekliyle elde edilen ağırlıklı ortalama modu olan mod 1-2-3'ün Şekil 5 B1 bölümünde verilen D_i profili tarafından da doğrulanmaktadır. Bu grafikte de görülebileceği üzere, özellikle baskın olarak görülen rezidüler, en hızlı üç modda tüm alel tiplerinde ortak olarak gözlenmektedir. Veri setindeki tüm HLA sınıf I yapıları arasındaki ortalama kare-kök sapma (RMSD) dağılımının 0.5 - 1 Å aralığında olması (Ek B), farklı aleller arasında katlanmış HLA yapısının belirgin şekilde farklılık göstermediğini göstermektedir. Ancak, RMSD dağılımının 0.7 Å etrafında kümelenmiş olması, yapılar arasında gözle görülemeyen farklılıkları da ifade edebilir. Bu olası farklılıkların, HLA protein dinamiğinde hızlı modlarda rol oynayan amino asitlerin pozisyonlarında değişikliğe yol açmadığı gözlemlenmiştir. Ancak, kullandığımız GNM yöntemin iri taneli bir hesapsal yöntem olduğu ve amino asit yan zincirleri arasındaki farklılığı doğrudan hesaba katmadığı da göz önünde bulundurulmalıdır.

Elde edilen bu rezidüler ve oluşturdukları iletişim patikasının HLA molekülünün fonksiyonundaki rolü literatürde net olarak ortaya konulmamıştır. 1-180 arası rezidülerin tamamı peptit bağlanmasında olası rol sahibi olabilir. Ancak, hızlı hareket modlarının proteinin katlanmış halini korumada önemli rol oynadığı ve dolayısıyla proteinin ne kadar kararlı olacağını belirleyen önemli faktörlerden olduğu raporlanmıştır [8, 16]. HLA-A02:01 alelinde Cys-101 ve Cys-164 arasında bir disülfid köprüsünün olduğu ve bu köprü mutasyon ile bozulduğunda protein kararlılığının etkilendiği [18]. Cys-164 ve bu rezidüye dizi komşuluğu bulunan rezidüler Cys-101 ile birlikte GNM hızlı modlarında önemli rol oynamaktadırlar.

HLAların insandaki en polimorfik protein olduğu dikkate alındığında, elde edilen iletişim patikasındaki rezidülerle bu rezidülerin korunurluk derecelerinin karşılaştırılması, doğal olarak oluşan HLA polimorfizminin protein işlevi açısından açıklanmasında faydalı olabilir. HLA yapısındaki her bir rezidü pozisyonunun korunurluk değerleri ConSurf sunucusu ile elde edilmiştir. Şekil 6, ConSurf sunucusundan elde edilen korunurluk değerlerini göstermektedir. Korunurluk değerleri görece yüksek olan GNM hızlı

modlarındaki iletişim patikasında yer alan veya dizi komşusu olan rezidüler, oklar ile belirtilmiştir. Buna göre, GNM hızlı modlarında rol oynayan rezidülerin belirgin bir şekilde dizideki diğer rezidülere göre daha korunmuş amino asitler oldukları gözlenmektedir.

3. Sonuç

HLA sınıf 1 molekülleri üzerine uygulanan GNM yöntemi sonucunda hızlı modlarda istisnasız aynı rezidülerin ön plana çıktığı görülmektedir. Bu rezidülerin (Arg-6, Tyr-27, Cys-101, Glu-163, Val-165) arasında bir iletişim patikası olduğu düşünülmektedir. Mod sayısının artması ile mevcut bu rezidülere ek olarak özellikle Ser-13, Gly-120, Thr-200 ve Thr-240 numaralı rezidülerin de profillerde baskın olarak görüldüğü gözlemlenmiştir. Aynı sisteme mensup 10 protein üzerinde yapılan önceki çalışmalarda da [15] aynı bölgelerin öne çıktığı bulunmuştu. Bu çalışma ile bu bölgelerin HLA sınıf 1 sistemi üzerinde genelleştirilmesi mümkün olabilmektedir. Bunun yanı sıra, elde edilen rezidülerin HLA yapısında görece olarak daha korunmuş amino asitler olduğu gözlemlenmiştir. Bu bulgu, protein dinamiğinde hızlı titreşim gösteren amino asitlerin protein kararlılığında önemli rol oynadığı değerlendirildiğinde, HLA proteininde görülen polimorfizm ile protein kararlılığı arasında var olduğu bilinen ilişkiyi hesapsal olarak da göstermektedir. Daha sonraki çalışmalarda bu iletişim patikasının farklı yöntemler ile de tayin ve tespitinin yanı sıra mutasyonların bu bölgelere etkileri üzerinde çalışılması mümkündür.

Teşekkür

Bu çalışma TÜBİTAK 113M293 nolu projenin sağladığı destek ile tamamlanmış olup, TÜBİTAK'a desteğinden dolayı teşekkür ederiz.

Kaynakça

- [1] Gough, S.C. ve M.J. Simmonds. 2007. The HLA Region and Autoimmune Disease: Associations and Mechanisms of Action. *Curr Genomics*, Cilt 8(7), s. 453-65.
- [2] Ghodke, Y. vd. 2005. HLA and disease. *Eur J Epidemiol*. Cilt 20(6), s. 475-88.
- [3] Schreuder, G.M. vd. 2001. The HLA Dictionary 2001: a summary of HLA-A, -B, -C, -DRB1/3/4/5 and -DQB1 alleles and their association with serologically defined HLA-A, -B, -C, -DR and -DQ antigens. *Eur J Immunogenet*, Cilt 28(6), s. 565-96.
- [4] Robinson, J. vd. 2011. The IMGT/HLA database. *Nucleic Acids Res*, 2011. Cilt 39(Database issue), s. D1171-6.
- [5] Starikov, E.B., L. Nilsson ve M. Hulsmeier. 2004. A single residue exchange between two HLA-B27 alleles triggers increased peptide flexibility. *Eur Biophys J*, Cilt 33(7), s. 651-5.
- [6] Ghosh, P., vd. 1995. The structure of an intermediate in class II MHC maturation: CLIP bound to HLA-DR3. *Nature*, Cilt 378(6556), s. 457-62.
- [7] Haliloglu, T., I. Bahar ve B. Erman. 1997. Gaussian dynamics of folded proteins. *Phys Rev Lett*, Cilt 79(16), s. 3090-3093.
- [8] Bahar, I. vd. 1998. Vibrational dynamics of folded proteins: Significance of slow and fast motions in relation to function and stability. *Phys Rev Lett*. Cilt 80(12), s. 2733-2736.
- [9] Ponomarenko, J. vd. 2011. IEDB-3D: structural data within the immune epitope database. *Nucleic Acids Res*, Cilt 39(Database issue), s. D1164-70.
- [10] Berman, H.M. vd. 2000. The Protein Data Bank. *Nucleic Acids Res*, Cilt 28(1), s. 235-42.

- [11] Vita, R. vd. 2015. The immune epitope database (IEDB) 3.0. Nucleic Acids Res, Cilt 43(Database issue), s. D405-12.
- [12] de Beer, T.A. vd. 2014. PDBsum additions. Nucleic Acids Res. Cilt 42(Database issue), s. D292-6.
- [13] Singh, H. vd. 2012. ccPDB: compilation and creation of data sets from Protein Data Bank. Nucleic Acids Res. Cilt 40(Database issue), s. D486-9.
- [14] Dereeper, A. vd. 2008. Phylogeny.fr: robust phylogenetic analysis for the non-specialist. Nucleic Acids Res, 2008. Cilt 36(Web Server issue), s. W465-9.
- [15] Haliloglu, T., A. Gul ve B. Erman. 2010. Predicting Important Residues and Interaction Pathways in Proteins Using Gaussian Network Model: Binding and Stability of HLA Proteins. Plos Computational Biology, Cilt 6(7).
- [16] Ozbek, P., S. Soner ve T. Haliloglu. 2013. Hot Spots in a Network of Functional Sites. PLoS One, Cilt 8(9).
- [17] Ashkenazy. H. vd., 2016. ConSurf 2016: an improved methodology to estimate and visualize evolutionary conservation in macromolecules. Nucleic Acids Research, Cilt 44(Web Server issue), s. W344-W350
- [18] Warburton RJ. vd., 1994. Mutation of the alpha 2 domain disulfide bridge of the class I molecule HLA-A*0201. Effect on maturation and peptide presentation. Hum Immunol, Cilt 39(4), s261-71.

Ekler

Ek A. Çalışmada Kullanılan Veri Seti
Ek B. Veri Setinde Bulunan Sınıf I Yapıları Arasındaki Alfa-Karbon RMSD Dağılımı

EK A – Çalışmada Kullanılan Veri Seti**Ek Tablo 1.** Çalışmada Kullanılan Veri Seti

PDB Adı	Sınıf	Allel Adı	Peptid	Çözünürlük (Å)
1W72	1	HLA-A*01:01	EADPTGHSY	2.15
3B08	1	HLA-A*01:01	EADPTGHSY	1.80
1AKJ	1	HLA-A*02:01	ILKEPVHGV	2.65
1A07	1	HLA-A*02:01	LLFGYPVYV	2.60
1B0G	1	HLA-A*02:01	ALWGFFPVL	2.50
1B0R	1	HLA-A*02:01	GILGFVFTL	2.90
1BD2	1	HLA-A*02:01	LLFGYPVYV	2.5
1DUY	1	HLA-A*02:01	LFGYPVYV	2.15
1DUZ	1	HLA-A*02:01	LLFGYPVYV	1.80
1EEY	1	HLA-A*02:01	ILSALVGIV	2.25
1EEZ	1	HLA-A*02:01	ILSALVGIL	2.30
1HHG	1	HLA-A*02:01	TLTSCNTSV	2.60
1HHH	1	HLA-A*02:01	FLPSDFFPSV	3.00
1HHI	1	HLA-A*02:01	GILGFVFTL	2.50
1HHJ	1	HLA-A*02:01	ILKEPVHGV	2.50
1HHK	1	HLA-A*02:01	LLFGYPVYV	2.50
1HLA	1	HLA-A*02:01	0	3.50
1I1F	1	HLA-A*02:01	FLKEPVHGV	2.80
1I1Y	1	HLA-A*02:01	YLKEPVHGV	2.20
1I4F	1	HLA-A*02:01	GVYDGREHTV	1.40
1I7R	1	HLA-A*02:01	FAPGFFPYL	2.20
1I7T	1	HLA-A*02:01	ALWGVFPVL	2.80
1I7U	1	HLA-A*02:01	ALWGFVPVL	1.80
1IM3	1	HLA-A*02:01	LLFGYPVYV	2.20
1JF1	1	HLA-A*02:01	ELAGIGILTV	1.85
1JHT	1	HLA-A*02:01	ALGIGILTV	2.15
1LP9	1	HLA-A*02:01	ALWGFFPVL	2.00
1OGA	1	HLA-A*02:01	GILGFVFTL	1.40
1P7Q	1	HLA-A*02:01	ILKEPVHGV	3.40
1QEW	1	HLA-A*02:01	FLWGPRALV	2.20
1QR1	1	HLA-A*02:01	IISAVVGIL	2.40
1QRN	1	HLA-A*02:01	LLFGYAVYV	2.80
1QSE	1	HLA-A*02:01	LLFGYPRYV	2.80
1QSF	1	HLA-A*02:01	LLFGYPVAV	2.80
1S8D	1	HLA-A*02:01	SLANTVATL	2.20
1S9W	1	HLA-A*02:01	SLLMWITQC	2.20

1S9X	1	HLA-A*02:01	SLLMWITQA	2.50
1S9Y	1	HLA-A*02:01	SLLMWITQS	2.30
1T1W	1	HLA-A*02:01	SLFNTIAVL	2.20
1T1X	1	HLA-A*02:01	SLYLT VATL	2.20
1T1Y	1	HLA-A*02:01	SLYNV VATL	2.00
1T1Z	1	HLA-A*02:01	ALYNTAAAL	1.90
1T20	1	HLA-A*02:01	SLYNTIATL	2.20
1T21	1	HLA-A*02:01	SLYNT VATL	2.19
1T22	1	HLA-A*02:01	SLYNT VATL	2.20
1TVB	1	HLA-A*02:01	ITDQVPFSV	1.80
1TVH	1	HLA-A*02:01	IMDQVPFSV	1.80
2BNQ	1	HLA-A*02:01	SLLMWITQV	1.70
2BNR	1	HLA-A*02:01	SLLMWITQC	1.90
2C7U	1	HLA-A*02:01	SLFNTIAVL	2.38
2CLR	1	HLA-A*02:01	MLLSVPLL LG	2.00
2F53	1	HLA-A*02:01	SLLMWITQC	2.10
2F54	1	HLA-A*02:01	SLLMWITQC	2.70
2GIT	1	HLA-A*02:01	LLFGKPVYV	1.70
2GJ6	1	HLA-A*02:01	LLFGKPVYV	2.56
2GT9	1	HLA-A*02:01	EAAGIGILTV	1.75
2GTW	1	HLA-A*02:01	LAGIGILTV	1.55
2GTZ	1	HLA-A*02:01	ALGIGILTV	1.70
2GUO	1	HLA-A*02:01	AAGIGILTV	1.90
2P5E	1	HLA-A*02:01	SLLMWITQC	1.89
2P5W	1	HLA-A*02:01	SLLMWITQC	2.20
2PYE	1	HLA-A*02:01	SLLMWITQC	2.30
2V2W	1	HLA-A*02:01	SLYNT VATL	1.60
2V2X	1	HLA-A*02:01	SLFNT VATL	1.60
2VLJ	1	HLA-A*02:01	GILGFVFTL	2.40
2VLK	1	HLA-A*02:01	GILGFVFTL	2.50
2VLL	1	HLA-A*02:01	GILGFVFTL	1.60
2VLR	1	HLA-A*02:01	GILGFVFTL	2.30
2X4N	1	HLA-A*02:01	KILGGVFXV	2.34
2X4O	1	HLA-A*02:01	KLTPLCVTL	2.30
2X4P	1	HLA-A*02:01	MILGGVFXV	2.30
2X4Q	1	HLA-A*02:01	MILGGVFXV	1.90
2X4R	1	HLA-A*02:01	NLVPM VATV	2.30
2X4S	1	HLA-A*02:01	AMDSNTLEL	2.55
2X4T	1	HLA-A*02:01	NLVPM VATV	2.30

2X4U	1	HLA-A*02:01	ILKEPVHGV	2.10
2X70	1	HLA-A*02:01	KILGGVFXV	2.00
3BGM	1	HLA-A*02:01	RQASLSISV	1.60
3BH8	1	HLA-A*02:01	RQASIELPSM	1.65
3BH9	1	HLA-A*02:01	RTYSGPMNKV	1.70
3BHB	1	HLA-A*02:01	KMDSFLDMQL	2.20
3D25	1	HLA-A*02:01	VLHDDLLEA	1.30
3D39	1	HLA-A*02:01	LLFGFPVYV	2.81
3D3V	1	HLA-A*02:01	LLFGFPVYV	2.80
3FQN	1	HLA-A*02:01	YLDSGIHSGA	1.65
3FQR	1	HLA-A*02:01	YLDSGIHSGA	1.70
3FQT	1	HLA-A*02:01	GLLGSPVRA	1.80
3FQU	1	HLA-A*02:01	GLLGSPVRA	1.80
3FQW	1	HLA-A*02:01	RVASPTSGV	1.93
3FQX	1	HLA-A*02:01	RVASPTSGV	1.70
3FT2	1	HLA-A*02:01	VLRDDLLEA	1.80
3FT3	1	HLA-A*02:01	VLHDDLLEA	1.95
3FT4	1	HLA-A*02:01	VLRDDLLEA	1.90
3GIV	1	HLA-A*02:01	SLFNTVATLY	2.00
3GJF	1	HLA-A*02:01	SLLMWITQV	1.90
3GSO	1	HLA-A*02:01	NLVPMVATV	1.60
3GSQ	1	HLA-A*02:01	NLVPSVATV	2.12
3GSR	1	HLA-A*02:01	NLVPVATV	1.95
3GSU	1	HLA-A*02:01	NLVPTVATV	1.80
3GSV	1	HLA-A*02:01	NLVPQVATV	1.90
3GSW	1	HLA-A*02:01	NLVPMVAHV	1.81
3GSX	1	HLA-A*02:01	NLVPMVAVV	2.10
3H7B	1	HLA-A*02:01	MLWGYLQYV	1.88
3H9S	1	HLA-A*02:01	MLWGYLQYV	2.70
3HAE	1	HLA-A*02:01	SLLMWITQV	2.90
3HG1	1	HLA-A*02:01	ELAGIGILTV	3.00
3HLA	1	HLA-A*02:01	0	2.60
3HPJ	1	HLA-A*02:01	RMFPNAPYL	2.00
3I6G	1	HLA-A*02:01	GLMWLSYFV	2.20
3I6K	1	HLA-A*02:01	TLACFVLAHV	2.80
3KLA	1	HLA-A*02:01	SLLMWITQL	1.65
3MGO	1	HLA-A*02:01	RLYQNPTTYI	2.30
3MGT	1	HLA-A*02:01	KLYQNPTTYI	2.20
3MR9	1	HLA-A*02:01	NLVPAVATV	1.93

3MRB	1	HLA-A*02:01	NLVPMVHTV	1.40
3MRC	1	HLA-A*02:01	NLVPMCATV	1.80
3MRD	1	HLA-A*02:01	NLVPMGATV	1.70
3MRE	1	HLA-A*02:01	GLCTLVAML	1.10
3MRF	1	HLA-A*02:01	GLCPLVAML	2.30
3MRG	1	HLA-A*02:01	CINGVCWTV	1.30
3MRH	1	HLA-A*02:01	CISGVCWTV	2.40
3MRI	1	HLA-A*02:01	CINMWCWTV	2.10
3MRJ	1	HLA-A*02:01	CINGMCWTV	1.87
3MRK	1	HLA-A*02:01	PLFQVPEPV	1.40
3MRL	1	HLA-A*02:01	CINGVVWTV	2.41
3MRM	1	HLA-A*02:01	KLVALGINAV	1.90
3MRN	1	HLA-A*02:01	LLFNILGGWV	2.30
3MRO	1	HLA-A*02:01	ELAGWGILTV	2.35
3MRP	1	HLA-A*02:01	ELAGLGINTV	2.10
3MRQ	1	HLA-A*02:01	ELAGLGINTV	2.20
3MRR	1	HLA-A*02:01	LLAGIGTVPI	1.60
3MYJ	1	HLA-A*02:01	YMFPNAPYL	1.89
3O3A	1	HLA-A*02:01	XLAXXLTV	1.80
3O3B	1	HLA-A*02:01	ELAXXLTV	1.90
3O3D	1	HLA-A*02:01	ELAXXLTV	1.70
3O3E	1	HLA-A*02:01	XLAXXLTV	1.85
3O4L	1	HLA-A*02:01	GLCTLVAML	2.54
3PWJ	1	HLA-A*02:01	LLYGFVNYV	1.70
3PWL	1	HLA-A*02:01	LGYGfVNYI	1.65
3PWN	1	HLA-A*02:01	LLYGFVNYI	1.60
3PWP	1	HLA-A*02:01	LGYGfVNYI	2.69
3QDG	1	HLA-A*02:01	ELAGIGILTV	2.69
3QDJ	1	HLA-A*02:01	AAGIGILTV	2.30
3QDM	1	HLA-A*02:01	ELAGIGILTV	2.80
3QEQ	1	HLA-A*02:01	AAGIGILTV	2.59
3QFD	1	HLA-A*02:01	AAGIGILTV	1.68
3QFJ	1	HLA-A*02:01	LLFGFPVYV	2.29
3REW	1	HLA-A*02:01	CLGGLLTMV	1.90
3TO2	1	HLA-A*02:01	LACFVLA AV	2.60
3UTQ	1	HLA-A*02:01	ALWGPD PAAA	1.67
3V5D	1	HLA-A*02:01	KVAELVHFL	2.00
3V5H	1	HLA-A*02:01	KVAEIVHFL	1.63
3V5K	1	HLA-A*02:01	KVAELVWFL	2.31

4E5X	1	HLA-A*02:01	LLFGYPVYV	1.95
4EUP	1	HLA-A*02:01	ALGIGILTV	2,88
4EUQ	1	HLA-A*02:01	EAAGIGILTV	2,69
4GKN	1	HLA-A*02:01	FATGIGITV	2.75
4GKS	1	HLA-A*02:01	FLTGIGITV	2.35
4I4W	1	HLA-A*02:01	ILAKFLHWL	1.77
3H9H	1	HLA-A*02:01	MLWGYLQYV	2.00
3IXA	1	HLA-A*02:01	LLFGYPVYV	2.10
2AV7	1	HLA-A*02:01	LLFGYPVYV	2.05
2J8U	1	HLA-A*02:01	ALWGFFPVL	2.88
2AV1	1	HLA-A*02:01	LLFGYPVYV	1.95
2UWE	1	HLA-A*02:01	ALWGFFPVL	2.40
2JCC	1	HLA-A*02:01	ALWGFFPVL	2.50
3OX8	1	HLA-A*02:03	FLPSDFFPSV	2,16
3OXR	1	HLA-A*02:06	FLPSDFFPSV	1,7
3OXS	1	HLA-A*02:07	FLPSDFFPSV	1,75
2XPG	1	HLA-A*03:01	KLIETYFSK	2.60
3RL1	1	HLA-A*03:01	AIFQSSMTK	2.00
3RL2	1	HLA-A*03:01	QVPLRPMTYK	2.39
1Q94	1	HLA-A*11:01	AIFQSSMTK	2.40
1QVO	1	HLA-A*11:01	QVPLRPMTYK	2.22
1X7Q	1	HLA-A*11:01	KTFPPTEPK	1.45
2HN7	1	HLA-A*11:01	AIMPARFYPK	1.60
2BCK	1	HLA-A*24:02	VYGFVRACL	2.80
3I6L	1	HLA-A*24:02	QFKDNVILL	2.40
3NFN	1	HLA-A*24:02	RYPLTFGWCF	2.39
3QZW	1	HLA-A*24:02	RYPLTFGWCF	2.80
4F7M	1	HLA-A*24:02	LYASPQLEGF	2.40
4F7P	1	HLA-A*24:02	FYRYGFVANF	1.90
4F7T	1	HLA-A*24:02	RYGFVANF	1.70
1HSB	1	HLA-A*68:01	AVA	1.90
1TMC	1	HLA-A*68:01	EVAPPEYHRK	2.30
2HLA	1	HLA-A*68:01	0	2.60
3VCL	1	HLA-B*07:02	RPHERNGFTVL	1.70
1AGB	1	HLA-B*08:01	GGRKKYKL	2.20
1AGC	1	HLA-B*08:01	GGKKKYQL	2.10
1AGD	1	HLA-B*08:01	GGKKKYKL	2.05
1AGE	1	HLA-B*08:01	GGKKKYRL	2.30
1AGF	1	HLA-B*08:01	GGKKRYKL	2.20

1M05	1	HLA-B*08:01	FLRGRAYGL	1.90
1MI5	1	HLA-B*08:01	FLRGRAYGL	2.50
3FFC	1	HLA-B*08:01	FLRGRAYGL	2.80
3SJV	1	HLA-B*08:01	FLRGRAYGL	3.10
3SKM	1	HLA-B*08:01	FLRGRAYVL	1.80
3SPV	1	HLA-B*08:01	RAKFKQLL	1.30
3SKO	1	HLA-B*08:01	FLRGRAYGL	1.60
3BXN	1	HLA-B*14:02	IRAAPPPLF	1.86
1XR8	1	HLA-B*15:01	LEKARGSTY	2.30
1XR9	1	HLA-B*15:01	ILGPPGSVY	1.79
3C9N	1	HLA-B*15:01	VQQESSFVM	1.87
1HSA	1	HLA-B*27:05	ARAAAAAAA	2.10
1JGE	1	HLA-B*27:05	GRFAAAIAK	2.10
1OGT	1	HLA-B*27:05	RRKWRRWHL	1.47
1UXS	1	HLA-B*27:05	RRRWRLTV	1.55
1W0V	1	HLA-B*27:05	RRLPIFSRL	2.27
2A83	1	HLA-B*27:05	RRRWHRWRL	1.40
2BSR	1	HLA-B*27:05	RRYDLIEL	2.30
2BSS	1	HLA-B*27:05	KRWIILGLNK	2.00
2BST	1	HLA-B*27:05	SRYWAIRTR	2.10
3B6S	1	HLA-B*27:05	RRKWRRWHL	1.80
3BP4	1	HLA-B*27:05	IRAAPPPLF	1.85
3DTX	1	HLA-B*27:05	RRKWRRWHL	2.10
3LV3	1	HLA-B*27:05	SRRWRRWNR	1.94
4G8G	1	HLA-B*27:05	KRWIILGLNK	2.4
4G8I	1	HLA-B*27:05	KRWIIMGLNK	1.6
4G9D	1	HLA-B*27:05	KRWIILGLNK	1.6
4G9F	1	HLA-B*27:05	KRWIIMGLNK	1.9
1JGD	1	HLA-B*27:09	RRLLRGHNQY	1.90
1K5N	1	HLA-B*27:09	GRFAAAIAK	1.09
1OF2	1	HLA-B*27:09	RRKWRRWHL	2.20
1UXW	1	HLA-B*27:09	RRRWRLTV	1.71
1W0W	1	HLA-B*27:09	RRLPIFSRL	2.11
3B3I	1	HLA-B*27:09	RRKWRRWHL	1.86
3BP7	1	HLA-B*27:09	IRAAPPPLF	1.80
3CZF	1	HLA-B*27:09	RRRWHRWRL	1.20
3D18	1	HLA-B*27:09	RRRWRLTL	1.74
3HCV	1	HLA-B*27:09	RRKWRRWHL	1.95
1A1N	1	HLA-B*35:01	VPLRPMTY	2.00

1A9B	1	HLA-B*35:01	LPPLDITPY	3.20
1A9E	1	HLA-B*35:01	LPPLDITPY	2,5
1CG9	1	HLA-B*35:01	LPPLDITPY	2.70
1XH3	1	HLA-B*35:01	LPAVVGLSPGEQEY	1.48
1ZHK	1	HLA-B*35:01	LPEPLPQGQLTAY	1.60
1ZSD	1	HLA-B*35:01	EPLPQGQLTAY	1.70
2AXG	1	HLA-B*35:01	APQPAPENAY	2.00
2CIK	1	HLA-B*35:01	KPIVVLHGY	1,75
2FYI	1	HLA-B*35:01	HPVGEADYFEY	1.50
2H6P	1	HLA-B*35:01	KPIVVLHGY	1.90
2NX5	1	HLA-B*35:01	EPLPQGQLTAY	2.70
3LKN	1	HLA-B*35:01	LPFERATIM	2.00
3LKO	1	HLA-B*35:01	LPFDRTTIM	1.80
3LKP	1	HLA-B*35:01	LPFDKSTIM	1.80
3LKQ	1	HLA-B*35:01	LPFDKTTIM	1.80
3LKR	1	HLA-B*35:01	LPFERATVM	2.00
3LKS	1	HLA-B*35:01	LPFEKSTVM	1.90
3MV7	1	HLA-B*35:01	HPVGEADYFEY	2.00
3MV8	1	HLA-B*35:01	HPVGEADYFEY	2.10
3MV9	1	HLA-B*35:01	HPVGEADYFEY	2.70
1ZHL	1	HLA-B*35:08	LPEPLPQGQLTAY	1.50
2AK4	1	HLA-B*35:08	LPEPLPQGQLTAY	2.50
2AXF	1	HLA-B*35:08	APQPAPENAY	1.80
2FZ3	1	HLA-B*35:08	HPVGEADYFEY	1.90
2NW3	1	HLA-B*35:08	EPLPQGQLTAY	1.70
3BW9	1	HLA-B*35:08	CPSQEPMSIYVY	1.75
3BWA	1	HLA-B*35:08	FPTKDVAL	1.30
3VFS	1	HLA-B*35:08	LPEPAPQGQLTAY	1.85
3VFT	1	HLA-B*35:08	LPEPLAQGQLTAY	1.95
3VFU	1	HLA-B*35:08	LPEPLPAGQLTAY	1.65
3VFW	1	HLA-B*35:08	LPEPLPQGALTAY	1.55
3VFW	1	HLA-B*35:08	LPEPLPQGQATAY	2.30
4JRX	1	HLA-B*35:08	LPEPLPQGQLTAY	2,3
4JRY	1	HLA-B*35:08	LPEPLPQGQLTAY	2,8
3KWW	1	HLA-B*35:08	LPEPLPQGQLTAY	2.18
3KXF	1	HLA-B*35:08	LPEPLPQGQLTAY	3.10
3LN4	1	HLA-B*41:03	AEMYGSVTEHSPSPSL	1.30
3LN5	1	HLA-B*41:04	HEEAVSVDRVL	1.90
1M6O	1	HLA-B*44:02	EEFGRAFSF	1.60

3DX6	1	HLA-B*44:02	EENLLDFVRF	1.70
3KPL	1	HLA-B*44:02	EEYLQAFTY	1.96
3KPM	1	HLA-B*44:02	EEYLKAWTF	1.60
3L3D	1	HLA-B*44:02	EEAGRAFSF	1.80
3L3G	1	HLA-B*44:02	EEFGAAFSF	2.10
3L3I	1	HLA-B*44:02	EEFGRAASF	1.70
3L3J	1	HLA-B*44:02	EEAGAAFSF	2.40
3L3K	1	HLA-B*44:02	EEFGAAASF	2.60
1N2R	1	HLA-B*44:03	EEFGRAFSF	1.70
1SYS	1	HLA-B*44:03	EEPTVIKKY	2.40
3DX7	1	HLA-B*44:03	EENLLDFVRF	1.60
3KPN	1	HLA-B*44:03	EEYLQAFTY	2.00
3KPO	1	HLA-B*44:03	EEYLKAWTF	2.30
1SYV	1	HLA-B*44:05	EEFGRAFSF	1.70
3DX8	1	HLA-B*44:05	EENLLDFVRF	2.10
3DXA	1	HLA-B*44:05	EENLLDFVRF	3.50
3KPP	1	HLA-B*44:05	EEYLQAFTY	1.90
3KPQ	1	HLA-B*44:05	EEYLKAWTF	1.84
3KPR	1	HLA-B*44:05	EEYLKAWTF	2.60
3KPS	1	HLA-B*44:05	EEYLQAFTY	2.70
1E27	1	HLA-B*51:01	LPPVVAKEI	2.20
1E28	1	HLA-B*51:01	TAFTIPSI	3.00
3W39	1	HLA-B*52:01	TAFTIPSI	3,1
1A1M	1	HLA-B*53:01	TPYDINQML	2.30
1A1O	1	HLA-B*53:01	KPIVQYDNF	2.30
2RFX	1	HLA-B*57:01	LSSPVTKSF	2.50
2YPK	1	HLA-B*57:01	KAFSPEVIPMF	1.95
3UPR	1	HLA-B*57:01	HSITYLLPV	2.00
3VH8	1	HLA-B*57:01	LSSPVTKSF	1.80
3VRI	1	HLA-B*57:01	RVAQLEQVYI	1.60
3VRJ	1	HLA-B*57:01	LTTKLTNTNI	1.90
2BVO	1	HLA-B*57:03	KAFSPEVIPMF	1.65
2BVP	1	HLA-B*57:03	ISPRTLDAW	1.35
2BVQ	1	HLA-B*57:03	KAFSPEVIP	2.00
2HJK	1	HLA-B*57:03	KGFNPEVIPMF	1.85
2HJL	1	HLA-B*57:03	KAFNPEIIPMF	1.50
2YPL	1	HLA-B*57:03	KAFSPEVIPMF	2.40
1EFX	1	HLA-C*03:04	GAVDPLLAL	3.00
1IM9	1	HLA-C*04:01	QYDDAVYKL	2.80

1QQD	1	HLA-C*04:01	QYDDAVYKL	2.70
3LQZ	2	HLA-DPA1*01:03/DP B1*02:01	RKFHYLPFLPST	3.25
4D8P	2	HLA-DQ	PQPEQPEQPFQPP	3.05
1JK8	2	HLA-DQ8	SHLVEALYLVCGERG	2.40
2NNA	2	HLA-DQ8	SGEGSFQPSQENP	2.10
1UVQ	2	HLA-DQA1*01:02/D QB1*06:02	MNLPSTKVSAAV	1.80
4GG6	2	HLA-DQA1*03:01/D QB1*03:02	QQYPSGQGSFQPSQQNPQ	3.20
1S9V	2	HLA-DQA1*05:01/D QB1*02:01	LQFPQPQLPY	2.22
1FYT	2	HLA-DR1	PKYVKQNTLKLAT	2.60
1JWM	2	HLA-DR1	PKYVKQNTLKLAT	2.70
1JWS	2	HLA-DR1	PKYVKQNTLKLAT	2.60
1JWU	2	HLA-DR1	PKYVKQNTLKLAT	2.30
1KG0	2	HLA-DR1	PKYVKQNTLKLAT	2.65
1KLG	2	HLA-DR1	GELIGILNAAKVPAD	2.40
1KLU	2	HLA-DR1	GELIGTLNAAKVPAD	1.93
1L05	2	HLA-DR1	PKYVKQNTLKLAT	3.20
1PYW	2	HLA-DR1	FVKQNAAL	2.10
1R5I	2	HLA-DR1	PKYVKQNTLKLAT	2.60
1SEB	2	HLA-DR1	AAAAAAAAAAAAA	2.70
1T5W	2	HLA-DR1	AAYSDQATPLLSR	2.40
1T5X	2	HLA-DR1	AAYSDQATPLLSR	2.50
2ICW	2	HLA-DR1	PKYVKQNTLKLAT	2.41
2OJE	2	HLA-DR1	PKYVKQNTLKLAT	3.00
3PDO	2	HLA-DR1	KPVSKMRMATPLLMQAL PM	1.95
3PGC	2	HLA-DR1	KMRMATPLLMQALPM	2.66
3PGD	2	HLA-DR1	KMRMATPLLMQALPM	2.72
3S4S	2	HLA-DR1	PKYVKQNTLKLAT	2.40
3S5L	2	HLA-DR1	PKYVKQNTLKLAT	2.10
4AEN	2	HLA-DR1	MKMRMATPLLMQALPM	2.20
4AH2	2	HLA-DR1	MKMRMATPLLMQAL	2.36
1BX2	2	HLA-DR2	ENPVVHFFKNIVTPR	2.60
1A6A	2	HLA-DR3	LPKPPKPVSKMRMATPLL MQALPM	2.75
1MUJ	2	HLA-DR3	GSHSRGLPKPPKPVSKMR MATPLLMQALPMGSGSGS	2.15
1J8H	2	HLA-DR4	PKYVKQNTLKLAT	2.40

3QXD	2	HLA-DRA*01:01 F54C mutasyonu/DR B1*01:0	PVSKMRMATPLLMQ	2.30
1AQD	2	HLA- DRA*01:01/DR B1*01:01	VGSDWRFLRGYHQA	2.45
1HXY	2	HLA- DRA*01:01/DR B1*01:01	PKYVKQNTLKLAT	2.60
1SJE	2	HLA- DRA*01:01/DR B1*01:01	PEVIPMFSALEGATP	2.45
1SJH	2	HLA- DRA*01:01/DR B1*01:01	PEVIPMFSALESEG	2.25
2FSE	2	HLA- DRA*01:01/DR B1*01:01	GIAGFKGEQGPKEGP	3.10
2G9H	2	HLA- DRA*01:01/DR B1*01:01	PKYVKQNTLKLAT	2.00
2IPK	2	HLA- DRA*01:01/DR B1*01:01	PKYVKQNTLKLAT	2.30
3L6F	2	HLA- DRA*01:01/DR B1*01:01	APPAYEKLSAEQSPP	2.10
3QXA	2	HLA- DRA*01:01/DR B1*01:01	PVSKMRMATPLLMQ	2.71
2SEB	2	HLA- DRA*01:01/DR B1*04:01	QYMRADQAAGGLR	2.50
3O6F	2	HLA- DRA*01:01/DR B1*04:01	FSWGAEGQRPGFG	2.80
2WBJ	2	HLA- DRA*01:01/DR B1*15:01	DFARVHFISALHGSG	3.00
1FV1	2	HLA- DRA*01:01/DR B5*01:01	NPVVHFFKNIVTPRTPPPS Q	1.90
1H15	2	HLA- DRA*01:01/DR B5*01:01	TGGVYHFVKKHVES	3.10
1HQR	2	HLA- DRA*01:01/DR B5*01:01	VHFFKNIVTPRTP	3.20
1ZGL	2	HLA- DRA*01:01/DR B5*01:01	VHFFKNIVTPRTP	2.80
1DLH	2	HLA- DRB1*01:01	PKYVKQNTLKLAT	2.80
2IAM	2	HLA- DRB1*01:01	GELIGILNAAKVPAD	2.80

2IAN	2	HLA- DRB1*01:01	GELIGTLNAAKVPAD	2.80
4FQX	2	HLA- DRB1*01:01	GKQNCLKLATK	2.60
4GBX	2	HLA- DRB1*01:01	GKQNCLKLATK	3.00
1D5M	2	HLA- DRB1*04:01	XRAMXSX	2.00
1D5X	2	HLA- DRB1*04:01	XRXXX	2.45
1D5Z	2	HLA- DRB1*04:01	XRAXSX	2.00
1D6E	2	HLA- DRB1*04:01	XRXTMASX	2.45
1YMM	2	HLA- DRB1*15:01	ENPVVHFFKNIVTPR	3.50
2Q6W	2	HLA- DRB3*01:01	AWRSDEALPLGS	2.25
3C5J	2	HLA- DRB3*03:01	QVIILNHPGQISA	1.80
4H1L	2	HLA- DRB3*03:01	HIRCNIPKRI	3.30
1MHE	1	HLA-E*01:01	VMAPRTVLL	2.85
2ESV	1	HLA-E*01:01	VMAPRTLIL	2.60
3BZE	1	HLA-E*01:01	VMAPRTLFL	2.50
3BZF	1	HLA-E*01:01	VMAPRALLL	2.50
3CDG	1	HLA-E*01:01	VMAPRTLFL	3.40
3CII	1	HLA-E*01:01	VMAPRTLFL	4.41
1KPR	1	HLA-E*01:03	VMAPRTVLL	2.80
1KTL	1	HLA-E*01:03	VTAPRTLIL	3.10
2DYP	1	HLA-G*01:01	RIIPRHLQL	2.50
1YDP	1	HLA-G*01:01	RIIPRHLQL	1.90
2D31	1	HLA-G*01:01	RIIPRHLQL	3.20
3KYN	1	HLA-G*01:01	KGPPAALTL	2.40
3KYO	1	HLA-G*01:01	KLPAQFYIL	1.70

Ek B. Veri Setinde Bulunan Sınıf I Yapıları Arasındaki Alfa-Karbon RMSD Dağılımı

