T.C.
SAĞLIK BİLİMLERİ ÜNİVERSİTESİ
HAMİDİYE SAĞLIK BİLİMLERİ ENSTİTÜSÜ

SARS-CoV-2 MUTASYONLARININ PROTEİN YAPILARI ÜZERİNDEKİ ETKİLERİNİN İN-SİLİKO YÖNTEMLER İLE ARAŞTIRILMASI

ZEHRA YAREN DÖNMEZ

TEZ DANIŞMANI
DR. ÖĞR. ÜYESİ BETÜL AKÇEŞME

TIBBİ BİYOLOJİ ANA BİLİM DALI
TIBBİ BİYOLOJİ YÜKSEK LİSANS PROGRAMI

YÜKSEK LİSANS TEZİ
EYLÜL/2021

İTHAF

“Annem, İncilay Dönmez’e ithaf ediyorum”

TEŞEKKÜR

Yüksek lisans eğitimim süresince bilgilerini ve tecrübelerini aktaran Sağlık Bilimleri Üniversitesi Tıbbi Biyoloji Ana Bilim Dalı Başkanı saygıdeğer hocam Prof. Dr. Ülkan KILIÇ'a

Tez çalışmam ve yüksek lisans eğitimimin her aşamasında bana yol gösteren, bilgi ve tecrübelerini benimle paylaşarak bana her konuda yardımcı olan saygı değer hocam Dr. Öğr. Üyesi Betül AKÇEŞME'ye

Yüksek lisans eğitimim süresince bilgilerini ve tecrübelerini aktaran saygı değer hocam Doç. Dr. Ender M. COŞKUNPINAR' a,

Tez çalışmam boyunca engin tecrübelerini ve manevi desteklerini üzerimden eksik etmeyen saygı değer hocam Dr. Öğr. Üyesi Faruk Berat AKÇEŞME'ye,

Tez çalışmam boyunca yanımda olarak desteklerini hiç esirgemeyen çok sevgili çalışma arkadaşlarıma,

Son olarak hayatım boyunca benden desteklerini ve sevgilerini esirgemeyen aileme teşekkür ederim.

İÇİNDEKİLER

İTHAF iv
TEŞEKKÜR v
İÇİNDEKİLER vi
TABLOLAR LİSTESİ viii
ŞEKİLLER LİSTESİ ix
SİMGELER VE KISALTMALAR x
ÖZET xii
ABSTRACT xiv
1. GİRİŞ VE AMAÇ 1
2. GENEL BİLGİLER 4
2.1. KORONAVİRÜSLER 4
2.2 SARS-CoV-2 6
2.3 NÜKLEOKAPSİD PROTEİNİ 9
2.3.1 NÜKLEOKAPSİD PROTEİN YAPISI 9
2.4 MUTASYONLARIN PROTEİN YAPISINA ETKİSİ 14
3. GEREÇ VE YÖNTEM 16
3.1 MUTASYONLARIN PROTEİN YAPILARINA ETKİLERİNİN İN-SİLİKO OLARAK ARAŞTIRILMASI 16
3.2 MUTASYON ANALİZİNDE KULLANILAN SUNUCULAR 18
3.2.1 Yapı temelli sunucular 18
3.2.2 Sekans temelli sunucular 20
3.3 ÇALIŞMADA KULLANILAN DİZİLERİN ELDE EDİLMESİ 21
3.3.1 Nükleokapsid Protein Mutasyonlarının Tespit Edilmesi 21
3.4 MUTASYONUN PROTEİN FONKSİYONU ÜZERİNDEKİ ETKİLERİNİN ANALİZİ 22
3.4.1 NÜKLEOKAPSİD PROTEİNİNİN ÜÇ BOYUTLU YAPISININ ELDE EDİLMESİ 22
3.4.2 DİZİ TEMELLİ ANALİZ 23
3.4.3 YAPI TEMELLİ ANALİZ 23
4. BULGULAR 24
4.1 TESPİT EDİLEN YANLIŞ ANLAMLI MUTASYONLAR 24
4.2 SEKANS TEMELLİ ANALİZ SONUÇLARI 30

4.3 YAPI TEMELLİ ANALİZ SONUÇLARI 37
4.3.1 STABİLİZASYON ANALİZ SONUÇLARI 37
4.3.2 ESNEKLİK ANALİZ SONUÇLARI 43
4.4 ANAHTAR MUTASYONLARIN PROTEİN YAPISINA ETKİSİNİN SONUÇLARI 49
4.4.1 Kimyasal Bağ Ve Etkileşim Analiz Sonuçları 50
4.4.2 Entropi Enerji Değişim Analizi Sonuçları 54
4.4.3 Atomik Kararsızlık Analiz Sonuçları 58
4.4.4 Deformasyon Enerjilerinin Analiz Sonuçları 59
5. TARTIŞMA 60
6. SONUÇLAR VE ÖNERİLER 65
KAYNAKLAR 66

TABLOLAR LİSTESİ

Tablo 3.1: Çalışmada kullanılan sekansa ve yapıya dayalı sunucular 33

Tablo 4.1: Nükleokapsid proteinin N-kolunda tespit edilen mutasyonlar 40

Tablo 4.2: Nükleokapsid proteinin N-Terminal Domaininde tespit edilen mutasyonlar..... 41

Tablo 4.3: Nükleokapsid proteinin bağlayıcı bölgesinde tespit edilen mutasyonlar..... 42

Tablo 4.4: Nükleokapsid proteinin C-Terminal Domaininde tespit edilen mutasyonlar..... 43

Tablo 4.5: Nükleokapsid proteinin C-kuyruğunda tespit edilen mutasyonlar..... 44

Tablo 4.6: Farklı sunucular ile elde edilen nötr ve deleterious (zararlı) mutasyonların dağılımı.. 46

Tablo 4.7: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteinin N-Kolu bölgesine ait mutasyon sonuçları (N: Nötr, D: Deleterious (zararlı)) 46

Tablo 4.8: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteinin N-Teminal bölgesine ait mutasyon sonuçları (N: Nötr, D: Deleterious (zararlı))..... 48

Tablo 4.9: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteinin bağlayıcı bölgeye ait mutasyon sonuçları (N: Nötr, D: Deleterious (zararlı)) 49

Tablo 4.10: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteinin C-Terminal Domainine ait mutasyon sonuçları (N: Nötr, D: Deleterious (zararlı)) 50

Tablo 4.11: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteinin C-Kuyruğuna ait mutasyon sonuçları (N: Nötr, D: Deleterious (zararlı)) 51

Tablo 4.12: DynaMut ağ sunucu ile elde edilen N proteinin N-Kolu bölgesine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları 53

Tablo 4.13: DynaMut ağ sunucu ile elde edilen N proteinin N-Terminal Domainine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları 54

Tablo 4.14: DynaMut ağ sunucu ile elde edilen N proteinin bağlayıcı bölgesine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları 55

Tablo 4.15: DynaMut ağ sunucu ile elde edilen N proteinin C-Terminal Domainine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları 56

Tablo 4.16: DynaMut ağ sunucu ile elde edilen N proteinin C-Kuyruğuna ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları 57

Tablo 4.17: DynaMut ağ sunucu ile elde edilen N proteinin N-kolu bölgesine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları 59

Tablo 4.18: DynaMut ağ sunucu ile elde edilen N proteinin N-Termianl Domainine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları..... 60

Tablo 4.19: DynaMut ağ sunucu ile elde edilen N proteinin bağlayıcı bölgesine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları 61

Tablo 4.20: DynaMut ağ sunucu ile elde edilen N proteinin C-Termianl Domainine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları..... 62

Tablo 4.21: DynaMut ağ sunucu ile elde edilen N proteinin C-Kuyruğuna ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları 63

ŞEKİLLER LİSTESİ

Şekil 2.1: Koronavirüslerinin sınıflandırılması 19

Şekil 2.2: SARS-CoV-2 virüsünün elektron mikroskobu görüntüsü 20

Şekil 2.3: SARS-CoV-2 virüsünün gen organizasyonu ve yapısal proteinlerinin şematik gösterimi 21

Şekil 2.4: SARS-CoV-2 virüsünün patogenezi 23

Şekil 2.5: RNA bağlanmasını kolaylaştıran N proteinlerinin yapısal bileşenleri 25

Şekil 2.6: RNA bağlanma domaini ve etkileşimlerinde yer alan yan zincirler 26

Şekil 2.7: Nükleokapsid proteinin gen organizayonu ile N-terminal domain (PDB:6M3M) ve C-Terminal Domaininin (PDB:6YUN) 3 boyutlu yapısı 27

Şekil 2.8: SARS-CoV-2 C-Terminal Domaini için topoloji diyagramı 28

Şekil 3.1: Nükleokapsid proteinin 3 boyutlu yapısı 37

Şekil 4.1: Bağ ve etkileşim tiplerinin gösterildiği renkler 64

Şekil 4.2: M234I mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 65

Şekil 4.3: A220V mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 65

Şekil 4.4: S194L mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 66

Şekil 4.5: S235F mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 66

Şekil 4.6: D128N mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 67

Şekil 4.7: D3L mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 67

Şekil 4.8: T362I mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu 68

Şekil 4.9: M234I mutasyonunun entropi enerji değişimi analizi sonucu 69

Şekil 4.10: A220V mutasyonunun entropi enerji değişimi analizi sonucu 69

Şekil 4.11: S194L mutasyonunun entropi enerji değişimi analizi sonucu 70

Şekil 4.12: S235F mutasyonunun entropi enerji değişimi analizi sonucu 70

Şekil 4.13: D3L mutasyonunun entropi enerji değişimi analizi sonucu 71

Şekil 4.14: T362I mutasyonunun entropi enerji değişimi analizi sonucu 71

Şekil 4.15: D128N mutasyonunun entropi enerji değişimi analizi sonucu 72

Şekil 4.16: Atomik kararsızlık analiz sonucu 73

Şekil 4.17: Deformasyon enerjilerinin analiz sonucu 74

SİMGELER VE KISALTMALAR

+ssRNA: Single Strand Ribonucleic Acid, Pozitif Tek Sarmallı Ribonükleik Asit

3B: 3 Boyutlu

ACE2: Angiotensin-Converting Enzyme 2, Anjiyotensin Dönüştürücü Enzimi

COVID-19: Corona Vırus Diseases-19, Korona Virüs Hastalığı-19

CTD: C-Terminal Domaini

DMV: Double-membrane vesicles, Çift Membranlı Vezikül

DSÖ: Dünya Sağlık Örgütü

E: Envelope, Zarf

ENCoM: Elastic Network Contact Model

ER: Endoplazmik Retikulum

ERGIC: Endoplasmic Reticulum-Golgi İntermediate Compartment**,** Endoplazmik Retikulum-Golgi Ara Bölmesi

GISAID: Global Initiative on Sharing All Influenza Data

gRNA: Genomik RNA

HCoV-229E: Human coronavirus 229E

HCoV-NL63: Human coronavirus NL63

HCoV-OC43: Human coronavirus OC43

IDR: Intrinsically Disordered Regions, Düzensiz Bölge

LKR: Linker, Bağlayıcı Bölge

M: Membran

MAPP: Multivarivate Analysis of Protein Polymorphism

mCSM: Mutation Cutoff Scanning Matrix

MERS-CoV: Middle East Respiratory Syndrome**,** Orta Doğu Solunum Sendromu

mRNA: Mesajcı RNA

N: Nükleokapsid

NSP: Nonstructural Protein, Yapısal Olmayan Proteinler

NsSNP: Non-synonymous Single Nucleotide Polymorphisms**,** Sinonim Olmayan Tek Nükleotid Polimorfizm

NTD: N-Terimal Domain

PANTHER: Protein Analysis Through Evolutionary Relationships

PCR: Polimerase Chain Reaction, Polimeraz Zincir Reaksiyonu

PhD-SNP: Predictor of Human Deleterious Single Nucleotide Polymorphisms

PolyPhen-2: Polymorphism Phenotyping version 2

PredictSNP: Predict Single Nucleotide Polymorphism

RNA: Ribonükleik asit

RNP: Ribonükleoprotein

RTC: Replication Transcription Complex**,** Replikasyon Transkripsiyon Kompleksi

RT-PCR: Real Time Polimerase Chain Reaction, Gerçek Zamanlı Polimeraz Zincir Reaksiyonu

S: Spike

SARS-CoV: Severe Acute Respiratory Syndrome Coronavirus, Şiddetli Akut Solunum Sendrom Koronavirüsü

SARS-CoV-2: Severe Acute Respiratory Syndrome Coronavirus 2**,** Şiddetli Akut Solunum Sendrom Koronavirüsü 2

SDM: Site Directed Mutator

SIFT: Sorting İntolerant From Tolerant

SNAP: Screening for non-Acceptable Polymorphisms

SR: Serin-Arginin

SVM: Support Vector Machines**,** Destek Vektör Makineleri

α: Alfa

β: Beta

ΔG: Gibbs Serbest Enerjisi

SARS-CoV-2 MUTASYONLARININ PROTEİN YAPILARI ÜZERİNDEKİ ETKİLERİNİN İN-SİLİKO YÖNTEMLER İLE ARAŞTIRILMASI

ÖZET

Amaç: SARS-CoV-2 virüsünde meydana gelen yanlış anlamlı mutasyonlar proteinlerin yapısında değişikliğe sebep olabilir. Meydana gelen bu yapısal değişikliklerin de ilaç etkileşimi, aşı ve ilaç keşfi ile birlikte hastalığın seyri için önemli olduğu bilinmektedir. Nükleokapsid proteini ilaç ve aşı için önemli bir hedeftir. Bu tezin temel amacı, SARS-CoV-2 nükleokapsid proteininde meydana gelen yanlış anlamlı mutasyoların tespit edilmesi ve bu mutasyonların protein yapısı üzerine etkisinin in-siliko yaklaşımlar kullanılarak ortaya koyulmasıdır.

Gereç ve Yöntemler: Çalışmamız da GISAID EpiCoV veri tabanından indirilen Türk populasyonuna ait 2286 örnekte N proteine ait 161 adet yanlış anlamlı mutasyon tespit edildi. Tespit edilen 161 adet yanlış anlamlı mutasyon, sekansa ve yapıya dayalı yöntemler kullanılarak analiz edildi. Sekansa dayalı analiz için; PROVEAN, PredictSNP ve SIFT ağ sunucuları ile MAPP ve SNAP ağ araçları kullanıldı. Yapıya dayalı analiz için ise; DynaMut, mCSM, SDM ağ sunucuları ile DUET ve ENCoM adlı sunucular kullanıldı.

Bulgular: Sekansa dayalı analiz sonucunda; PROVEAN ve PredictSNP ağ sunucularında ortak olarak 39 adet mutasyon deleterious (zararlı), 16 adet mutasyon nötr olarak tespit edildi DynaMut ağ sunucusunda yapılan protein stabilizasyon analizi sonucunda 103 adet mutasyonun proteinin 3 boyutlu yapısının stabilizasyonu arttırdığı ve 58 adet mutasyonun ise protein yapısındaki stabilitizasyonu azalttığı tespit edildi. DynaMut ağ sunucusunda yapılan protein entropi analizi sonucunda; 77 adet yanlış anlamlı mutasyonun proteinin 3 boyutlu yapısının esnekliğini arttırdığı ve 84 adet yanlış anlamlı mutasyonun proteinin esnekliğini azalttığı tespit edildi. Kimyasal bağ ve etkileşim analizi sonucunda M234I, A220V, S194L, S235F, D128N ve T362I mutasyonlarının etkileşim değişimlerine yol açtığı D3L mutasyonunun değişimlerine yol açmadığı tespit edildi. Mutasyonlar sonucu N proteininin 3 boyutlu yapsıında deformasyon enerjisinde ve atomik kararsızlıkta bir değişiklik tespit edilmemiştir.

Sonuç: In siliko yöntemler kullanılarak elde edilen bulgularımız, SARS-CoV-2 N proteininde meydana gelen bazı yanlış anlamlı mutasyonların proteinin yapısını ve bağlanma afinitesini değiştirebileceğini ortaya koymaktadır. SARS-CoV-2 virüsüne karşı aşı ve ilaç geliştirme araştırmalarına katkısı olacağı düşünülen çalışmamızın in vitro ve in vivo yöntemler ile doğrulanması gerekmektedir.

Anahtar kelimeler: Nükleokapsid proteini, SARS-CoV-2, Yanlış anlamlı mutasyon, Protein stabilitesi

THE INVESTIGATION OF THE EFFECT OF THE SARS-CoV-2 MUTATIONS ON PROTEIN STRUCTURES BY USING IN-SILICO APPROACHES

ABSTRACT

Aim: Missense mutations in the SARS-CoV-2 virus may cause changes in the structure of proteins. It is known that these structural changes are important for drug interaction, vaccine and drug discovery, and the course of the disease. The nucleocapsid protein is an important target for drugs and vaccines. The main purpose of this thesis is to detect missense mutations in the SARS-CoV-2 nucleocapsid protein and to reveal the effects of these mutations on protein structure using in-silico approaches.

Material and Methods: In our study, 161 missense mutations of the N protein were found in 2286 samples of the Turkish population which were downloaded from the GISAID EpiCoV database. The detected 161 missense mutations were analyzed using sequence and structure-based methods. For sequence-based analysis; PROVEAN, PredictSNP and SIFT web servers and MAPP and SNAP network tools were used. For structure-based analysis; DynaMut, mCSM, SDM web servers and DUET and ENCoM servers were used.

Results: As a result of the sequence-based analysis, by PROVEAN and PredictSNP web server, 39 mutations were found to be deleterious while 16 mutations were neutral. As a result of protein stability analysis performed on the DynaMut web server; it was found that 103 missense mutations had increased the stability of the 3D structure of the protein whereas 58 missense mutations had decreased its stability. As a result of the protein stabilization analysis performed on the DynaMut web server, it was found that 77 mutations had increased the flexibility of the 3D structure of the protein and 84 mutations had decreased the flexibility of the protein structure. As a result of the chemical bond and interaction analysis, it was determined that M234I, A220V, S194L, S235F, D128N and T362I mutations caused interaction changes but D3L mutation did not cause any interaction changes. There is no change in the deformation energy and atomic instability in the 3-dimensional structure of the N protein as a result of mutations.

Conclusion: Our results obtained by using in silico approaches revealed that some missense mutations in the SARS-CoV-2 N protein can change the structure of the protein

and its binding affinity. Our study, which is thought to contribute to vaccine and drug development research against SARS-CoV-2 virus, needs to be validated with in vitro and in vivo methods.

Keywords: Nucleocapsid protein, SARS-CoV-2, Missense mutations, Protein stability

1. GİRİŞ VE AMAÇ

Koronavirüsler, büyük (80-220 nm çapında), zarflı, Pozitif Tek Sarmallı Ribonükleik Asit (+ssRNA) virüsleridir. En uzun koronavirüs genomu 33,5 kb'ye kadar olabilmektedir. Şiddetli Akut Solunum Sendrom Koronavirüsü 2 (SARS-CoV-2) 'nin genomu Şiddetli Akut Solunum Sendrom Koronavirüsü (SARS-CoV) genomu ile %80 homolojiye sahiptir. SARS-CoV-2, Beta (β) koronavirüs alt cinsine aittir (1). Bu virüslerin önemli bir özelliği, genetik varyasyonlara duyarlılıktaki değişkenliğidir (2).

Çin'in Wuhan şehrinde, 2019 yılının sonunda SARS-CoV-2 olarak adlandırılan yeni bir koronavirüs ortaya çıkmıştır ve viral pnömoni salgınına neden olmuştur. Koronavirüs Hastalığı 2019 (COVID-19) olarak da bilinen bu yeni koronavirüs hastalığı, tüm dünyaya hızla yayılmıştır. Dünya Sağlık Örgütü (DSÖ) 11 Mart'ta, COVID-19'u pandemi olarak tüm dünyaya ilan etmiştir (3). SARS-CoV-2 en yaygın olarak solunum yoluyla bulaşmaktadır (1). COVID-19'un semptomları arasında; yüksek ateş, öksürük, nefes darlığı, yorgunluk, kas ve baş ağrısı, koku ve tat kaybı yer almaktadır (2)

SARS-CoV-2'nin genomu yaklaşık 30.000 nükleotitten oluşmaktadır. Genomik RNA'sı (gRNA), 14 Açık Okuma Çerçevesini (ORF) içermektedir. ORF1a ve ORF1b, genomun üçte ikisini kapsamaktadır ve sırasıyla Poliprotein 1a (PP1a) ve Poliprotein 1ab (PP1ab) çevrilir. Bu poliproteinler viral proteazlar tarafından Yapısal Olmayan Proteinler (NSP), Nsp1-Nsp16'ya transkribe edilir. Bazı NSP'ler, birkaç konakçı faktörle birlikte, Çift Membranlı Vezikül (DMV)'ler içinde bir Replikasyon Transkripsiyon Kompleksi (RTC) oluşturmaktadır. RTC'ler, viral genom replikasyonu ve transkripsiyonunun merkezidir. Genomun geri kalan üçte biri, dört ana yapısal proteini kodlayan ORF'lere sahiptir. ORF'ler, 4 yapısal protein ve bazı yardımcı proteinleri üretir. Bunlar; Spike (S), Membran (M), Zarf (E), Nükleokapsid (N) ve bazı yardımcı proteinler (ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF9 ve ORF10)'dir (4,5).

Diğer RNA virüslerine benzer şekilde, SARS-CoV-2'deki genetik çeşitlilik, virüsün hayatta kalması ve patogenezi için kritik öneme sahiptir. RNA virüslerindeki en önemli evrim mekanizmalarından biri sürekli mutasyona uğramasıdır. SARS-CoV-2'nin genomik varyasyonlarını saptamak için şimdiye kadar çeşitli çalışmalar yapılmıştır. Bu çalışmalar sonucunda, SARS-CoV-2 virüslerinde genetik varyasyonlar tespit edilmiştir.

Yapılan çalışmalara göre, yanlış anlamlı ve nötral mutasyonlar, SARS-CoV-2 genomu boyunca en sık görülen mutasyon tipleridir. Viral replikasyon, bulaşmaya karşı uyarılmış bağışıklık tepkilerinde daha çok mutasyona uğrayan bölgelerin rolü nedeniyle, bu mutasyonların SARS-CoV-2 salgınındaki etkilerini belirlemek için daha fazla araştırma yapılması gerekmektedir (6, 9).

Nükleokapsid proteini, koronavirüsler için önemli bir yapısal proteindir. Koronavirüslerde bol miktarda bulunur ve yüksek oranda immünojeniktir (2, 66). N proteinlerinin; konak hücreye girme, viral RNA genomuna bağlanma ve Ribonükleoprotein (RNP) çekirdeğini oluşturma fonksiyonları vardır. SARS-CoV-2 N proteini, %90 dizi özdeşliği ile SARS-CoV N proteini ile yüksek homoloji paylaşır (2). N proteini, RNA paketlenmesine ve virüs partikül salınımına katılan koronavirüs için önemli bir antijendir. N protein, aşı geliştirme ve serolojik analizlerde de sıklıkla kullanılmaktadır (5,10). SARS-CoV-2 N proteini, Serin-Arginin (SR) açısından zengin bir alan olan ve oligomerizasyonu sağlayan merkez bağlayıcı bölge (LKR), RNA bağlanma alanı içeren N-Terminal Domaini (NTD) ve C-Terminal Domaini (CTD)'den oluşur (5, 11, 12). SARS-CoV-2 N proteinlerinin moleküler özelliklerini konu edinen birçok çalışma yapılmasına rağmen N proteinin özellikleri henüz tam olarak bilinmemektedir.

RNA virüsünün genomu, virüs yayıldıkça kolayca mutasyonlar oluşturabilmektedir. Ayrıca koronavirüsler, diğer RNA virüslerine kıyasla yüksek mutasyon oranlarına sahiptir. Bu durum virüslerin adaptasyon süreçleri ile ilişkilendirilmiştir (13, 14, 15). SARS-CoV-2'deki mutasyonların COVID-19 hastalığının şiddeti, bulaşma ve virüsün stabilitesi üzerinde önemli etkileri olduğu düşünülmektedir (16). Viral mutasyonların biyolojik karakterizasyonu, viral ilaç direncini, bağışıklık kaçışını ve patogenezle ilgili mekanizmaları değerlendirmek için değerli bilgiler sağlamaktadır (17). SARS-CoV-2'de sürekli olarak yeni mutasyonların ortaya çıkması, ilaç ve aşı tasarımlarının gelişimi için en büyük zorluktur. N proteininin viral yaşam döngüsünün birçok önemli işlevine dahil olması onu COVID-19'da aday bir ilaç ve aşı hedefi haline getirmektedir (18, 19). Global Initiative on Sharing All Influenza Data (GISAID) veri bankasında 29/06/2021 tarihinde, 2,132,114 adet SARS-CoV-2 genom sekansı bulunmaktadır (20). SARS-CoV-2'nin dizilenen tüm genom verilerinin sayılarındaki artış, virüsün genomunda oluşan mutasyonların tespitinde ve bu

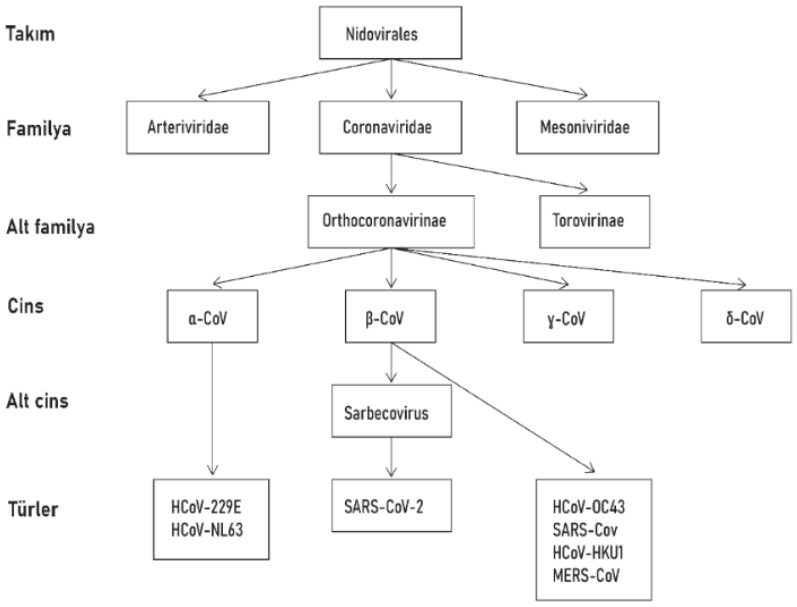
mutasyonların protein yapısı üzerindeki etkilerinin araştırılmasına imkan sağlamıştır. Nokta mutasyonunun proteinin stabilitesini nasıl değiştirdiğini tahmin etmek, farmasötik ilaç ve aşı tasarımı girişimlerine rehberlik etmektedir. Fiziksel proteinler ile ıslak laboratuvar çalışmalar yapmak, zaman ve finansal kaynaklar nedeniyle engelleyicidir (21). Ancak protein yapısı üzerindeki bir mutasyonun etkilerini tahmin etmeye yönelik biyoinformatik çalışmalar, yanlış anlamlı mutasyonların protein fonksiyonu ve yapısı üzerindeki etkilerini kolaylıkla ölçebilmekte ve zaman ve finansal kaynaklar içinde sorun teşkil etmemektedir (15,21,22).

Aşı ve antiviral ilaç tasarımı için N proteine ait mutasyonların protein stabilitesine ve proteine bağlanma etkisini belirleyen kesin sonuçlar bulunmamaktadır. Bu tezin temel amacı, in-siliko yaklaşımlarla SARS-CoV-2 nükleokapsid proteininde meydana gelen yanlış anlamlı mutasyoların tespit edilmesi ve bu mutasyonlarin protein yapısı üzerine etkisini ortaya koymaktir. Elde edilen sonuçların SARS-CoV-2 virüsüsün bulaşması, teşhisi ve COVID-19 tedavisi üzerinde etkileri olabileceğini düşünülmektedir. Bulgularımız SARS-CoV-2 virüsüne karşı ilaç aşı geliştirme sürecine yardımcı olacaktır.

2. GENEL BİLGİLER

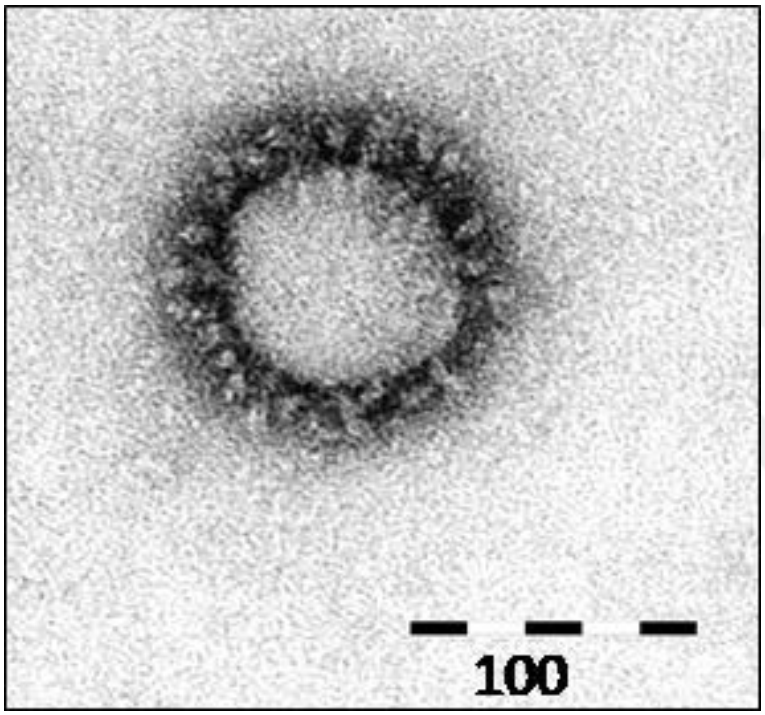
2.1. KORONAVİRÜSLER

Koronavirüsler, Nidovirales takımında, Coronaviridae ailesinde ve Orthocoronavirinae alt ailesinde bulunmaktadır. Orthococonavirinae alt ailesi 4 cins virüs içerir; Alphakoronavirus, Betacoronavirus, Gammacoronavirus ve Deltacoronavirus'tür (23). SARS-CoV-2, β-koronavirus alt cinsine aittir (Şekil 2.1) (5). İlk olarak Human coronavirus 229E (HCoV-229E) ve Human coronavirus OC43 (HCoV-OC43), genellikle soğuk algınlığının nedenlerinden sadece biri olarak görülmekteydi. İlerleyen yıllarda 2002'deki SARS-CoV salgını ve daha sonra 2012'de Orta Doğu Solunum Sendromu (Middle East respiratory syndrome-MERS-CoV) salgını olmuştur. 31 Aralık 2019'da Çin'in Hubei Eyaleti, Wuhan şehrindeki etiyolojisi bilinmeyen pnömoni vakaları DSÖ bildirilmeye başlanmıştır. 7 Ocak 2020'de bu vakalarının sebebinin yeni bir koronavirüs olduğu ilan edilmiştir. 11 Şubat 2020'de DSÖ, yeni koronavirüs kaynaklı pnömoniyi COVID-19 olarak adlandırmıştır ve Uluslararası Virüs Sınıflandırma Komisyonu, virüsün adını SARS-CoV-2 olarak adlandırıldığını duyurmuştur (23). Virüse bu adın verilme sebebi SARS-CoV-2 virüsünün SARS-CoV virüs genomuna %80 benzerlikte olmasıdır (5). SARS-CoV-2'den önce, SARS-CoV ve MERS-CoV dahil olmak üzere altı koronavirüsün insanlarda hastalığa neden olduğu bilinmekteydi (24).



Şekil 2.1: Koronavirüslerinin sınıflandırılması (23)

Virüsün morfolojisi elektron mikrograf yöntemi ile incelenmiştir ve güneş benzeri küresel bir yapıya benzediği bulunmuştur (Resim 2.2) (24). Koronavirüsler ssRNA virüsleridir. Çapının yaklaşık olarak 60-140 nanometre (nm) arasında, sivri uçlarının ise 9-12 nm arasında olduğu bilinmektedir. Yapısındaki sivri uçlardan yola çıkılarak koronavirüs adı koyulmuştur (5)

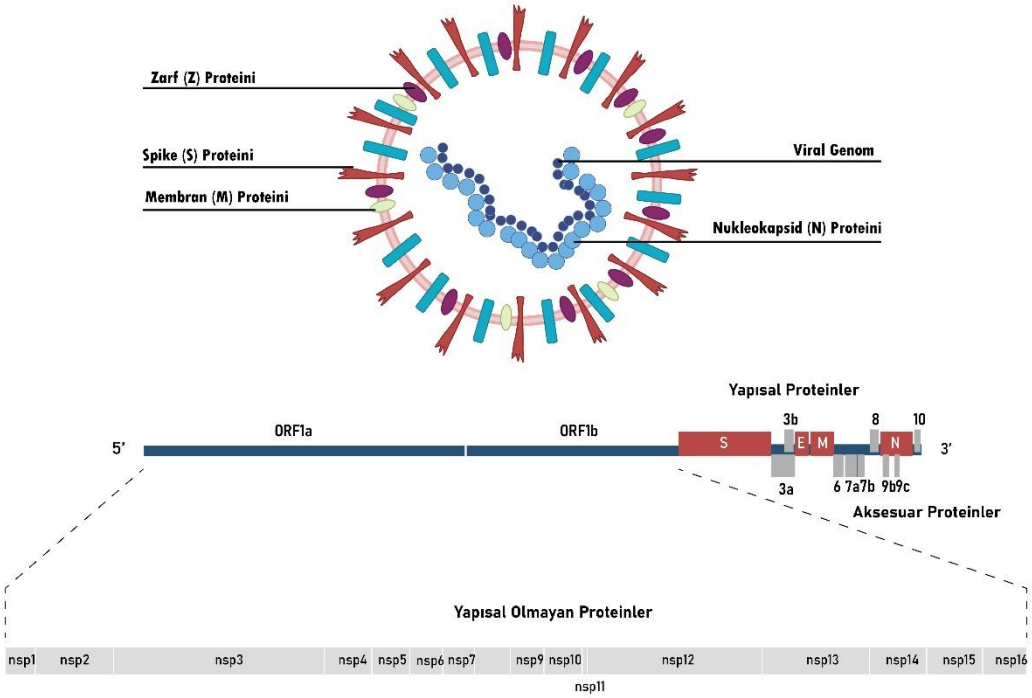


Şekil 2.2: SARS-CoV-2 virüsünün elektron mikroskobu görüntüsü (25)

Koronavirüsler beş tür protein içermektedir. Bunlar; S, E, M, N proteinleri ve hemaglutinin esteraz dimeridir. S, E, M ve HE proteinleri virüsün yüzeyinde bulunurken, genetik materyal virionun içindedir ve N proteininin genetik materyale bağlı olduğu bilinmektedir. Tüm genom yaklaşık olarak 26,000-32,000 baz uzunluğundadır ve +ssRNA'ya sahiptir (5).

2.2 SARS-CoV-2

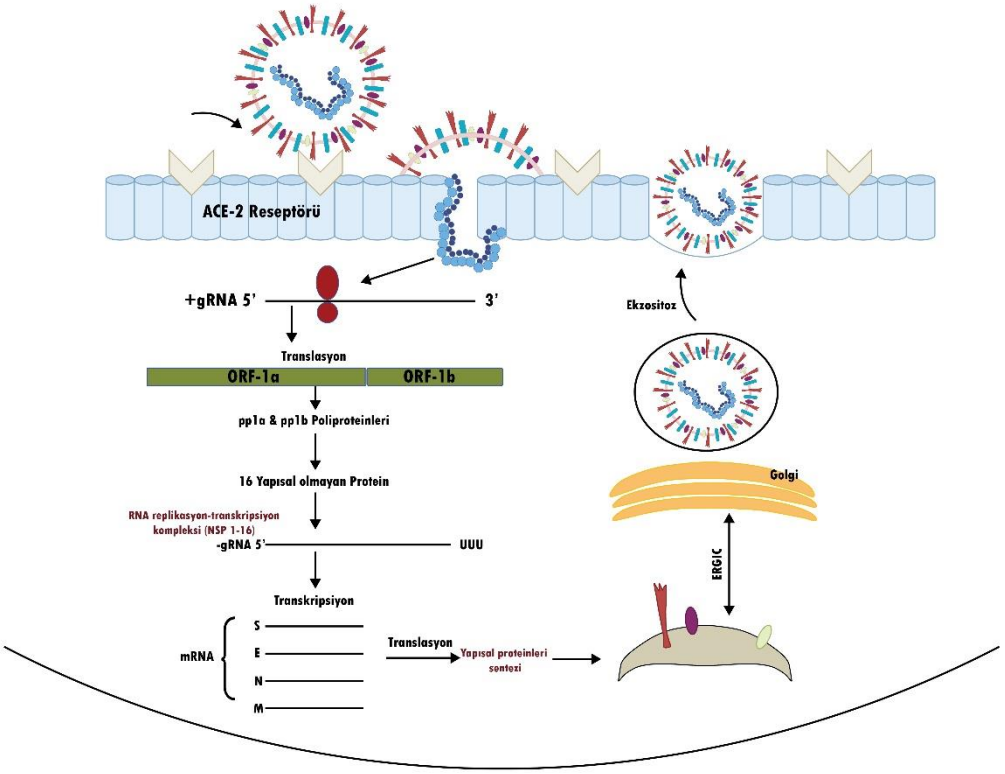
SARS-CoV-2 adı verilen yeni koronavirüsün neden olduğu COVID-19, Aralık 2019'da Çin'in Wuhan Hubei şehrinde başlamıştır. Giderek küresel olarak artan ve yayılan bir solunum hastalığıdır. SARS-CoV-2 virüsünün neden olduğu COVID-19 salgınını DSÖ tarafından 11 Mart'ta pandemi olarak ilan edilmiştir (26). Türkiye'deki ilk vaka, 10 Mart'ta bildirilirmiş olup 29/06/2021 tarihine kadar 5,420,156 vaka tespit edilmiştir (27, 28). SARS-CoV-2 virüsü taşıyıcı ya da hasta kişiler tarafından öksürme ve hapşırma sırasında üretilen damlacıklar yoluyla bulaşmaktadır. Kuluçka süresi 2 ila 14 gün arasında değişmektedir. Hasta kişilerde; yüksek ateş, öksürük, boğaz ağrısı, yorgunluk, baş ağrısı, miyalji ve nefes darlığı görülür ancak taşıyıcılar asemptomatik olarak da hastalığı atlatabilmektedirler (29).



Şekil 2.3: SARS-CoV-2 virüsünün gen organizasyonu ve yapısal proteinlerinin şematik gösterimi (30)

SARS-CoV-2'nin genomu yaklaşık 30.000 nükleotidden oluşmaktadır (18). SARS-CoV-2'nin yapısal proteinleri, S proteini, M proteini, E proteini ve N proteinidir. SARS-CoV-2'nin RNA genomu 5 'ila 3' yönünde 6 gene sahiptir: Açık Okuma Çerçevesi 1a (ORF1a), Açık Okuma Çerçevesi 1b (ORF1b), S, E, M ve N'dir (Şekil2.3). RNA

genomunun üçte iki kısmı, PP1a ve PP1ab olan iki viral replikaz proteinini üreten ORF1a ve ORF1b oluşturmaktadır (Şekil 2.3) (30). Viral enfeksiyon sırasında, SARS-CoV-2 genomunu endozomlar veya viral zarfın konak hücre zarı ile doğrudan füzyonu yoluyla konakçı hücreye enjekte eder, bu da S proteininin insan Anjiyotensin Dönüştürücü Enzimine (ACE2) bağlanmasıyla gerçekleşir. Konakçı hücreye girişin ardından kaplanmamış viral RNA konak hücre sitoplazmasına salınır ve konakçı ribozomları tarafından çevrilir. Translasyon ürünleri, poliproteinler PP1a ve PP1ab, viral proteazlar, tarafından yapısal olmayan proteinler Nsp1-16'ya proteolitik olarak bölünür. RNA (+) ipliği önce RNA (-) ipliğine kopyalanır ve daha sonra negatif iplik, ya yeni virion oluşumu için RNA (+) ipliğine replikasyon için veya alt genomik Mesajcı RNA (mRNA)'ların transkripsiyonu için kullanılır. Bu alt genomik mRNA'lar yapısal proteinlere (S, M, E, N) ve yardımcı proteinlere (ORF3a, ORF6, ORF7a, ORF7b, ORF8, ORF9 ve ORF10) çevrilir. S, M ve E proteinleri endoplazmik retikuluma (ER) girer ve N proteini, nükleoprotein kompleksi üretmek için genomic RNA (gRNA) (+) ipliğine bağlanır. Nükleoprotein kompleksi ve yapısal proteinler, viryonların toplandığı, olgunlaştığı ve golgi'den küçük kesecikler şeklinde tomurcuklandığı ER-Golgi Ara Bölmesi (ERGIC)'e hareket eder. Bu veziküller, ekzositoz yoluyla hücre dışı bölgeye salındıkları konak hücre zarına giderler. Salınan viryonlar, hastalığın ilerlemesine yol açan yeni bir hücre grubunu enfekte eder (Şekil 2.4) (4).



Şekil 2.4: SARS-CoV-2 virüsünün patogenezi (30)

COVID-19 hastalığının tanısında kullanılan yöntemlerden biri Gerçek Zamanlı Polimeraz Zincir Reaksiyonu (Real Time Polimerase Chain Reaction-RT-PCR)'dır. RT-PCR, gRNA tespiti için özel olarak geliştirilmiş bir Polimeraz Zincir Reaksiyonu (PCR) yöntemidir. RT-PCR, yeterince güvenilir ve hızlı bir tekniktir ayrıca birkaç saat içinde yüksek verimli bir şekilde sonuç vermektedir (31). SARS-CoV-2 için PCR testleri Ocak 2020'den beri mevcuttur. RT-PCR için örnekler; nazofaringeal, orofaringeal veya nazal swablar, üst ve alt solunum yolu aspiratları, bronkoalveolar lavaj ve balgam dahil olmak üzere vücudun çeşitli enfekte bölgelerinden alınabilmektedir. RT-PCR yöntemi sonuçlarında özellikle erken vakalarda %30'unda yalancı negatif sonuç gözlemlenebilmektedir. Ayrıca, COVID-19 hastalığının tanısında kullanılan serolojik testlerde mevcut olup bunlar; İmmünoglobulin M (IgM) / İmmünoglobulin G (IgG) testi ve antijen testidir (19).

SARS-CoV-2 için hala spesifik bir ilaç bulunmamaktadır. Ancak önceden var olan ilaçların yeniden konumlandırılması ve dozaj ayarı yapılması ile tedavi için kullanılmaktadır. Tedavide güncel olarak; antiviral, antibakteriyel, antimalaryal ajan,

immünomodülatör, anjiyotensin II reseptör bloklayıcı, bradykinin B2 reseptör antagonisti, anthelmintik, antiprotozoal, H2 engelleyici, antikoagülan gibi ilaç grupları kullanılmaktadır (32). COVID-19 küresel salgını için, etkili aşıların geliştirilmesi ve akut terapötik seçenekler gibi iki yönlü klinik müdahaleler gereklidir. Aşıların etkili ve başarılı bir şekilde geliştirildikleri takdirde, COVID-19 pandemisine karşı küresel mücadelede en etkili strateji olacağı vurgulanmaktadır. Bir araştırmada SARS-CoV-2'ye karşı 52 farklı aşının gelişme sürecinde olduğu bildirilmektedir (33). Ülkemizde 16 aşıya ilişkin çalışma yürütülmektedir (34). COVID-19 pandemisi ile savaşmak için yeni ve aynı zamanda etkili, aşı ile ilaç keşfi denemeleri devam etmektedir. Ancak şu an bulaşma riskini azaltmak ve yayılımı önlemek için mesafenin korunması, ellerin yıkanması ve maskelerinin kullanılması en önemli tedbirlerdir (35).

2.3 NÜKLEOKAPSİD PROTEİNİ

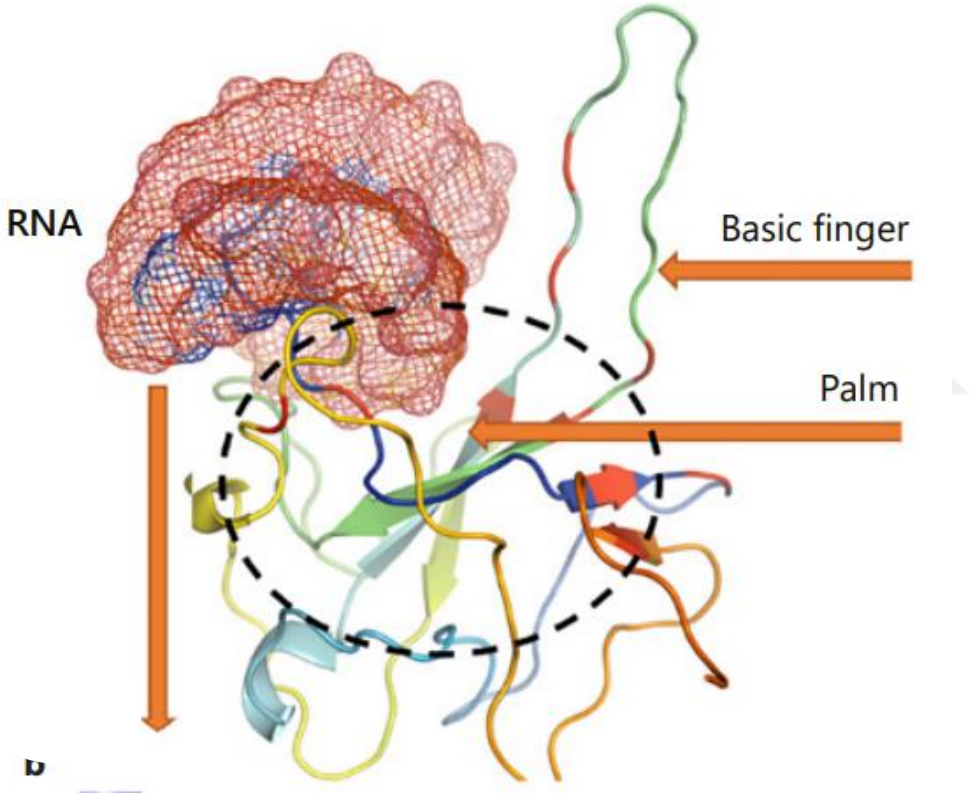
Koronavirüs ailesinde N proteini, viral RNA transkripsiyonu ve replikasyonu için gerekli olan çok işlevli bir RNA bağlayıcı proteindir. Viral RNA genomunun paketlenmesi sırasında helikal RNP'lerin oluşturulmasında, viral mRNA transkripsiyonu ve replikasyonunda rol almaktadır. Ayrıca hücre iskeleti organizasyonu ve immün regülasyon gibi enfekte hücre metabolizmasının düzenlenmesinde birçok önemli rol oynamaktadır (36). N proteini viral genomun replikasyonunu ve transkripsiyonunu kolaylaştırmak diğer bir deyişle RNA aktiviteleri için, oldukça düzenli olan RNA yapısını korumaktadır (37). N proteinin birincil işlevi, viral RNA genomuna bağlanmak ve RNP kompleksini oluşturmaktır. Ayrıca N proteininin, konakçı RNA susturma aracılı antiviral tepkilerle savaşmak için çift sarmallı RNA bağlama aktivitesini kullanmaktadır. RNA susturma, viral bir baskılayıcısı olarak işlev görmektedir. Çift sarmallı RNA, enfeksiyon sırasında çok sayıda eksprese edilmektedir ve enfeksiyon sonrasında hem humoral hem de hücresel bağışıklık tepkilerini indükleyebildiği bildirilmiştir (38). Ek olarak N proteinini; aktin organizasyonu, konak hücre döngüsü ve apoptoz gibi konakçı-patojen etkileşimlerini düzenlediğini gösteren çalışmalar da mevcuttur (36).

2.3.1 NÜKLEOKAPSİD PROTEİN YAPISI

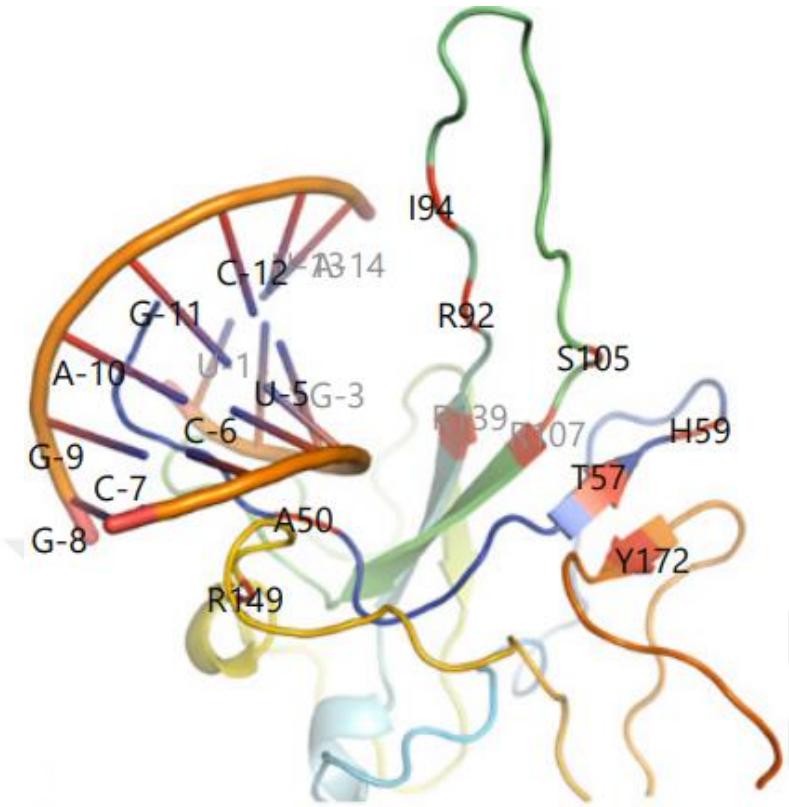
Nükleokapsid proteini; genişletilmiş bir merkezi loop'a ve β-sheet merkezden oluşan sağ el benzeri katlanmış bir yapıya sahiptir. Çekirdek bölge; β4-β2-β3-β1-β5

topolojisine sahip, beş strandden oluşan “U” şeklinde sağ el benzeri bir anti-paralel β-sheet platformundan ve bu yapının yanında 2 kısa Alfa (α) -heliks’den oluşmaktadır. Yapının göze çarpan bir özelliği, β2 – β3 arasında genişleyen bir looptur ve bunun sonucunda uzun bir β-hairpin oluşmaktadır (36, 37,39)

RNA dupleksinin, RNA ile doğrudan temas eden yüksek oranda pozitif arginin yan zincirlerini (R92, R107 ve R149) içeren NTD'nin basic finger ve palm arasında bağlı olduğu bulunmuştur (Şekil 2.5). Ayrıca RNA bağlanmasına katkıda bulunan I94 ve L104 yan zincirleri ile hidrofobik etkileşim içinde olduğu bulunmuştur. Çevredeki domainlerdeki; A50, T57, H59, R89, R92, I94, S105, R107, R149 ve Y172 yan zincirleri RNA ile etkileşime girebilmektedir (Şekil 2.6) (37).

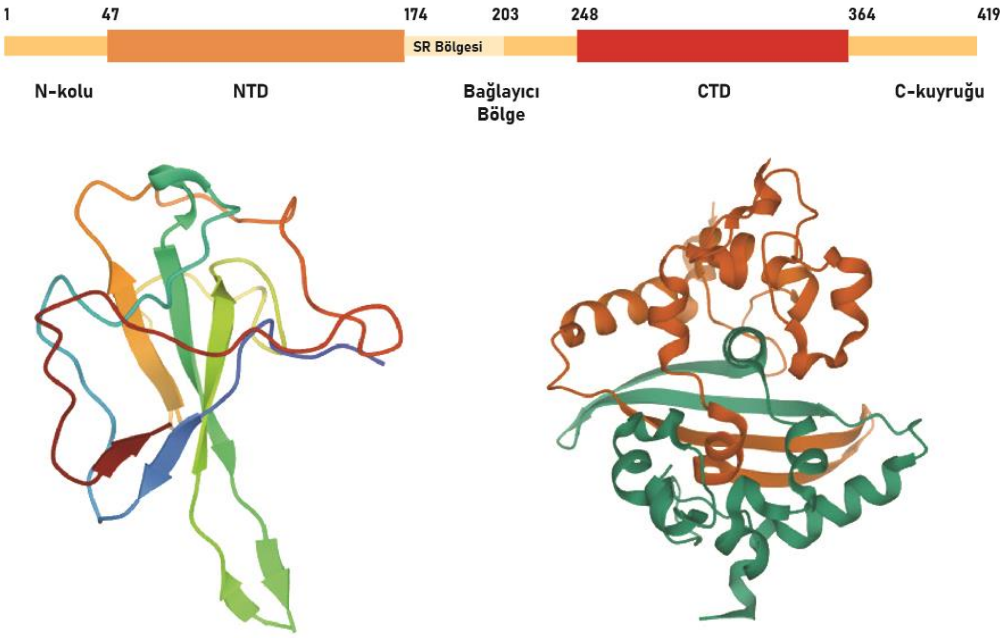


Şekil 2.5: RNA bağlanmasını kolaylaştıran N proteinlerinin yapısal bileşenleri (37)



Şekil 2.6: RNA bağlanma domaini ve etkileşimlerinde yer alan yan zincirler (37)

Nükleokapsid proteini, NTD ve CTD olmak üzere iki yapısal domaine sahiptir (Şekil 2.7). Bu iki domain bağlayıcı bölge adı verilen Düzensiz Bölge (IDR) ile birbirine bağlanmaktadır. Ayrıca NTD ve CTD, N-kolu ve C-kolu olarak adlandırılan başta ve sonda olmak üzere iki IDR'ye sahiptir. Bağlayıcı bölge, fosforilasyon bölgelerini içeren bir SR bölgesine de sahiptir. NTD RNA bağlanma afinitesi gösterdiği, CTD oligomerizasyonu sağladığı ve SR açısından zengin bağlayıcı bölgede de birincil fosforilasyon rollerine sahip olduğu bilinmektedir (Şekil 2. 5). Bunların her biri, antiviral inhibitörler tarafından hedeflenebilecek potansiyel ilaç hedefidir (37, 38).

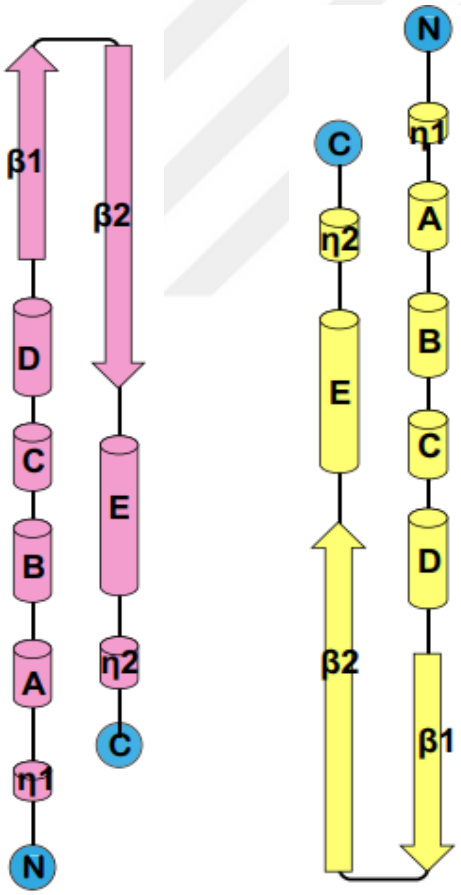


Şekil 2.7: Nükleokapsid proteinin gen organizayonu ile N-terminal domain (PDB:6M3M) ve C-Terminal Domaininin (PDB:6YUN) 3 boyutlu yapısı (36,40,41)

Her NTD molekülü sağ el dönüşlü, 4 antiparalel β strand içermektedir. Çekirdekteki alt domain 4 antiparalel β strandden oluşan bir β sheet yapısından oluşmaktadır. Ayrıca β hairpinler çekirdekten dışarı doğru çıkıntılar oluşturmaktadır. Yüzeydeki yük dağılımına bağlı olarak, loop kısımları pozitif olarak yüklenmekte ve RNA bağlanması için bir alan sağlamaktadır. Çekirdek domaini, insanları enfekte eden farklı koronavirüsler arasında korunan bir RNA bağlama bölgesi içermektedir (38).

SARS-CoV-2 NTD'nin genel yapısı, diğer CoV ailesindeki virüslerin NTD yapılarına güçlü bir şekilde benzediği bulunmuştur. Yapılan bir çalışma da SARS-CoV-2 NTD yapısını, SARS-CoV, MERS-CoV, HCoV-OC43 ve Human coronavirus NL63 (HCoV-NL63) gibi insanda enfeksiyona neden olan koronavirüslerin NTD yapılarıyla karşılaştırılmış ve yapısal farklılıkları olan iki bölge bulunmuştur. Bunlardan ilki N-terminal bölgelerinin yönelimleri farkıdır. SARS-CoV-2, SARS-CoV ve MERS-CoV virüslerinin NTD bölgesinin N-terminal loopları dışarı doğru, HCoV-OC43 ve HCoV-NL63 virüslerinin ise NTD'den loopları çekirdek alt domaine doğru döner. İkinci olarak, çıkıntı yapan β-hairpin bölgeleri esnektir; HCoV-OC43 ve HCoV-NL63 N-NTD yapılarında görünmemektedir (38).

SARS-CoV-2 CTD, sıkı bir homodimer yapı oluşturur ve genel olarak dikdörtgen bir slab şekli göstermektedir. Her protomer, "C" harfi şekline sahiptir ve beş α-heliks, iki β-strand ve iki heliksten oluşmaktadır (Şekil 2.8). Bir protomerden gelen β-hairpin, diğer protomerin boşluğuna girmekte ve bu, dimer ara yüzünde dört şeritli β-strand ile antiparalel β-sheet oluşumuyla sonuçlanmaktadır. β-sheet, slab dimerin bir yüzünü oluşturmakta, dimerin zıt yüzeyi de α-heliksler ve looplar tarafından oluşturulmaktadır. İki hairpin yapısı arasındaki kapsamlı hidrojen bağı etkileşimleri ve β-sheet ile α-heliks arasındaki hidrofobik etkileşimler, dimerik yapıyı oldukça kararlı hale getirmektedir. Bu, CTD'nin çözelti içinde bir dimer olarak var olduğuna dair yapılan biyokimyasal karakterizasyon çalışması sonuçları ile tutarlı olduğunu tespit etmişlerdir (38).



Şekil 2.8: SARS-CoV-2 C-Terminal Domaini için topoloji diyagramı (38)

2.4 MUTASYONLARIN PROTEİN YAPISINA ETKİSİ

Mutasyon, virüslerin genetik seçilimin bir sonucu olarak sürekli değişime uğradığının ana mekanizmalarından biridir. Tüm mutasyonlar arasında, yanlış anlamlı mutasyonlar aminoasit değişimlerine neden olur ve daha sonra virüs proteininin yapılarını değiştirerek viral üremeyi etkileyebilir, bulaşı arttırabilmekte, yanlış negatif tanılara ve ilaç direncine yol açabilmekte yani virüsün evrimsel olarak bazı avantajlara sahip olmasını sağlamaktadır (42).

Yanlış anlamlı mutasyonlar proteinin yapısında değişikliğe sebep olabilmektedir. Ayrıca, bu mutasyonlar da protein-protein etkileşimlerini etkileyebilmektedir. Moleküler docking (yerleştirme) ve Gibbs Serbest Enerjisi (ΔG) değerinin hesaplanması ve deleterious (zararlı) tahmini ile protein bağlanma afinitesininde değişikler saptanabilmektedir. Dahası, protein-protein etkileşimlerinde kritik rol oynayan hot spot noktalar ilaç ve aşı hedefi olarak kullanılabilmekte ancak bu noktalar çalışılırken coğrafi bölgelerde sık olarak görülen mutasyonların etkileri de araştırılmalıdır. Ayrıca mutasyona uğramış virüs proteinin, insan proteinleri ya da virüs proteinleriyle etkileşimini araştırmak, mutasyona uğramış virüsün konakçı hücrelere bulaşabilirliğini analiz etmek, aşıların ve ilaçların geliştirilmesi sırasında SARS-CoV-2 mutasyonlarının etkileşimini doğrulamak için önemlidir (42, 43).

Proteinin stabilitesi mutasyon nedeniyle bozulduğunda, mutasyonun protein işlevi üzerindeki etkisini yapıya ve sekansa dayalı sunucular ile tahmin etmek mümkündür (45). Elastic Network Contact Model (ENCoM) sunucusu; mutasyonun tüm proteinin stabilitesi üzerine olan etkisini tahmin etmektedir. Mutation Cutoff Scanning Matrix (mCSM) ağ sunucusu; bir aminoasit yan zincirini çevreleyen atomik mesafe kalıpları üzerindeki etkisini tahmin etmektedir. Site Directed Mutator (SDM) ağ sunucusu; mutasyonun homolog protein ailesi içinde ortama özgü aminoasit frekanslarını kullanıp tahmin yapmaktadır. DUET sunucusu; mCSM ve SDM ağ sunucularını birlikte kullanıp mutasyonun yapısal özelliklerini ve mutasyonu çevreleyen 3 Boyut (3B)'lu ortamına göre yorum yapmaktadır. ENCoM sunucusu ise titreşim entropi enerjisindeki değişimi kullanıp ve mutasyonun yapının esnekliğine olan etkisi hesaplamaktadır (45, 46, 47, 48, 49).

Protein Variation Effect Analyzer (PROVEAN), insersiyon ve delesyona bağlı olarak tek veya çoklu aminoasit yer değiştirmelerine tahmin oluşturmak için hizalamaya dayalı puan yaklaşımını kullanan ağ sunucusudur (50). Predict Single Nucleotide Polymorphism (PredictSNP) bu web sunucusu en iyi altı tahmin aracını derler ve deleterious (zararlı) veya nötr mutasyon türleri için doğru tahmin güven puanları verir. PredictSNP'deki algoritmalar, mutasyon etkisini tahmin etmek için farklı yaklaşımlar kullanır. Bunlar; sırasıyla fizikokimyasal özellikler ve hizalama puanı, destek vektör makinesi, uzman ampirik kurallar seti, hizalama puanı ve sinir ağı yöntemleridir (15).

Yapılan bir çalışmada GISAID olarak bilinen bir kaynakta depolanan yaklaşık 290.000 SARS-CoV-2 proteom dizisini kullanarak, proteinlerin farklı oranlarda mutasyona uğradığı bulunmuştur. Spesifik olarak, COVID-19 aşıları ve terapötiklerinin ana hedefleri olan S ve N proteinlerinin en yüksek mutasyona sahip olan proteinler olduğu tespit edilmiştir (51). Bu nedenle çalışmamızda DynaMut, mCSM, SDM ağ sunucuları ile DUET ve ENCoM adlı sunucular tarafından yapı temelli analizinin ve PROVEAN, PredictSNP ve Sorting İntolerant From Tolerant (SIFT) ağ sunucuları ile Multivarivate Analysis of Protein Polymorphismv(MAPP) ve Screening for non-Acceptable Polymorphisms (SNAP) ağ araçları tarafından sekansa dayalı analizinin yapılarak, Türkiye'deki SARS-CoV-2 genomlarında bulunan N proteinine ait yanlış anlamlı mutasyonların nükleokapsid proteinin 3 boyutlu yapısı üzerine etkilerinin araştırılması amaçlandı.

3. GEREÇ VE YÖNTEM

3.1 MUTASYONLARIN PROTEİN YAPILARINA ETKİLERİNİN İN-SİLİKO OLARAK ARAŞTIRILMASI

Bir proteinin aminoasit dizisi, yapısını ve sonuç olarak işlevini belirlemektedir. Tek bir aminoasit mutasyonu bile bir proteinin yapısını değiştirebilmekte ve bu da hastalıkların nedeni olabilmektedir (21). Yanlış anlamlı mutasyonların gen ekspresyonu ve protein işlevi üzerindeki etkilerinin anlaşılması önemlidir. Yanlış anlamlı mutasyonlar, sadece protein stabilitesini değiştirmekle kalmamakta, aynı zamanda proteinler, nükleik asitler veya polisakaritler, ligandları ve metal iyonları da dahil olmak üzere diğer biyolojik moleküllerle etkileşimini bozabilmekte hatta işlev kaybına yol açabilmektedir (44). Bu nedenle mutasyonların protein yapısı üzerine etkisinin araştırılması önemlidir. Fiziksel proteinlerde mutagenez çalışmaları yapmak, amino asit yer değişimlerinin etkileri hakkında fikir verebilir, ancak bu tür ıslak laboratuvar çalışmaları, gereken zaman ve finansal kaynaklar nedeniyle engelleyicidir. Bir protein yapısı üzerindeki bir mutasyonun etkilerini tahmin etmeye yönelik biyoinformatik çalışmalar, ıslak laboratuvar çalışmaları için hazırlık olabilmekte ve daha sonra tamamlayabilmektedir (21).

Protein yapısı, aminoasitlerin arasındaki etkileşimle ve çevre ile etkileşimi sonucu belirlenmektedir. Böylece, protein yapısı kararlı bir 3B yapı ile sonuçlanmaktadır. Katlanma işlemi sonucunda proteinin kararlı 3B yapısı, protein çevre sisteminin ΔG'sine karşılık gelmektedir. Minimum Gibbs serbest enerjisi; hidrofobik etkiler ve protein konfigürasyonları gibi entropik etkileşimlerle Van der Waals etkileşimleri, hidrojen ve elektrostatik bağlanma gibi protein içi etkileşim enerjilerini içermektedir. Yanlış anlamlı mutasyonlar sonucunda protein stabilitesi artabilir ya da azalabilir bunun sonucunda proteinin işlevi değişebilmektedir. Bir protein domainin yüzeyinde meydana gelen mutasyonlar genellikle nötrdür, çekirdekte meydana gelen mutasyonlar ise katlanma domaindeki kararlılığı değiştirebilmektedir. Ayrıca protein domainin yüzeyinde meydana gelen mutasyonlar, diğer proteinler için bağlanma afinitelerini veya çok alanlı proteinlerdeki alanlar arası bağı etkileyebilmektedir (52).

Yanlış anlamlı mutasyonların, protein stabilitesi üzerindeki etkileri, katlanmamış ΔG cinsinden ölçülmektedir. ΔG; sıcaklık, pH değerleri, tuz konsantrasyonları, organik çözücüler, üre ve diğer kimyasal maddeler gibi çeşitli faktörlerin doğrusal olmayan bir fonksiyonudur. Deneysel bakış açısından, mutasyona uğramış ve doğal tip proteinler ($\Delta\Delta G^{u}$) arasındaki açığa çıkan serbest enerjinin farkıdır. Böylece $\Delta\Delta G^{u} = \Delta G^{u}_{mutantı} - \Delta G^{u}_{doğal\ tip}$, mutasyona uğramış ve doğal tip proteinler arasındaki katlanmış yapının serbest enerjisinin farkıdır. ΔΔG işareti, varyasyonun protein stabilitesini azalttığını ($\Delta\Delta G^{u} < 0$) veya arttırdığını ($\Delta\Delta G^{u} > 0$) göstermektedir (52).

$$\Delta\Delta G^{u} = \Delta G^{u}_{mutant} - \Delta G^{u}_{doğal\ tip}$$

Mevcut tahmin algoritmaları, protein stabilitesini karakterize eden; yapısal temelli özellikler (yan zincir temas ağları, yan zincir/atom mesafeleri, protein geometrisi), sekansa dayalı özellikler, (korunmuş sekanslara ve amino asit pozisyonlarına dayalı), enerji temelli özellikler (Van der Waals etkileşimleri, çözme enerjisi, ekstra stabilize edici serbest enerji vb.) ve moleküler temelli özellikler (arayüzün çözücüyle erişilebilen yüzey alanı, hidrofobik ve hidrofilik alan) kullanmaktadır (52). Yukarıda bahsedilen özelliklerin tümü mevcut yöntemlerin çoğunda kullanılmaktadır ancak bazı yöntemler sadece bir alt kümeyi kullanmaktadır. Protein stabilitesi için ilk öngörücü yaklaşımlar, moleküler dinamikler veya Monte Carlo simülasyonları ile birleştirilen moleküler mekanikten türetilen fiziksel serbest enerji fonksiyonlarına dayalı kuvvet alanlarının uygulanması yoluyla geliştirilmiştir (44, 52). Mevcut tahmin algoritmaları, sekansa ve yapıya dayalı olmak üzere iki grubu ayrılmaktadır (44) (Tablo 3.1).

Sekansa dayalı yöntemler, proteinlerin amino asit dizisini kullanmakta ve stabilitedeki değişiklikleri tahmin etmek için makine öğrenme tekniklerini uygulamaktadır (53). Yapıya dayalı yöntemler, sekans ve yapısal bilgilerin bir kombinasyonunu kullanılabilmektedir (38).

Tablo 3.1: Çalışmada kullanılan sekansa ve yapıya dayalı sunucular

SEKANSA DAYALI SUNUCULAR	YAPIYA DAYALI SUNUCULAR
PROVEAN	DynaMut
PREDICT SNP	mCSM
MAPP	DUET
SIFT	SDM
SNAP	ENCoM

3.2 MUTASYON ANALİZİNDE KULLANILAN SUNUCULAR

3.2.1 Yapı temelli sunucular

DynaMut ağ sunucusu; protein dinamiklerini ve nokta mutasyonlarının protein stabilitesi üzerindeki etkisini analiz etmek için kullanılmaktadır. Ayrıca konformasyonları örnekleyerek protein dinamiklerini analiz etmek, görselleştirmek ve titreşimsel entropi değişikliklerinden kaynaklanan protein dinamikleri ve stabilite üzerindeki mutasyonların etkisini değerlendirmede kullanmaktadır. Ayrıca doğal tipli yan zincirinin çevre özelliklerine ilişkin bir dizi tamamlayıcı bilgi (örneğin, görece çözücü erişilebilirliği, yan zincir derinliği ve ikincil yapı) ve doğal tip yapıyı temsil eden grafik tabanlı sonuçlar içermektedir. DynaMut ağ sunucusu, Arpeggio tarafından hesaplanan kovalent olmayan moleküler etkileşimlerin, doğal tip ve mutant kalıntıların ilgili 3B yapılarındaki deformasyon enerjilerinin ve atomik dalgalanmalarının görselleştirilmesini de sağlamaktadır (46).

mCSM ağ sunucusunun temel amacı; bir mutasyonun etkisiyle ilişkilendirilebilecek bir yan zincirinin çevresindeki atomik mesafe modellerinin etkisini göstermektir. Belirli bir mutasyon bölgesi için, doğal tipte yan zincir ortamını, geometrik merkezinden r mesafesi içindeki atomlar tarafından tanımlanmaktadır. Belirlenen bölgenin atomları arasındaki ikili mesafe hesaplamasında, yakından uzağa belirlenen bölgenin mesafe spektrumunu açıklayan bir atom mesafe matrisi oluşturmaktadır. Ayrıca mutasyonun neden olduğu atom değişikliklerini hesaba katmak için, farmakofor sayımı yapılmaktadır. Farmakofor türleri; hidrofobik, pozitif, negatif, hidrojen alıcısı, hidrojen donörü, aromatik, kükürt ve nötrdür. Yirmi aminoasit yan zincirinin her biri farklı bir vektörle temsil edilir. Buradaki her pozisyon, yan zincirdeki belirli bir farmakoforun frekansını belirtmektedir. Doğal tip ve mutant farmakofor vektörleri arasındaki fark da

sonuca eklenmektedir. mCSM ağ sunucusu ile elde edilen mutasyonun etkisinin sonuçlarının, bir aminoasit yan zincirinin çevreleyen atomik mesafe modelleriyle ilişkilidir. Bu sonuçlar; protein yapısal sınıflandırması, fonksiyon tahmini ve reseptör bazlı protein ligand tahmini dahil olmak üzere birçok çalışmada kullanılmaktadır (48).

SDM ağ sunucusu, Topham ve arkadaşları tarafından geliştirilen istatistiksel bir potansiyel enerji fonksiyonudur. SNP'lerin proteinlerin stabilitesi üzerindeki etkisini tahmin etmektedir. Doğal tip ve mutant protein arasındaki serbest enerji farkına benzer olan bir stabilite skorunu hesaplamak için kullanılan yer değiştirme verileri, evrimleri sırasında çoklu mutasyonları olarak kabul edilen homolog protein ailelerinden alınmaktadır. Sonuçlar, doğal tip ve mutant aminoasit yan zincirlerinin lokal yapısal ortamı hakkında bilgi, stabilite skoru tahmini ve hastalık ilişkisi tahminini içermektedir (45, 49).

DUET sunucusu; sıralı minimal optimizasyon ile geliştirilmiş destek vektör makineleri (SVM) kullanılan optimize edilmiş bir sunucuda ayrı yöntemlerin sonuçlarını birleştirerek mCSM ve SDM ağ sunucusunu birleştirmektedir. Bir protein yapısındaki tek nokta mutasyonu verildiğinde; iki yöntemden gelen tahminleri doğrusal olmayan bir şekilde birleştirmekte ve hesaplamaktadır. Bunu yapmak için, SDM ağ sunucusu ile elde edilen ikincil yapıyı ve mCSM ağ sunucusu ile elde edilen doğal tip ve mutant yan zincir arasındaki değişiklikleri açıklamakta, mutasyona ilişkin tamamlayıcı bilgileri de hesaplamakta ve analiz etmektedir. Sonuçlar, Gibbs Serbest Enerji ($\Delta\Delta G$) türünden ifade edilir. Negatif sonuçlu değerler, stabil olmayan mutasyonları ifade etmektedir. Sonuç kısmında; göreceli solvent erişilebilirliği (relative solvent accesibilty), yan zincir hidrojen bağı memnuniyeti (side-chain hydrogen bond satisfaction) ve ikincil yapı ek bilgileri de sunulmaktadır (45).

ENCoM sunucusu; tek nokta mutasyonlarının protein dinamikleri ve titreşimsel entropi değişikliklerinden kaynaklanan termostabilite üzerindeki etkisini tahmin etmek için kaba taneli normal mod analiz yöntemini kullanmaktadır. Ayrıca ENCoM sunucusu, her yan zincirin proteinin esnekliği açısından etkisini anlamada b-faktörünü kullanmaktadır. Her mutasyonun tüm proteinin esnekliği üzerindeki etkisini doğal tipe göre göstermektedir. Yan zincirlerde daha az esnek olanlar mavi, daha esnek olanlar ise kırmızı renk ile gösterilmektedir. Proteindeki sertleşmenin azalması ya da artması proteinin aktivasyonundaki değişimle yakından ilişkilidir (47).

3.2.2 Sekans temelli sunucular

SNAP, sekans bazlı elde edilen bilgileri kullanarak proteindeki tüm Sinonim Olmayan Tek Nükleotid Polimorfizm (Non-synonymous single nucleotide polymorphisms-nsSNP)'leri nötr olmayan (fonksiyon üzerindeki etki) ve nötr (etki yok) olarak sınıflandırabilen bir ağ aracıdır. SNAP her bir örnek için güvenilirlik indeksi denilen belirli bir tahminin güven seviyesini yansıtan iyi kalibre edilmiş bir ölçü sağlar. Ayrıca, nsSNP'lerin işlevsel etkilerini, evrimsel bilgileri (dizi aileleri içinde yan zincir koruma), protein yapısının tahmin edilen yönlerini (ikincil yapı, çözücü erişilebilirliği) ve diğer ilgili bilgileri dahil ederek doğru bir şekilde tahmin eden bir ağ aracıdır (54).

SIFT, bir aminoasit mutasyonunun deleterious (zararlı) olup olmadığını tahmin etmek için korunan sekansları ve amino asit özelliklerini kullanan çok adımlı bir algoritmadır. İlk olarak, protein veritabanında benzer dizileri arar ve %90 özdeşlikte konsensüs dizilerini kümeler. İkinci olarak, seçilen konsensüs dizileri bir medyan dizi koruma eşiğine ulaşana kadar PSI-BLAST ile yinelemeli olarak en yakından ilişkili konsensüs dizisi seçmektedir. Son olarak aminoasit mutasyonları için olasılıkları hesaplamaktadır (55).

PROVEAN, tüm genomda bulunan çok sayıda varyasyonu toplu şekilde analiz etmektedir. Bu araç, tek nükleotid polimorfizmleri, çoklu nükleotid yer değişimlerini, insersiyon veya delesyon gibi genomik varyasyonların bir listesini kabul eder. İlk olarak, genomik varyasyonlar, referans genom dizisine ve Ensembl gen açıklamasına dayalı olarak kodlayan ve kodlamayan varyasyonlar olarak sınıflandırılır. İkinci olarak, kodlanan varyantları, amino asit yer değiştirmeleri, insersiyon, delesyon, anlamsız mutasyonlar veya çerçeve kaymaları olarak protein sekansı seviyesinde sınıflandırmakta ve daha sonra bunlar için PROVEAN puanları hesaplanmaktadır (56).

PredictSNP mutasyonların protein fonksiyonu üzerindeki etkilerini tahmin etmek için sekiz programı kullanan ve konsensüs dizileri sınıflandıran bir ağ sunucusudur. PredictSNP ağ sunucusunda kullanılan; nsSNPAnalyzer, Predictor of Human Deleterious Single Nucleotide Polymorphisms (PhD-SNP), Polymorphism Phenotyping version 2 (PolyPhen-2) ve SNAP çeşitli makine öğrenimi yöntemleriyle eğitilmiş bir karar modeli kullanmaktadır. SIFT ve Protein Analysis Through Evolutionary Relationships (PANTHER) yalnızca evrimsel bilgileri kullanmaktadır. MAPP doğal tip ve mutant aminoasitler arasındaki fizikokimyasal özelliklerdeki farklılıkları da dikkate almaktadır.

Polymorphism Phenotyping version 1 (PolyPhen-2), sınıflandırma için uzman bir ampirik kurallar seti kullanmaktadır (43).

MAPP, bir protein dizisi hizalamasının bir sütununda bulunan fizikokimyasal varyasyonu dikkate alır ve bu varyasyona dayanarak, tüm olası amino asit ikamelerinin proteinin işlevi üzerindeki etkisini tahmin etmektedir. Elde edilen değerlerde, sapma ne kadar büyük olursa mutasyonun proteinin işlevini bozma olasılığı ve proteinin işlevi üzerindeki etkisi o kadar yüksek olarak belirlenmektedir (57).

3.3 ÇALIŞMADA KULLANILAN DİZİLERİN ELDE EDİLMESİ

SARS-CoV-2 genom dizileri, 01 Mart 2020 - 15 Mart 2021 tarihleri arasında GISAID EpiCoV veri tabanına yüklenmiş verilerden elde edildi. Bu diziler, sadece Türkiye bölgesindeki verilere ulaşmak için "Avrupa/Türkiye" lokasyonu ile filtrelendi ve 2557 genom dizisiyle sonuçlandı. Analiz edilecek dizilerin kalitesini artırmak amacı ile de filtreler uygulandı. Yalnızca insan konakçılardan alınan, genomun tam uzunluğu, (sekans boyutu 29 kilo baz veya daha büyük olan), yüksek kapsama (high coverage)((< 0.05% özgün mutasyon, insersiyon ya da delesyon belirtilmemiş, < 0.1% sekans) ve düşük kapsamaların dışlandığı (low coverage excuded) (>5% Ns) filtreler uygulandı. Bunun sonucunda, 2286 viral genom dizisi GISAID EpiCoV veri tabanından indirilmiştir.

3.3.1 Nükleokapsid Protein Mutasyonlarının Tespit Edilmesi

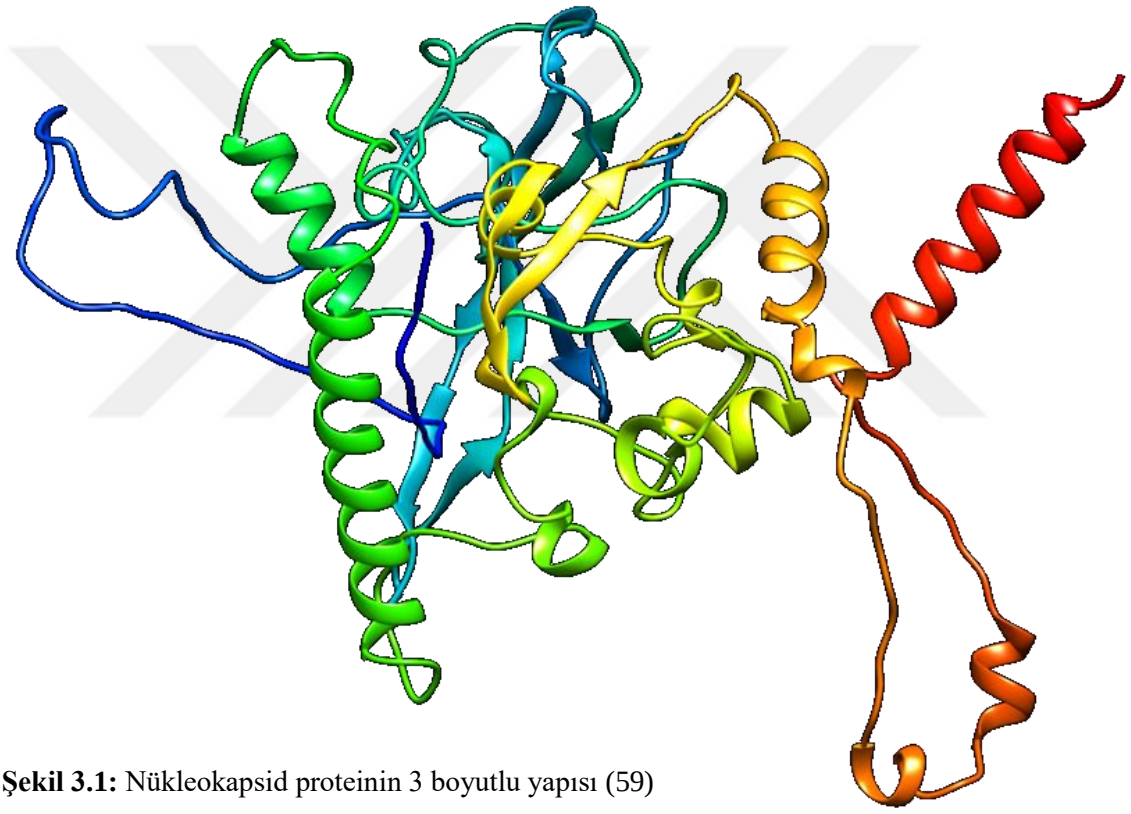
Türkiye'den elde edilen genom dizilerinde bulunan, nukleokapsid proteini yanlış anlamlı mutasyonları Nextclade ağ aracı kullanılarak tespit edildi (58).

Nextclade, analiz edilmek istenen SARS-CoV-2 genom dizlerindeki mutasyonları MN908947.3 referans genomu kullanarak tanımlayan bir ağ aracıdır. Filtrelenerek elde edilmiş 2286 genom dizisini içeren fasta formatındaki dosya Nextclade çevrim içi ağ aracına yüklendi. Daha sonra sonuçlar csv ve tsv formatında kaydedildi.

3.4 MUTASYONUN PROTEİN FONKSİYONU ÜZERİNDEKİ ETKİLERİNİN ANALİZİ

3.4.1 NÜKLEOKAPSİD PROTEİNİNİN ÜÇ BOYUTLU YAPISININ ELDE EDİLMESİ

Üç boyutlu N protein yapısı I-Tasser sunucusundan indirildi (59). I-tasser sunucusu bu yapıyı elde edebilmek için 6M3M ve 6YUN pdb kodlu 3 boyutlu protein yapılarını kullanmaktadır. I-Tasser sunucusundan indirilen nükleokapsid proteininin kodu QHD43423.pdb'dır (Şekil 3.1). Bu yapının TM değeri 0.97'dir.



Şekil 3.1: Nükleokapsid proteinin 3 boyutlu yapısı (59)

3.4.2 DİZİ TEMELLİ ANALİZ

P0DTC9 kodlu SARS-CoV-2 nükleokapsid protein dizisi NCBI veri bankasından indirildi (60). Dizi temelli analiz için PROVEAN ve PredictSNP ağ sunucuları kullanıldı. PROVEAN ve PredictSNP ağ sunucularını kullanmak için P0DTC9 kodlu SARS-CoV-2 nükleokapsid protein dizisi ile 161 adet yanlış anlamlı mutasyonu içeren txt biçiminde bir dosya yüklendi.

GISAID EpiCoV veri tabanından indirilen 2286 örnekte bulanan N proteine ait 161 adet yanlış anlamlı mutasyon analiz edildi (61).

3.4.3 YAPI TEMELLİ ANALİZ

Yapı temelli analiz, DynaMut ağ sunucusu ile yapıldı (62). DynaMut ağ sunucusuna referans olarak [QHD43423.pdb] kodlu N protein yapısı ve 161 adet yanlış anlamlı mutasyonu içeren txt biçiminde bir dosya yüklendi.

GISAID EpiCoV veri tabanından indirilen 2286 örnekte bulanan N proteine ait 161 adet yanlış anlamlı mutasyon analiz edildi.

4. BULGULAR

4.1 TESPİT EDİLEN YANLIŞ ANLAMLI MUTASYONLAR

GISAID EpiCoV veri tabanından indirilen 2286 örnekte bulanan N proteine ait 161 adet yanlış anlamlı mutasyon tespit edildi. Mutasyonların 39 adeti N-kolunda, 29 adeti NTD'de, 44 adeti bağlayıcı bölgede, 23 adeti ise CTD'de ve 26 adeti C-kuyruğunda yer almaktadır (Tablo 4.1, Tablo 4.2, Tablo 4.3, Tablo 4.4, Tablo 4.5).

Tablo 4.1: Nükleokapsid proteinin N-kolunda tespit edilen mutasyonlar

MUTASYON	GEN BÖLGESİ	GÖRÜLME SIKLIĞI
D3L	A28281T, T28282A	382
D3Y	G28280T	33
P13S	C28310T	26
A35V	C28377T	17
Q7K	C28292A	7
G25C	G28346T	6
Q9H	G28300T	5
D22N	G28337A	4
G34W	G28373T	4
S2Y	C28278A	4
N8T	A28296C	3
P13L	C28311T	3
R40C	C28391T	3
G25V	G28347T	2
G44C	G28403T	2
S23L	C28341T	2
S33I	G28371T	2
A12G	G28308G	1
A12S	G28307T	1
D22Y	G28337T	1
D3N	G28280A	1
D3Q	G28280C, T28282A	1
E31G	A28365G	1
G19A	G28329C	1
N29S	A28359G	1
N8Y	A28295T	1
Q28K	C82A	1
Q7L	A28293T	1
Q9R	A28299G	1
R10L	G28302T	1
R14C	C28313T	1
R32H	G28368A	1
S21P	T28334C	1
S2F	C28278T	1
S33G	A28370G	1
S37P	T28382C	1
S37T	T28382A	1
T16M	C28320T	1
T24A	A28343G	1

Tablo 4.2: Nükleokapsid proteinin N-Terminal Domaininde tespit edilen mutasyonlar

MUTASYON	GEN BÖLGESİ	GÖRÜLME SIKLIĞI
D128N	G28655A	11
P67L	C28473T	8
D128Y	G28655T	6
L139F	G28690T	4
A119S	G28628T	3
A90S	G28541T	3
T135I	C28677T	3
T166I	C28770T	3
D144H	G28703C	2
E136Q	G28679C	2
H145Y	C28706T	2
L167F	G28774T	2
N126S	C28473T	2
P80R	C28512G	2
A152S	G28727T	1
A173V	C28791T	1
D128E	C28657A	1
D63Y	G28460T	1
E118Q	G28625C	1
M101I	G28576T	1
N140S	A28692G	1
P142S	C28697T	1
P151A	C28724G	1
P151S	C28724T	1
P67S	C28472T	1
Q163K	C28760A	1
T141I	C28695T	1
T49I	C28419T	1
V72L	G28487C	1

Tablo 4.3: Nükleokapsid proteinin bağlayıcı bölgesinde tespit edilen mutasyonlar

MUTASYON	GEN BÖLGESİ	GÖRÜLME SIKLIĞI
R203K	G28881A, G28882A	1158
G204R	G28883C	1117
S235F	C28977T	387
S194L	C28854T	164
T205I	C28887T	118
T198I	C28866T	42
M234I	G28975T	28
G204L	G28883C, G28884T	23
S202N	G28878A	22
A220V	C28932T	21
A211V	C28905T	19
P199S	C28868T	16
G238V	G28986T	15
N196Y	A28859T	13
S188P	T28835C	8
S193T	G28851C	8
R195I	G28857T	7
P199L	C28869T	5
G212V	G28908T	4
G238C	G28985T	4
G214C	G289113T	3
R191L	G28845T	3
M210I	G28903T	2
P207H	C28893A	2
P207S	C28892T	2
G215S	G28916A	2
A208S	G28895T	1
A211G	28906	1
A218V	C28926T	1
G204P	G28883C, G28884C	1
G204Q	G28883C, G28884A	1
M210N	28906	1
N213S	A28911G	1
Q229R	A28959G	1
Q241L	A28995T	1
R185C	C28826T	1
R185G	C28826G	1
R209S	A28900C	1
S180C	A28811T	1
S201C	A28874T	1
S202C	A28877T	1
S206F	C28890T	1
S232N	G28968A	1
V246A	T29010C	1

Tablo 4.4: Nükleokapsid proteinin C-Terminal Domaininde tespit edilen mutasyonlar

MUTASYON	GEN BÖLGESİ	GÖRÜLME SIKLIĞI
T362I	C29358T	41
T334I	C29274T	9
M322I	G29239A	7
I320F	A29231T	3
A308S	G29195T	2
H300Y	C29171T	2
M317T	T29223C	2
P364S	A29147C	2
T271I	C29085T	2
T334A	A1000G	2
V324I	G29243A	2
V350G	T29322G	2
I292L	A29147C	1
I337V	A29282G	1
K299N	A29170C	1
L339F	G29290T	1
P326S	C29249T	1
P364Q	C29364A	1
Q289H	G29140C	1
S327L	C29253T	1
T296I	C29160T	1
V350A	T29322C	1
V350L	G29321C	1

Tablo 4.5: Nükleokapsid proteinin C-kuyruğunda tespit edilen mutasyonlar

MUTASYON	GEN BÖLGESİ	GÖRÜLME SIKLIĞI
D377Y	G29402T	22
D401Y	G29474T	21
E378Q	G29405C	17
P383S	C29420T	14
A376T	G29399A	5
P383L	C29421T	5
T379I	C29409T	5
P365S	C29366T	4
A414S	G29513T	3
E378A	A29406C	3
D377A	A29403C	2
Q380L	C29411T, C29412T	2
Q389H	G1167C	2
S413T	G29511C	2
S416L	C29520T	2
T417I	C29523T, T29524C	2
K375E	A29396G	1
P365L	C29367T	1
P365Q	C29367A	1
Q389L	A29439T	1
Q390H	A29443C	1
Q418H	G29527T	1
S413I	G29511T	1
T366I	C29370T	1
T379A	A29408G	1
T391I	C29445T	1

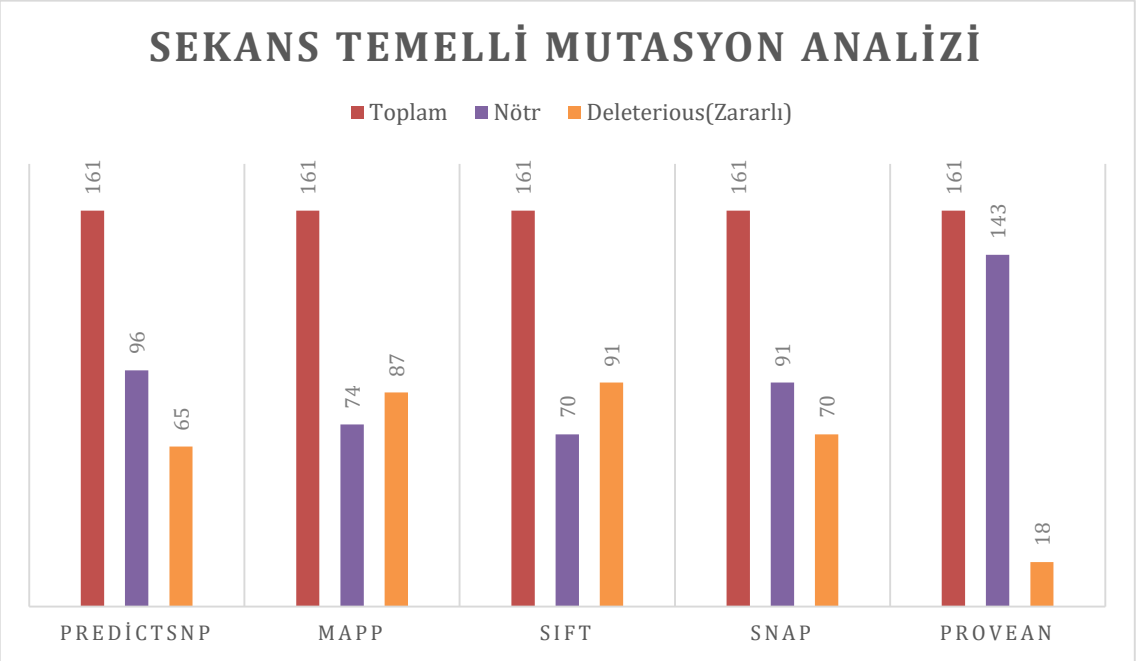
4.2 SEKANS TEMELLİ ANALİZ SONUÇLARI

Dizi temelli analiz için PROVEAN ve PredictSNP ağ sunucuları kullanıldı. PROVEAN ve PredictSNP ağ sunucularının sonuçları mail ile elde edildi.

PROVEAN skorları için; -2,5 veya daha düşük puana sahip mutasyonlar “deleterious (zararlı)” olarak ve -2,5'in üzerinde puan alan mutasyonlar ise "nötr" olarak kabul edildi (56). SIFT skorları 0 ile 1 arasında değişmektedir. SIFT skoru; 0,05’den daha düşük puana sahip mutasyonlar “deleterious (zararlı)” olarak ve SIFT skoru 0,05 ya da 0,05’in üzerinde puana sahip mutasyonlar ise nötr olarak kabul edildi (55). SNAP skorları -100 ile +100 arasında değişmektedir. SNAP skoru; 0 ile -100 arasında olan mutasyonlar “nötr”, SNAP skoru 0 ile +100 arasında olan mutasyonlar için etkili olarak kabul edildi (54). PredictSNP skorları -1 ile +1 arasında değişmektedir. PredictSNP skoru; 0 ile -1 arasında olan mutasyonlar “nötr”, PredictSNP skoru 0 ile +1 arasında olan mutasyonlar için etkili olarak kabul edildi (43). MAPP skoru 4,55 değerinden büyük olan mutasyonların proteinin fonksiyonunu etkilemektedir. Yüksek skorlar, gözlemlenen evrimsel varyasyona fizikokimyasal farklılığı nedeniyle potansiyel olarak deleterious bir varyantı tanımlanmaktadır (57).

PROVEAN ağ sunucusunda yapılan analiz sonucunda 18 adet yanlış anlamlı mutasyon deleterious (zararlı) olarak ve 143 adet yanlış anlamlı mutasyon ise nötr olarak tespit edildi. PredictSNP ağ sunucu ile yapılan analiz sonucunda 65 adet yanlış anlamlı mutasyon deleterious (zararlı) olarak ve 96 adet yanlış anlamlı mutasyon ise nötr olarak tespit edildi. SNAP ağ aracı ile yapılan analiz sonucunda 70 adet yanlış anlamlı mutasyon deleterious (zararlı) olarak ve 91 adet yanlış anlamlı mutasyon ise nötr olarak tespit edildi. SIFT ağ sunucusu ile yapılan analiz sonucunda 91 adet yanlış anlamlı mutasyon deleterious (zararlı) olarak ve 70 adet yanlış anlamlı mutasyon ise nötr olarak tespit edildi. MAPP aracı ile yapılan analiz sonucunda 87 adet yanlış anlamlı mutasyon deleterious (zararlı) olarak ve 74 adet yanlış anlamlı mutasyon ise nötr olarak tespit edildi (Tablo 4.6, Tablo 4.7, Tablo 4.8, Tablo 4.9, Tablo 4.10, Tablo 4.11).

Tablo 4.6: Farklı sunucular ile elde edilen nötr ve deleterious (zararlı) mutasyonların dağılımı



Tablo 4.7: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteininin N-kolu bölgesine ait mutasyon sonuçları (N: Nötr D: Deleterious (zararlı))

MUTASYON	PROVEAN	PredictSNP	MAPP	SIFT	SNAP
D3L	N	D	D	D	D
D3Y	N	D	D	D	D
P13S	N	D	N	D	D
A35V	N	D	D	D	D
Q7K	N	N	D	N	N
G25C	N	D	D	D	D
Q9H	N	D	D	D	D
D22N	N	D	N	D	D
G34W	N	D	D	D	D
S2Y	N	D	N	D	D
N8T	N	N	N	D	N
P13L	N	D	D	D	D
R40C	N	D	D	D	D
G25V	N	D	D	D	D
G44C	N	D	D	D	D
S23L	N	N	D	D	D
S33I	N	D	D	D	D

Tablo 4.7: (devam)

A12G	N	N	N	D	N
A12S	N	N	N	D	N
D22Y	N	D	D	D	D
D3N	N	D	N	D	D
D3Q	N	D	N	D	D
E31G	N	N	D	N	N
G19A	N	N	D	D	N
N29S	N	N	N	N	D
N8Y	N	N	N	D	N
Q7L	N	N	D	D	D
Q9R	N	D	D	D	D
R10L	N	N	D	D	D
R14C	N	D	D	D	D
R32H	N	D	D	D	D
S21P	N	D	D	D	D
S2F	N	D	D	D	D
S33G	N	N	N	N	N
S37P	N	N	N	N	N
S37T	N	N	D	N	N
T16M	N	D	D	D	D
T24A	N	N	D	D	D
Q28K	N	N	N	N	N

Tablo 4.8: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteininin N-Terminal Domainine ait mutasyonların analiz sonuçları (N: Nötr D: Deleterious (zararlı))

MUTASYON	PROVEAN	PredictSNP	MAPP	SIFT	SNAP
D128N	N	N	N	N	N
P67L	N	D	D	D	N
D128Y	D	D	D	D	D
L139F	N	N	N	N	N
A119S	N	D	D	D	N
A90S	N	N	N	N	N
T135I	N	N	D	D	N
T166I	N	N	N	N	N
D144H	N	D	D	D	D
E136Q	N	N	N	N	N
H145Y	N	N	-	D	N
L167F	N	D	D	D	N
N126S	N	N	N	N	N
P80R	N	N	D	N	D
A152S	N	N	N	N	N
A173V	N	N	N	N	N
D128E	N	N	N	N	N
D63Y	N	D	D	D	D
E118Q	N	N	N	N	N
M101I	N	N	N	N	N
N140S	N	N	N	N	N
P142S	N	N	N	D	N
P151A	D	D	D	D	D
P151S	D	D	D	D	D
P67S	N	N	N	N	N
Q163K	N	N	D	N	N
T141I	N	N	D	D	N
T49I	N	N	D	D	D
V72L	N	D	D	D	D

Tablo 4.9: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteininin bağlayıcı bölgeye ait mutasyonların analiz sonuçları (N: Nötr D: Deleterious (zararlı))

MUTASYON	PROVEAN	PredictSNP	MAPP	SIFT	SNAP
R203K	N	N	N	N	D
G204R	N	D	N	D	D
S235F	N	D	N	D	D
S194L	D	D	D	D	D
T205I	N	N	D	N	N
T198I	N	D	D	D	D
M234I	N	N	N	N	N
G204L	D	D	D	D	D
S202N	N	D	D	N	D
A220V	N	N	D	D	N
A211V	N	N	N	N	N
P199S	N	N	N	N	N
G238V	N	N	-	D	N
N196Y	D	D	D	D	D
S188P	D	D	D	D	D
S193T	N	N	N	D	N
R195I	D	D	D	D	D
P199L	N	N	N	N	N
G212V	N	D	N	D	D
G238C	N	D	-	D	D
G214C	N	D	D	D	N
G215S	N	N	N	N	N
R191L	D	D	D	D	D
M210I	N	N	N	N	N
P207H	N	D	D	D	D
P207S	N	D	N	D	D
A208S	N	N	D	N	N
A211G	N	N	N	N	N
A218V	N	N	D	N	N
G204P	N	D	D	D	D
G204Q	N	N	N	N	N
M210N	N	N	N	N	N
N213S	N	N	N	N	N
Q229R	N	N	D	N	N
Q241L	N	D	D	D	D
R185C	D	D	D	D	D
R185G	D	N	N	N	D
R209S	N	N	N	N	N
S180C	D	D	D	D	D
S201C	D	D	D	D	D
S202C	D	D	D	D	D
S206F	D	D	D	D	D
S232N	N	D	D	D	N
V246A	N	D	D	D	N

Tablo 4.10: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteininin C-Terminal Domainine ait mutasyonların analiz sonuçları (N: Nötr D: Deleterious (zararlı))

MUTASYON	PROVEAN	PredictSNP	MAPP	SIFT	SNAP
T362I	N	N	N	D	D
T334I	D	D	D	D	D
M322I	N	N	N	N	N
I320F	N	N	N	N	N
A308S	N	D	D	D	D
H300Y	N	N	D	D	D
M317T	N	N	N	N	D
P364S	N	D	D	N	D
T271I	D	D	D	D	D
T334A	N	N	D	N	D
V324I	N	N	D	N	N
V350G	N	D	D	D	D
I292L	N	N	N	N	N
I337V	N	N	D	N	N
K299N	N	N	N	D	D
L339F	N	N	N	N	N
P326S	N	N	N	N	N
P364Q	N	D	D	D	D
Q289H	N	D	D	N	N
S327L	D	N	D	D	D
T296I	N	N	N	N	N
V350A	N	N	D	N	D
V350L	N	N	N	N	N

Tablo 4.11: PROVEAN ve PredictSNP ağ sunucu ile elde edilen N proteininin C-kuyruğuna ait mutasyonların analiz sonuçları (N: Nötr D: Deleterious (zararlı))

MUTASYON	PROVEAN	PredictSNP	MAPP	SIFT	SNAP
D377Y	N	D	D	D	D
D401Y	N	D	D	D	D
E378Q	N	N	N	D	N
P383S	N	N	N	D	N
A376T	N	N	N	N	N
P383L	N	N	D	D	N
T379I	N	N	N	N	N
P365S	N	N	N	N	N
A414S	N	D	N	N	D
E378A	N	N	N	D	N
D377A	N	D	N	D	D
Q380L	N	N	D	N	N
Q389H	N	D	N	D	N
S413T	N	N	-	D	N
S416L	N	D	D	D	D
T417I	N	N	N	N	N
K375E	N	N	N	N	N
P365L	N	D	D	D	N
P365Q	N	N	N	N	N
Q389L	N	N	D	D	N
Q390H	N	D	N	N	N
Q418H	N	D	D	D	D
S413I	N	D	-	D	N
T366I	N	D	D	D	N
T379A	N	N	N	N	N
T391I	N	N	N	D	N

4.3 YAPI TEMELLİ ANALİZ SONUÇLARI

Yapı temelli analiz, DynaMut ağ sunucusu ile yapıldı (62). DynaMut ağ sunucusunun sonuçları mail ile elde edildi.

ΔΔG SDM, ΔΔG ENCoM, ΔΔG DUET, ΔΔG mCSM ve ΔΔG DynaMut değerleri için 0 kcal/mol'dan küçük olan değerler yapıyı kararsız hale getirmektedir ve 0 kcal/mol'dan büyük olan değerler ise yapıyı stabilize etmektedir. ΔΔSVib ENCoM değerleri için de 0 kcal/mol'dan küçük olan değerler yapının esnekliğini azaltmakta ve 0 kcal/mol'dan büyük olan değerler ise yapının esnekliğini arttırmaktadır.

4.3.1 STABİLİZASYON ANALİZ SONUÇLARI

DynaMut ağ sunucusunda yapılan protein stabilizasyon analizi sonucunda 103 mutasyonun proteinin 3 boyutlu yapısının stabilizasyonu arttırdığı ve 58 mutasyonun ise protein yapısındaki esnekliğini azalttığı tespit edildi. N proteinin 3 boyutlu protein yapısının esnekliğini arttıran yanlış anlamlı mutasyonların 22 tanesi N-kolunda, 17 tanesi NTD'de, 36 tanesi bağlayıcı bölgede, 10 tanesi de CTD'de ve 18 tanesi C-kuyruğunda bulunmaktadır. N proteinin 3 boyutlu protein yapısının kararsızlaştıran yanlış anlamlı mutasyonların ise 17 tanesi N-kolunda, 12 tanesi NTD'de, 8 tanesi bağlayıcı bölgede, 13 tanesi de CTD'de ve 8 tanesi C-kuyruğunda bulunmaktadır (Tablo 4.12, Tablo 4.13, Tablo 4.15, Tablo 4.16).

Tablo 4.12: DynaMut ağ sunucu ile elde edilen N proteininin N-kolu bölgesine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları

MUTASYON	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
D3L	-0,548	0,194	-0,190	0,380
D3Y	-0,382	-0,065	-0,430	-0,153
P13S	-0,272	-0,745	-0,840	-0,562
A35V	1,13	-0,270	1,150	0,296
Q7K	-0,411	-0,316	0,040	-0,083
G25C	-0,092	-0,966	-0,420	-0,842
Q9H	0,062	-0,500	0,740	-0,199
D22N	0,605	0,564	0,390	0,826
G34W	-1,112	-1,143	-2,140	-1,531
S2Y	1,318	-0,682	0,160	-0,560
N8T	-0,344	-0,207	-0,400	0,032
P13L	1,523	-0,306	-0,040	0,008
R40C	0,137	0,043	-0,400	-0,005
G25V	0,063	-0,494	0,150	-0,146
G44C	-0,843	-1,079	-0,320	-0,938
S23L	-0,459	-0,426	0,940	0,033
S33I	0,216	-0,226	1,370	0,356
A12G	-0,880	-1,353	-0,390	-1,199
A12S	0,413	-1,411	-1,620	-1,419
D22Y	1,464	0,094	0,540	0,267
D3N	-0,217	0,321	-0,070	0,532
D3Q	-0,463	0,311	-0,740	0,465
E31G	-0,094	-0,635	-0,940	-0,683
G19A	0,441	-0,454	0,160	-0,101
N29S	-0,043	0,023	-0,530	0,258
N8Y	0,219	-0,280	-0,180	-0,337
Q7L	0,477	0,017	1,280	0,436
Q9R	0,411	-0,045	0,540	0,303
R10L	0,025	-0,188	0,450	0,052
R14C	-0,223	0,023	-0,240	0,003
R32H	0,233	-0,624	0,120	-0,529
S21P	-0,140	-0,187	-0,050	0,024
S2F	1,005	-0,920	0,360	-0,740
S33G	1,158	-0,142	1,450	0,455
S37P	0,448	-0,162	-0,230	0,024
S37T	0,333	-0,275	0,570	0,189
T16M	-0,316	-0,187	0,620	0,045
T24A	0,195	-0,265	0,610	0,114
Q28K	0,413	-0,053	0,520	0,413

Tablo 4.13: DynaMut ağ sunucu ile elde edilen N proteininin N-Terminal Domainine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları

MUTASYON	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
D128N	-0,402	0,004	-0,200	0,082
P67L	0,244	-0,385	-0,040	-0,201
D128Y	-0,036	0,237	0,120	0,196
L139F	-0,137	-0,938	-0,410	-0,942
A119S	0,640	-0,995	-0,850	-0,971
A90S	0,030	-1,159	1,040	0,938
T135I	1,482	-0,390	1,470	0,255
T166I	0,880	-0,174	1,070	0,316
D144H	0,210	-0,377	-0,040	-0,385
E136Q	0,034	0,347	-0,870	0,367
H145Y	0,809	1,127	-0,400	1,035
L167F	1,096	-0,989	-0,460	-1,006
N126S	-0,080	-0,026	-0,440	0,209
P80R	1,025	0,419	1,850	0,783
A152S	-0,090	-0,202	-0,850	-0,131
A173V	0,848	-0,546	1,740	0,260
D128E	-0,003	-0,357	0,250	-0,080
D63Y	2,299	1,688	0,330	1,484
E118Q	-0,743	-0,109	-0,360	0,001
M101I	0,093	-0,321	0,680	0,396
N140S	-0,279	-0,244	-1,150	-0,199
P142S	-0,546	-0,548	-0,840	-0,394
P151A	-0,129	-0,525	0,510	-0,198
P151S	-0,323	-0,999	-1,050	-1,016
P67S	0,078	-1,091	-0,840	-1,055
Q163K	-0,309	-0,459	-0,430	-0,329
T141I	1,026	-0,036	1,280	0,521
T49I	0,450	0,416	0,650	0,798
V72L	0,348	-0,362	-1,210	-0,325

Tablo 4.14: DynaMut ağ sunucu ile elde edilen N proteininin bağlayıcı bölge bölgesine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları

MUTASYON	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
R203K	0,749	-0,179	0,100	0,007
G204R	1,064	-0,139	-1,910	-0,155
S235F	0,312	-0,968	0,580	-0,820
S194L	0,647	-0,446	2,270	0,313
T205I	0,867	-0,141	1,070	0,371
T198I	0,398	-0,275	1,370	0,169
M234I	0,237	0,053	0,440	0,692
G204L	0,013	-0,441	-1,860	-0,565
S202N	0,681	0,013	0,780	0,348
A220V	0,104	-0,278	-1,030	-0,124
A211V	-0,165	-0,134	0,010	0,130
P199S	0,263	-0,454	-0,580	-0,342
G238V	0,108	-0,408	0,090	-0,069
N196Y	-1,169	-0,641	1,109	-0,542
S188P	0,335	0,492	-1,620	0,375
S193T	0,052	-0,336	0,690	0,070
R195I	0,481	0,002	0,280	0,199
P199L	0,287	-0,336	0,700	0,043
G212V	0,014	-0,274	0,120	0,106
G238C	0,088	-0,960	-0,230	-0,792
G214C	1,523	-0,765	0,530	-0,343
R191L	-0,358	-0,854	0,710	-0,380
G215S	0,187	-0,430	-0,640	-0,136
M210I	0,051	0,198	0,550	0,782
P207H	0,053	-0,015	0,910	0,198
P207S	-0,708	-0,336	-0,260	-0,157
A208S	-0,221	-0,499	-0,850	-0,442
A211G	0,271	-0,051	-0,150	0,180
A218V	0,267	-0,631	-1,460	-0,617
G204P	0,684	-0,327	-3,610	-0,844
G204Q	0,324	-0,444	-2,780	-0,623
M210N	0,814	0,800	0,420	1,168
N213S	0,180	0,148	-0,480	0,463
Q229R	1,086	-0,009	0,340	0,386
Q241L	0,346	0,166	1,170	0,693
R185C	-1,292	-2,110	0,330	-1,721
R185G	0,109	-1,437	0,310	-1,059
R209S	0,307	-0,144	-1,480	0,475
S180C	0,265	-0,076	0,830	0,239
S201C	0,321	-0,212	1,180	0,174
S202C	-0,377	-0,206	0,620	0,048
S206F	0,390	-0,864	0,940	-0,621
S232N	0,153	-0,125	0,590	0,367
V246A	-0,099	-0,963	-0,470	-0,781

Tablo 4.15: DynaMut ağ sunucu ile elde edilen N proteininin C-Terminal Domainine ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları

MUTASYON	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
T362I	0,994	-0,146	1,950	0,606
T334I	-0,065	-0,267	0,640	0,200
M322I	-0,546	-0,498	0,510	0,177
I320F	0,092	-0,872	-0,590	-0,849
A308S	0,159	-0,891	-1,140	-0,774
H300Y	0,502	0,845	-0,390	0,862
M317T	0,316	0,444	-1,390	0,660
P364S	-0,537	-0,588	-0,840	-0,392
T271I	0,569	-0,219	0,890	0,336
T334A	-0,653	-0,311	-0,450	-0,152
V324I	-0,260	-0,684	-0,280	-0,356
V350G	-0,544	-1,871	-0,960	-1,855
I292L	0,144	-1,208	0,200	-0,814
I337V	-0,318	-0,249	-0,120	0,049
K299N	-0,446	-0,288	0,280	-0,112
L339F	-0,134	-1,047	0,120	-0,874
P326S	-0,487	-0,582	-0,260	-0,413
P364Q	-0,247	-0,400	0,640	0,072
Q289H	-0,157	-1,039	0,060	-0,842
S327L	0,753	-0,317	1,030	0,098
T296I	1,195	-0,405	1,930	0,355
V350A	-0,250	-1,430	0,980	-0,843
V350L	0,740	-0,270	0,610	0,280

Tablo 4.16: DynaMut ağ sunucu ile elde edilen N proteininin C-kuyruğuna ait mutasyonların ΔΔG DynaMut, ΔΔG mCSM, ΔΔG SDM ve ΔΔG DUET analiz sonuçları

MUTASYON	ΔΔG DynaMut (kcal/mol)	ΔΔG mCSM (kcal/mol)	ΔΔG SDM (kcal/mol)	ΔΔG DUET (kcal/mol)
D377Y	0,953	-0,389	0,770	-0,046
D401Y	1,253	-0417	0,780	-0,068
E378Q	-0,247	0,027	-1,190	0,017
P383S	0,194	-0,887	0,010	-0,478
A376T	0,088	-0,757	-1,330	-0,623
P383L	0,612	-0,503	1,570	0,248
T379I	-0,278	-0,141	0,900	0,332
P365S	0,211	-0,416	-0,840	-0,214
A414S	-0,115	-0,159	-2,040	-0,148
E378A	-0,256	-0,511	-0,680	-0,558
D377A	-0,062	-0,175	1,850	0,555
Q380L	0,154	0,188	1,400	0,094
Q389H	0,584	-0,410	0,700	-0,239
S413T	0,333	-0,003	0,270	0,477
S416L	-0,308	-0,347	0,980	0,054
T417I	0,043	-0,268	0,900	0,116
K375E	0,351	0,012	0,990	0,621
P365L	0,055	-0,359	-0,040	-0,032
P365Q	0,177	-0,386	-0,810	-0,177
Q389L	0,705	0,227	1,110	0,602
Q390H	-0,322	-0,055	0,850	0,235
Q418H	0,350	-0,212	0,910	-0,001
S413I	0,175	-0,417	1,020	0,168
T366I	0,340	-0,031	0,460	0,314
T379A	-0,330	-0,348	0,680	0,069
T391I	0,245	-0,170	1,030	0,311

4.3.2 ESNEKLİK ANALİZ SONUÇLARI

DynaMut ağ sunucusunda yapılan protein entropi analizi sonucunda; 77 yanlış anlamlı mutasyonun proteinin 3 boyutlu yapısının esnekliğini arttırdığı ve 84 yanlış anlamlı mutasyonun esnekliğini azalttığı tespit edildi. N proteinin üç boyutlu protein yapısının esnekliğini arttıran mutasyonların; 17 tanesi N-kolunda, 11 tanesi, NTD'de 21 tanesi bağlayıcı bölge bölgesinde, 12 tanesi de CTD'de ve 16 tanesi C-kuyruğunda bulunmaktadır. N proteinini üç boyutlu protein yapısının esnekliğini azaltan mutasyonların ise; 22 tanesi N-kolunda, 18 tanesi NTD'de, 23 tanesi bağlayıcı bölgesinde, 11 tanesi de CTD'de ve 10 tanesi C-kuyruğunda bulunmaktadır (Tablo 4.17, Tablo 4.18, Tablo 4.19, Tablo 4.20, Tablo 4.21).

Tablo 4.17: DynaMut ağ sunucu ile elde edilen N proteininin N-kolu bölgesine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları

MUTASYON	ΔΔS ENCoM ($kcal^{-1}/mol^{-1}$)	ΔΔSVibENCoM ($kcal^{-1}/mol^{-1}$)
D3L	0,044	0,044
D3Y	-0,030	-0,030
P13S	0,472	0,472
A35V	-0,174	-0,174
Q7K	0,509	0,509
G25C	-0,022	-0,022
Q9H	-0,365	-0,365
D22N	0,162	0,162
G34W	0,073	0,073
S2Y	-0,692	-0,692
N8T	-0,051	-0,051
P13L	-0,693	-0,693
R40C	0,883	0,883
G25V	0,010	0,010
G44C	0,186	0,186
S23L	-0,034	-0,034
S33I	-0,055	-0,055
A12G	0,420	0,420
A12S	-0,029	-0,029
D22Y	-1,928	-1,928
D3N	0,059	0,059
D3Q	0,042	0,042
E31G	0,616	0,616
G19A	-0,015	-0,015
N29S	0,454	0,454
N8Y	-0,191	-0,191
Q7L	0,843	0,843
Q9R	-0,397	-0,397
R10L	0,838	0,838
R14C	0,852	0,852
R32H	-0,055	-0,055
S21P	-0,358	-0,358
S2F	-0,712	-0,712
S33G	-0,134	-0,134
S37P	-0,565	-0,565
S37T	-0,242	-0,242
T16M	-0,034	-0,034
T24A	0,040	0,040
Q28K	-0,150	-0,150

Tablo 4.18: DynaMut ağ sunucu ile elde edilen N proteininin N-Termianl Domainine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları

MUTASYON	ΔΔS ENCoM ($kcal^{-1}/mol^{-1}$)	ΔΔSVibENCoM ($kcal^{-1}/mol^{-1}$)
V72L	-0,626	-0,626
D128N	0,209	0,209
P67L	-0,169	-0,169
D128Y	0,244	0,244
L139F	-0,197	-0,197
A119S	-0,222	-0,222
A90S	0,008	0,008
T135I	-0,422	-0,422
T166I	-0,226	-0,226
D144H	-0,322	-0,322
E136Q	-0,015	-0,015
H145Y	0,144	0,144
L167F	-0,543	-0,543
N126S	-0,024	-0,024
P80R	-0,540	-0,540
A152S	-0,137	-0,137
A173V	-0,511	-0,511
D128E	0,170	0,170
D63Y	-0,43	-0,943
E118Q	0,390	0,390
M101I	0,419	0,419
N140S	0,030	0,030
P142S	-0,060	-0,060
P151A	0,297	0,297
P151S	-0,055	-0,055
P67S	-0,009	-0,009
Q163K	0,296	0,296
T141I	-0,315	-0,315
T49I	0,152	0,152

Tablo 4.19: DynaMut ağ sunucu ile elde edilen N proteininin bağlayıcı bölge bölgesine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları

MUTASYON	ΔΔS ENCoM ($kcal^{-1}/mol^{-1}$)	ΔΔSVibENCoM ($kcal^{-1}/mol^{-1}$)
R203K	-0,107	-0,107
G204R	-2,522	-2,522
S235F	-0,135	-0,135
S194L	0,044	0,044
T205I	-0,289	-0,289
T198I	-0,070	-0,070
M234I	0,057	0,057
G204L	-0,901	-0,901
S202N	-0,215	-0,215
A220V	-0,018	-0,018
A211V	-0,074	-0,074
P199S	-0,440	-0,440
G238V	-0,041	-0,041
N196Y	0,334	0,334
S188P	0,141	0,141
S193T	0,010	0,010
R195I	0,349	0,349
P199L	-0,021	-0,021
G212V	0,003	0,003
G238C	-0,004	-0,004
G214C	-1,357	-1,357
G215S	-0,259	-0,259
R191L	0,134	0,134
M210I	0,469	0,469
P207H	-0,051	-0,051
P207S	-0,010	-0,010
A208S	0,023	0,023
A211G	0,088	0,088
A218V	-0,388	0,388
G204P	-0,572	-0,572
G204Q	-0,938	-0,938
M210N	0,810	0,810
N213S	-0,162	-0,162
Q229R	-0,360	-0,360
Q241L	0,089	0,089
R185C	1,042	1,042
R185G	1,753	1,753
R209S	1,731	1,731
S180C	-0,059	-0,059
S201C	0,176	0,176
S202C	0,315	0,315
S206F	0,016	0,016
S232N	0,046	0,046
V246A	0,311	0,311

Tablo 4.20: DynaMut ağ sunucu ile elde edilen N proteininin C-Termianl Domainine ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları

MUTASYON	ΔΔS ENCoM ($kcal^{-1}/mol^{-1}$)	ΔΔSVibENCoM ($kcal^{-1}/mol^{-1}$)
T334A	0,149	0,149
T362I	-0,230	-0,230
T334I	0,267	0,267
M322I	0,175	0,175
I320F	-0,680	-0,680
A308S	-0,104	-0,104
H300Y	-0,175	-0,175
M317T	0,342	0,342
P364S	0,003	0,003
T271I	-0,398	-0,398
V324I	-0,114	-0,114
V350G	0,296	0,296
I292L	0,079	0,079
I337V	0,464	0,464
K299N	0,383	0,383
L339F	0,130	0,130
P326S	0,012	0,012
P364Q	-0,024	-0,024
Q289H	-0,050	-0,050
S327L	-0,122	-0,122
T296I	-0,422	-0,422
V350A	0,265	0,265
V350L	-0,181	-0,181

Tablo 4.21: DynaMut ağ sunucu ile elde edilen N proteininin C-kuyruğuna ait mutasyonların ΔΔS ENCoM ve ΔΔSVib ENCoM analiz sonuçları

MUTASYON	ΔΔS ENCoM ($kcal^{-1}/mol^{-1}$)	ΔΔSVibENCoM ($kcal^{-1}/mol^{-1}$)
D377Y	-0,545	-0,545
D401Y	-0,487	-0,487
E378Q	0,072	0,072
P383S	-0,006	-0,006
A376T	-0,041	-0,041
P383L	-0,058	-0,058
T379I	-0,043	-0,043
P365S	0,115	0,115
A414S	-0,031	-0,031
E378A	0,217	0,217
D377A	0,249	0,249
Q380L	0,094	0,094
Q389H	-0,119	-0,119
S413T	0,143	0,143
S416L	0,086	0,086
T417I	0,073	0,073
K375E	0,127	0,127
P365L	0,131	0,131
P365Q	0,141	0,141
Q389L	0,010	0,010
Q390H	0,019	0,019
Q418H	-0,222	-0,222
S413I	0,077	0,077
T366I	-0,238	-0,238
T379A	0,143	0,143
T391I	0,052	0,052

4.4 ANAHTAR MUTASYONLARIN PROTEİN YAPISINA ETKİSİNİN SONUÇLARI

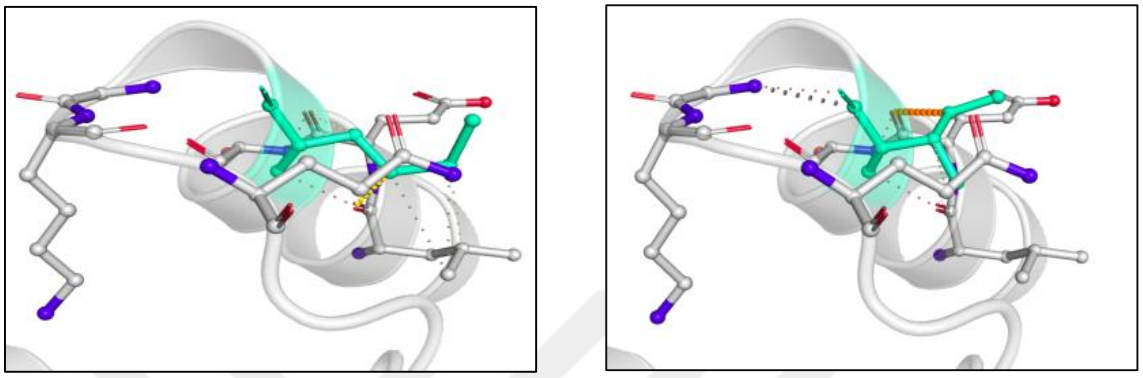
Çalışmamızda Türk populasyonunda görülen D3L, M234I, A220V, S194L, S235F, T362I ve D128N mutasyonları ele alındı. Titreşimsel entropi ve protein stabilitesi üzerindeki mutasyon etkisi gibi protein yapısı stabilite parametreleri de DynaMut ağ sunucusu tarafından tahmin edildi. Kırmızı renk hidrojen bağlarını, turuncu renk zayıf hidrojen bağlarını, mavi renk halojen bağlarını, sarı renk iyonik etkileşimleri, açık mavi renk aromotik etkileşimi, yeşil hidrofobik teması, pembe karbonil etkileşimlerini, gri renkte ise Van der waals etkileşimlerini göstermektedir. Doğal tip ve mutant yan zincirler açık yeşil renktedir. Aralarındaki etkileşimler ve bağlar kesik çizgi ile gösterilmektedir. Mavi renk yapının sertleşmesini, kırmızı renk ise esneklikte bir artış olduğunu göstermektedir (Şekil 4.1).

Bağ ve Etkileşim Tipleri	Renk	Bağ ve Etkileşim Tipleri	Renk
Hidrojen bağı		İyonik etkileşimler	
Su aracılı hidrojen bağı		Metal kompleks etkileşimler	
Zayıf hidrojen bağı		Aromatik etkileşimler	
Su aracılı zayıf hidrojen bağı		Hidrofobik etkileşimler	
Halojen bağı		Kabonil etkileşimler	

Şekil 4.1: Bağ ve etkileşim tiplerinin gösterildiği renkler (63)

4.4.1 Kimyasal Bağ Ve Etkileşim Analiz Sonuçları

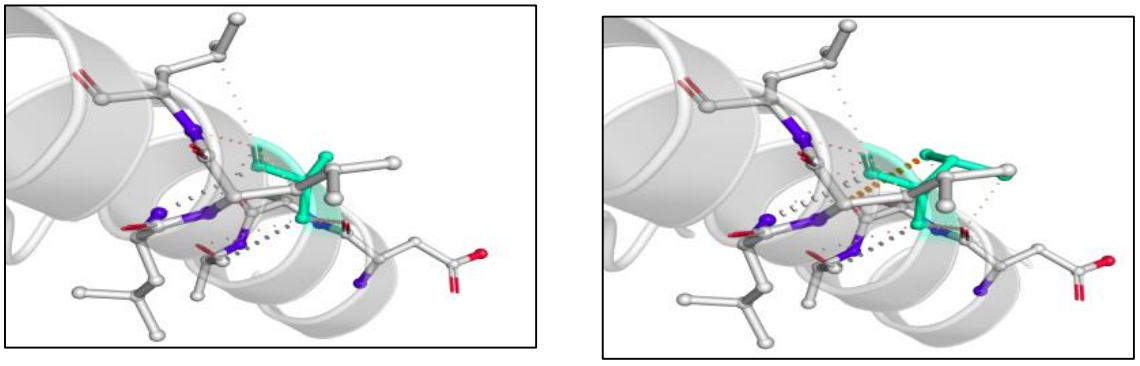
DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde M234I mutasyonu sonucunda iyonik bağ etkileşiminin bozulduğu ve mutant yapıda hidrojen bağı ve hidrofobik etkileşimlere yol açtığı tespit edildi (Şekil 4.2).



Doğal tip yapısı Mutant Yapı

Şekil 4.2: M234I mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

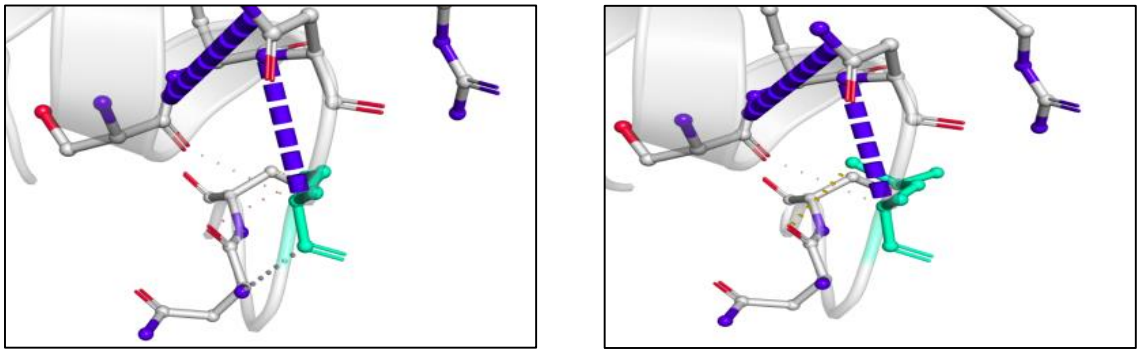
DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde A220V mutasyonu sonucunda mutant yapıda hidrojen bağı ve iyonik bağ oluşumu tespit edildi (Şekil 4.3).



Doğal tip yapısı Mutant Yapı

Şekil 4.3: A220V mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

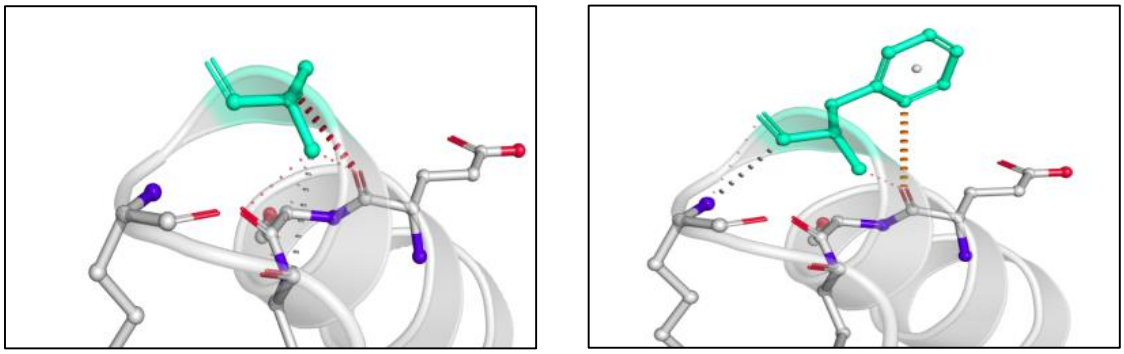
DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde S194L mutasyonu sonucunda doğal tip yapısındaki karbonil bağ ve iyonik etkileşiminin bozulduğu ve mutant yapıda hidrojen bağı oluşumu tespit edildi (Şekil 4.4).



Doğal tip yapısı Mutant Yapı

Şekil 4.4: S194L mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

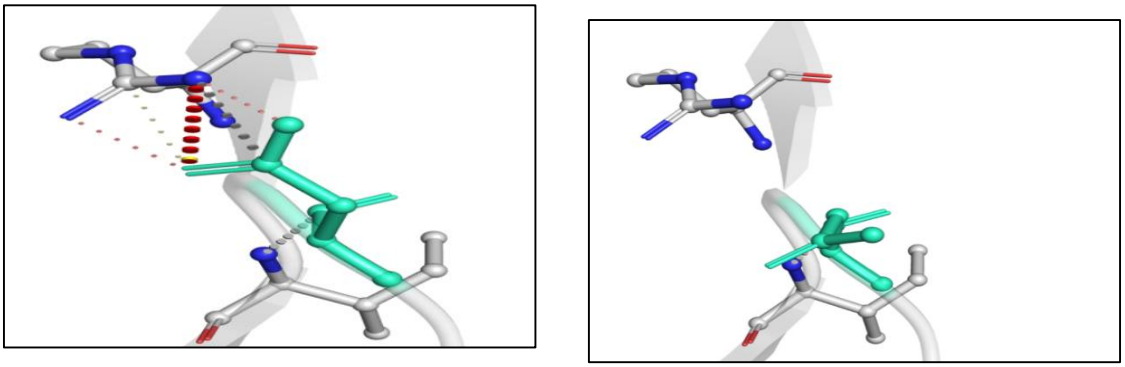
DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde S235F mutasyonu sonucunda doğal tip yapıdaki hidrojen bağının bozulduğu ve mutant yapıda zayıf hidrojen oluşumu tespit edildi (Şekil 4.5).



Doğal tip yapısı Mutant Yapı

Şekil 4.5: S235F mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

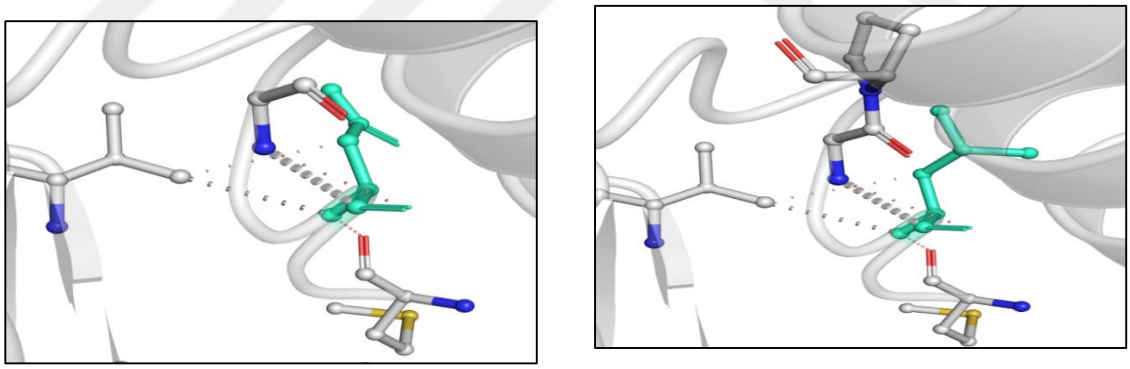
DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde D128N mutasyonu sonucunda doğal tipte bulunan bağ ve etkileşimlerinin mutant tipte bozulduğu tespit edildi (Şekil 4.6).



Doğal tip yapısı Mutant Yapı

Şekil 4.6: D128N mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

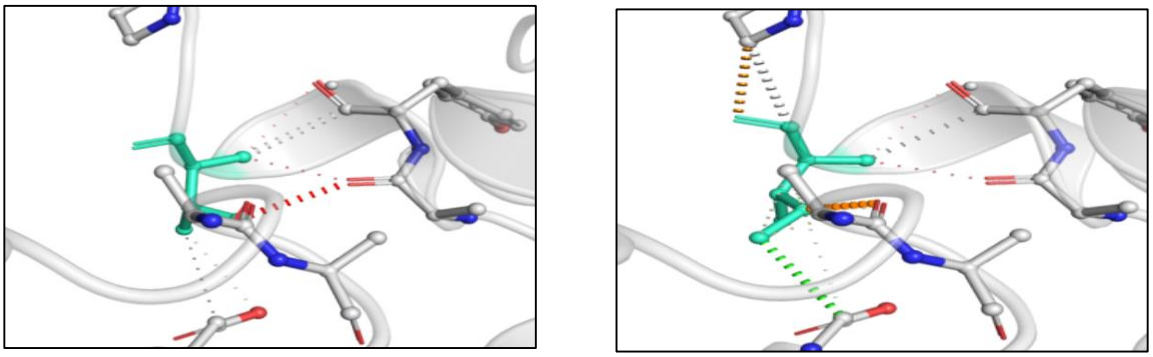
DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde D3L mutasyonu sonucunda bir değişim tespit edilmedi (Şekil4.7).



Doğal tip yapısı Mutant Yapı

Şekil 4.7: D3L mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde T362I mutasyonu sonucunda doğal tip yapıdaki zayıf hidrojen bağının bozulduğu ve mutant yapıda hidrojen bağı ile hidrofobik etkileşim oluşumu tespit edildi (Şekil 4.8).

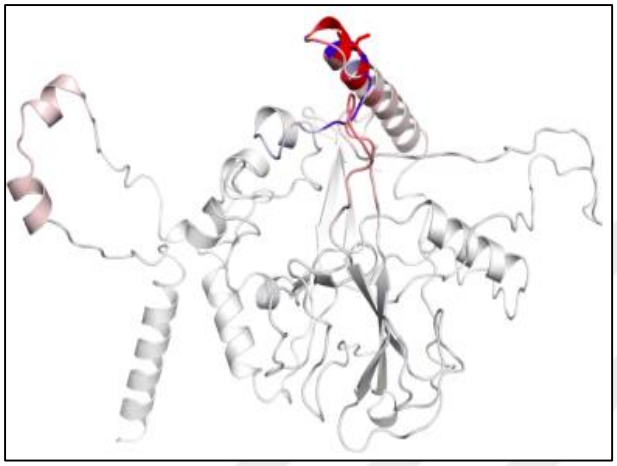


Doğal tip yapısı Mutant Yapı

Şekil 4.8: T362I mutasyonunun doğal tip ve mutant yapıdaki kimyasal bağ ve etkileşim analizi sonucu (62)

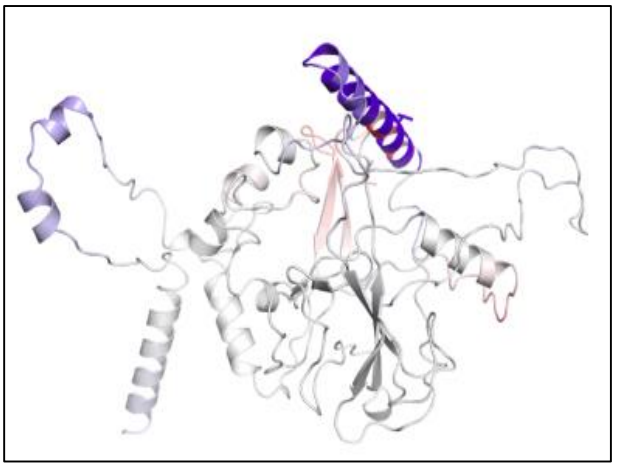
4.4.2 Entropi Enerji Değişim Analizi Sonuçları

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda, M234I mutasyonun N proteinin esnekliğini arttırdığı tespit edildi (ΔΔSVibENCoM: 0,057 kcal.mol^{-1}.K^{-1}) (Şekil 4.9).



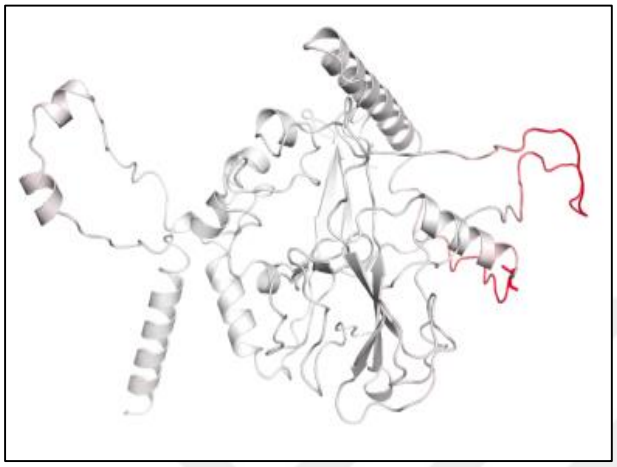
Şekil 4.9: M234I mutasyonunun entropi enerji değişimi analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda A220V mutasyonunun, N proteinin esnekliğini azalttığı tespit edildi. (ΔΔSVibENCoM: -0,018 kcal.mol^{-1}.K^{-1}) (Şekil 4.10).



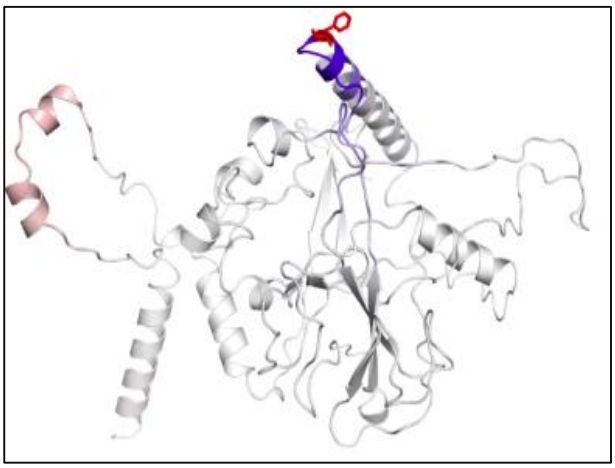
Şekil 4.10: A220V mutasyonunun entropi enerji değişimi analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda S194L mutasyonunun, N proteinin esnekliğini arttırdığı tespit edildi. (ΔΔSVibENCoM: 0,044 kcal.mol^{-1}.K^{-1}) (Şekil 4.11).



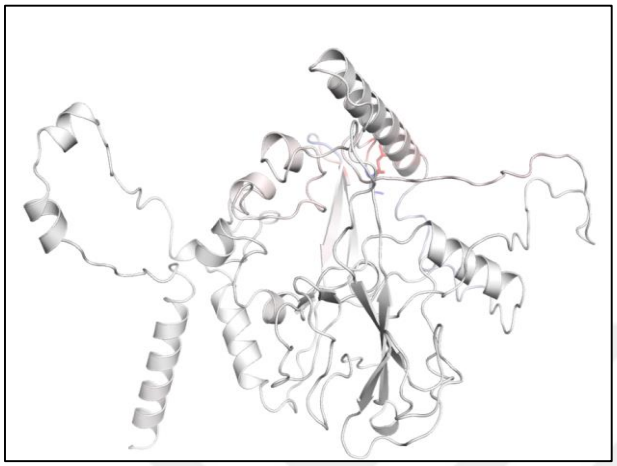
Şekil 4.11: S194L mutasyonunun entropi enerji değişimi analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda S235F mutasyonu N proteinin esnekliğini azalttığı tespit edildi. (ΔΔSVibENCoM: -0,135 kcal.mol^{-1}.K^{-1}) (Şekil 4.12).



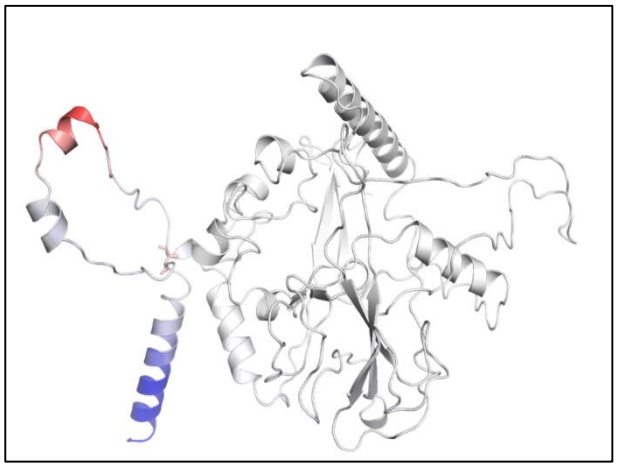
Şekil 4.12: S235F mutasyonunun entropi enerji değişimi analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda D3L mutasyonu N proteinin esnekliğini arttırdığı tespit edildi. (ΔΔSVibENCoM: 0,044 kcal.mol-1.K-1) (Şekil 4.13).



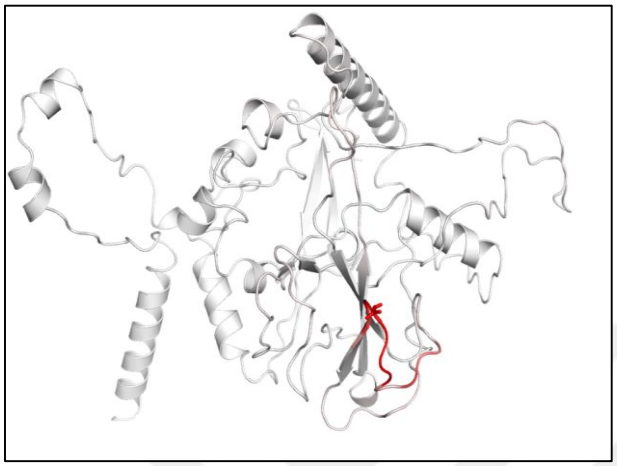
Şekil 4.13: D3L mutasyonunun entropi enerji değişimi analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda T362I mutasyonu N proteinin esnekliğini azalttığı tespit edildi. (ΔΔSVibENCoM: -0,230 kcal.mol-1.K-1) (Şekil 4.14).



Şekil 4.14: T362I mutasyonunun entropi enerji değişimi analizi sonucu (62)

DynaMut ağ sunucusu ile yapılan entropi enerji değişimi analizi sonucunda D128N mutasyonu N proteinin esnekliğini arttırdığı tespit edildi. (ΔΔSVibENCoM: 0,209 kcal.mol-1.K-1) (Şekil 4.15).

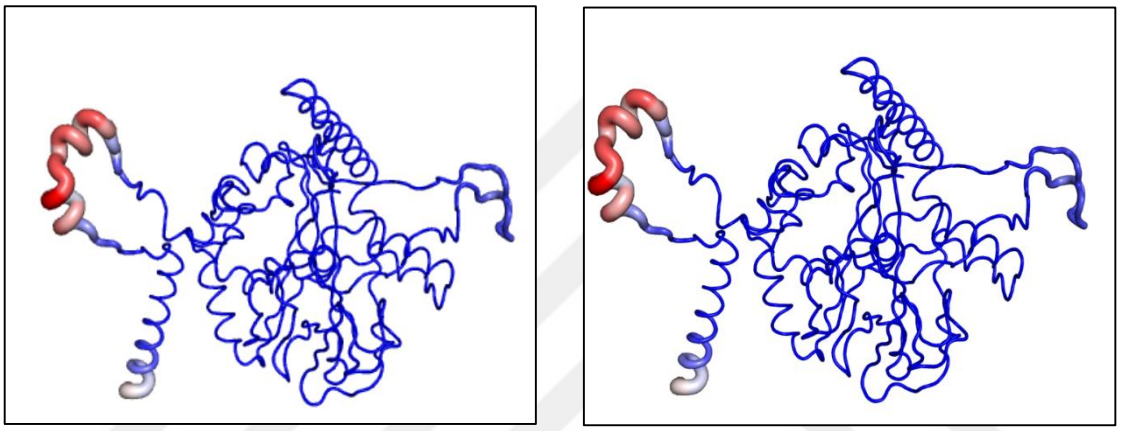


Şekil 4.15: D128N mutasyonunun entropi enerji değişimi analizi sonucu (62)

4.4.3 Atomik Kararsızlık Analiz Sonuçları

Atomik kararsızlık, mutlak atomik hareketin genliğini sağlamaktadır. Molekülün ilk 10 modu üzerinden hesaplama yapılmaktadır. Kararsızlığın büyüklüğü iki şekilde gösterilmektedir. İlk olarak ince ve kalın tüp şeklinde gösterimidir. İkinci olarak ise renklere bağlı gösterimidir. Mavi renk düşük kararsızlığı, beyaz renk orta kararsızlığı ve kırmızı renk ise yüksek kararsızlığı temsil etmektedir.

DynaMut ağ sunucusu ile yapılan atomik kararsızlık analizi sonucunda, doğal ve mutant tip arasında bir değişiklik tespit edilmedi (Şekil 4.16).



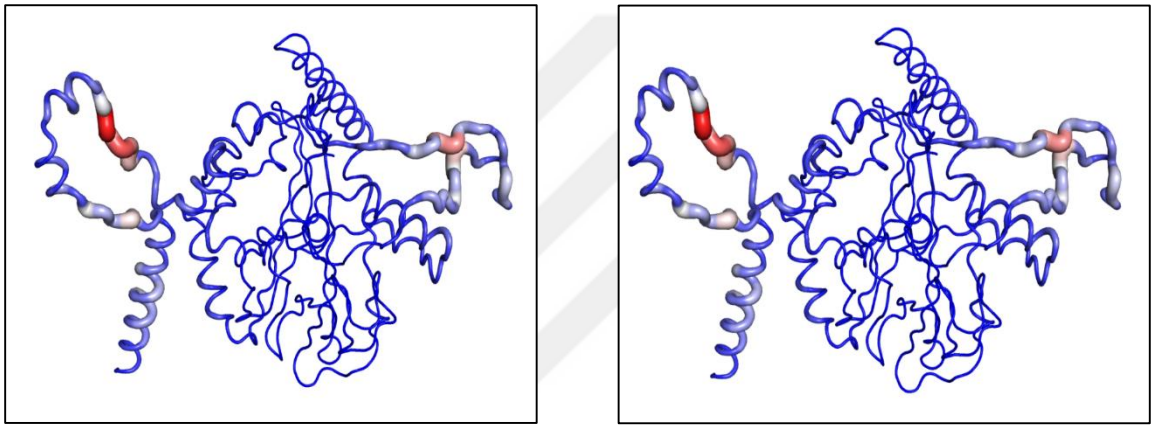
Doğal tip yapısı Mutant Yapı

Şekil 4.16: Atomik kararsızlık analiz sonucu (62)

4.4.4 Deformasyon Enerjilerinin Analiz Sonuçları

Deformasyon enerjisi, proteindeki yerel esneklik miktarı için bir ölçü sağlamaktadır. Molekülün ilk 10 modu üzerinden hesaplama yapılmaktadır. Deformasyon enerjilerinin büyüklüğü iki şekilde gösterilmektedir. İlk olarak ince ve kalın tüp şeklinde gösterimidir. İkinci olarak ise renklere bağlı gösterimidir. Mavi renk düşük deformasyon enerjisini, beyaz renk orta deformasyon enerjisini ve kırmızı renk ise yüksek deformasyon enerjisini temsil etmektedir.

DynaMut ağ sunucusu ile yapılan atomik kararsızlık analizi sonucunda, doğal ve mutant tip arasında bir değişiklik tespit edilmedi (Şekil 4.17).



Doğal tip yapısı Mutant Yapı

Şekil 4.17: Deformasyon enerjilerinin analiz sonucu (62)

5. TARTIŞMA

Dünya tarihindeki en büyük salginlardan biri olan COVID-19 pandemisine neden olan SARS-CoV-2 virüsünün yapısal proteinleri bilim dünyasının odağında olmayı sürdürmektedir (38). Coronaviridae virüs ailesinde N proteini, viral RNA transkripsiyonu ve replikasyonu için gerekli olan çok işlevli bir RNA bağlayıcı proteindir (16). Dünya tarihindeki en büyük salgınlardan biri olan COVID-19 pandemisine neden olan SARS-CoV-2 virüsünün yapısal proteinleri bilim dünyasının odağında olmayi sürdürmektedir.

Bazı araştırmalar, N ve E proteini arasındaki etkileşimin SARS-CoV virüsününün hücreden çıkışında bir işlevi olduğunu ve koronavirüsün oluşumu için N–M protein kompleksinin gerekli olduğunu göstermektedir (63-66). Bu nedenle, her SARS-CoV-2 N protein mutantı ile E proteini arasındaki etkileşim, virüs salınımını artırabilirken, S194L N-M protein kompleksinin azalan bağlanma afinitesi virüs oluşumunu zayıflatabilir (42). Çalışmamızdaki yapıya bağlı protein stabilizasyonu analizi sonucunda; ΔΔG DynaMut değerleri S194L mutasyonunun N proteinin 3B yapısının stabilizasyonunu attırdığı (Tablo 4.14) ve ΔΔSvib ENCoM değeri de N proteinin 3B yapısının esnekliğini arttırdığı (Şekil 4.10) tespit edildi. Sekansa bağlı mutasyon patojenite analizi sonucunda ise S194L mutasyonu deleterious (zararlı) olarak (Tablo 4.9) tespit edildi. DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde S194L mutasyonu sonucunda doğal tip yapısındaki karbonil bağ ve iyonik etkileşiminin bozulduğu ve mutant yapıda hidrojen bağı oluşumu tespit edildi (Şekil 4.5). Yapılan analizler sonucunda S194L mutasyonunun N-M protein kompeksinin bağlanma afinitesine ve N ile E protein arasındaki etkileşime etki edeceği düşünülmektedir. Yapılmış olan in siliko analizler konu ile ilgili yapilacak olan in vitro ve in vivo çalışmalara yön vermesi açısından önemli görüldüğü gibi tespit edilen yapısal değişikliklerin bu çalışmalarla bulguları güçlendirmesi beklenmektedir.

Yapılan çalışmalarda, SR açısından zengin olan bağlayıcı bölgenin, esas olarak serin yan zincirlerinde fosforilasyon yoluyla hücre içi sinyalleşme yolaklarında önemli bir rol oynadığı bulundu (67-69). Çalışmamızda incelenen mutasyonların 43 tanesi bu bölgede tespit edilmiştir (Tablo 4.9). Çalışmamız sonucunda; S194L, G204L, N196Y, S188P, R195I, R191L, R185C, S180C, S201C, S202C ve S206F mutasyonları deleterious (zararlı) olarak (Tablo 4.9) tespit edildi. DynaMut ağ sunucusu ile yapılan analizde; G204P, G204Q mutasyonlarının ΔΔG DUET değerleri N proteinin 3B yapısını kararsız

hale getirdiği (Tablo 4.14) ve ΔΔSVib ENCoM değerleri sonucunda yapının esnekliğinin azaldığı (Tablo 4.19) tespit edildi. R185C mutasyonunun, ΔΔG DynaMut değeri sonucunda N proteinin 3B yapısını kararsız hale getirdiği (Tablo 4.14) görülmüştür. Bunun sonucunda, serin yan zincirlerindeki bu mutasyonların fosforilasyona bağlı sinyalleşmeyi etkileyebileceği düşünülmektedir.

Nükleokapsid proteinindeki NTD ve CTD, RNA bağlamasında görev aldığı bilinmektedir (70). Yapılan bir çalışmada PROVEAN ağ sunucusu kullanılarak N proteininin CTD alanında 13 ve NTD alanında 12 deleterious (zararlı) mutasyon tespit edildi (71). NTD'de bulunan D128Y ile P151S (Tablo 4.8) ve CTD'de bulunan T334I ile T271I (Tablo 4.10) mutasyonlarının deleterious (zararlı) tespiti bizim çalışmamızla tutarlıdır. Ayrıca çalışmamız PredictSNP ağ sunucu ile yapılan deleterious(zararlı) mutasyon tespiti sonucunda; NTD'de 6 ve CTD'de 5 adet deleterious(zararlı) mutasyon tespit edildi. Bunun, SARS-CoV-2 virüsünde RNA paketlemesinde sorunlara yol açabileceği tahmin edilmektedir.

Azad yaptığı bir çalışmada, B ve T hücre epitoplarından birinin N proteininin 305-340 yan zincirleri arasında yer aldığını ortaya koymaktadır. 305-340 yan zincirleri arasında yer alan mutasyonlar nedeniyle, konakçının immünolojik tepkisini etkileyebilecek epitop özellikleri değişebilmektedir (11). Çalışmamızda 305-340 aminoasitleri arasında CTD'de bulunan 11 adet mutasyon tespit edildi. Bu mutasyonlardan T334I, PROVEAN ve PredictSNP ağ sunucuları ile yapılan analiz sonucunda deleterious (zararlı) olarak tespit edildi. DynaMut ağ sunucusu ile yapılan analiz sonucunda; T334A, V234I ve P326S mutasyonların N proteinin 3B yapısını kararsızlaştırdığı tespit edildi. Bu nedenle, SARS-COV-2 N proteinini hedef alan aşıların ve ilaçların geliştirilmesinde, çeşitli popülasyonlarda ve bölgelerde meydana gelen mutasyonların göz önünde bulundurulması gerektiği düşüncesini güçlendirmektedir. Çalışmamızda bu fikri desteklemektedir.

Düzensiz bölgeler, iyi çözümlenmiş bir üçüncül yapıya sahip değildir; bununla birlikte, koronavirüs N proteinlerinin IDR'leri, viral genomik RNA ile bağlanmada önemli bir rol oynadığı bulunmuştur (72, 73). Çalışmamızda IDR bölgelerinde bulunan 108 adet mutasyon tespit edildi (Tablo 4.1, Tablo 4.3, Tablo 4.5). PROVEAN ve PredictSNP ağ sunucuları ile yapılan analiz sonucunda iki sunucuda da ortak olarak S194L, G204L, N196Y, S188P, R195I, R191L, R185C, S180C, S201C, S202C ve S206F

mutasyonları deleterious (Tablo 4.9) olarak tespit edildi. ΔΔG DUET değeri sonucunda G204L mutasyonunun N proteinin yapısnının stabilitesini azalttığını, R195I mutasyonunun N proteinin yapısını stabilitesini arttırdığı tespit edildi (Tablo 4.14). Ayrıca R195I mutasyonunun ΔΔSVib ENCoM değeri sonucunda N proteinin 3B yapısının esnekliğini arttırdığı tespit edildi (Tablo 4.19). Bu mutasyonların, viral genomun N proteinine bağlanma stabilitesini değiştirebileceği düşünülmektedir.

Bir çok çalışma, koronavirüs N proteinlerinin aynı zamanda RNA şaperonları gibi davrandığını ve RNA genomunun uygun şekilde katlanmasına yardımcı olduğunu bildirmiştir. Bu nedenle, N protein fonksiyonuna müdahale edebilen ilaçlar, farmakolojik açıdan büyük ilgi görmektedir. N proteininin oligomerizasyonu sağlayan CTD bölgesini engelleyen ilaçlar, virüs oluşumunu inhibe etmek için umut vadeden adaylardır. SARS-CoV-2 N proteini ile yakın zamanda yapılan bir çalışmada; Conivaptan, Ergotamine, Venetoclax, Rifapentine ve diğerleri dahil olmak üzere işlevini engelleyebilecek 10 adet umut verici ilaç tanımladı (73-77) Çalışmamızda N proteinin CTD'sinde T334I ve T271I mutasyonları PROVEAN ile PredictSNP ağ sunucularıyla yapılan analiz sonucunda ortak olarak deleterious olarak tespit edildi (Tablo 4.10). Bu bölgede tespit edilen mutasyonlar ilaç-protein etkileşimi ile ilgili yapılacak olan çalışmalarda hedef olma özelliğine sahiptir.

Hu ve arkadaşlarının yaptıkları çalışmada, yaklaşık 3,142 klinik aşamadaki ve FDA onaylı küçük molekülleri N proteinini hedefleyen aday terapötik inhibitörleri belirlemek için kullanmışlardır. Ceftriaxone, arbidol ve Lopinavir ilaçlarının N proteininde etkileşim yerleri belirlenmiştir. Arbidol N proteinin NTD bölgesine, Lopinavir ise CTD bölgesine bağlandığı bulunmuştur (78). Çalışmamızda lopinavir ilacının etkileştiği T334 yan zincirinde mutasyon tespit edildi (Tablo 4.4) ve ΔΔG DynaMut değeri sonucunda bu mutasyonun N proteini kararsızlaştırdığı (Tablo 4.15), ΔΔSVib ENCoM değeri sonucunda N proteinin 3B yapısının esnekliğini arttırdığı tespit edildi (Tablo 4.20). Ayrıca arbidol ilacının etkileştiği T166 yan zinciriyle L167 yan zincirinde mutasyon tespit edildi (Tablo 4.2) ve ΔΔSVib ENCoM değeri sonucunda bu mutasyonların N proteinin yapısının esnekliğini azalttığı bulundu (Tablo 4.18). Bu spesifik yapısal alanlarda bulunan mutasyonların, bu alanlarla bahsi geçen küçük moleküllerin etkileşimlerinin ne şekilde değişeceğinin tespit edilmesi önemlidir.

Yapılan başka bir çalışmada ise 4E1RCat, Silmitasertib, TMCB, Sapanisertib ve Rapamycin adlı 5 antiviral ilacın CTD'ye yönelik potansiyel bağlanma ve inhibitör

aktivite mekanizmalarını araştırılmıştır. Tüm ilaçlar için yerleştirme hesaplamalarından sonra elde edilen sonuçlar, CTD ile H-bağları oluşturan en yaygın yan zincirlerin Arg277, Asn285 ve Tyr333 olduğunu tespit edilmiştir (79). Çalışmamızda, bu yan zincirlerin hiç birinde mutasyon tespit edilmedi.

Vılar ve Isom'un yaptığı bir çalışmada, 290,000 tam SARS-CoV-2 proteom genomu ile çalışılmış ve çeşitli proteinlere ait mutasyonlar analiz edilmiştir. Bu çalışmada en sık görülen R203K ve G204R mutasyonlarının, yan zincir pozisyonlarında ilk formuna geri döndüğünü tespit edilmiştir. Çalışmamızda da ele alınan A220V, D3L, S235F, M234I ve S194L mutasyonları son zamanlarda SARS-CoV-2 genomunda daha çok tespit edilmekte ve bunun çalışılmasının önemli olduğu vurgulanmaktadır (51). Bu sebeple çalışmamızda, D3L, M234I, A220V, S194L, S235F mutasyonları ile NTD ve CTD'de en sık görülen T362I ve D128N mutasyonları anahtar mutasyonlar olarak seçildi. D3L, M234I, A220V, S194L, S235F, T362I ve D128N mutasyonlarının; kimyasal bağ ve etkileşim analizi, N proteinin atomik kararsızlık analizi ile deformasyon enerjilerinin analizi için çevrimiçi bir program sunucusu olan DynaMut ağ sunucusu kullanıldı.

Yapılan çalışmalarda N proteinde; R203K ve G204R en sık tespit edilen mutasyonlardır (Tablo 4.3) (80-82). Bizim bulgularımız da bu sonucu desteklemektedir.

Ul Alam ve arkadaşlarının yaptıkları çalışma sonucunda, A220V mutasyonunun RNA bağlanmasını etkileyebileceği bildirilmiştir (83). Çalışmamızda A220V mutasyonu sonucunda mutant yapıda hidrojen bağı ve iyonik bağ oluşumu tespit edildi (Şekil 4.3). ΔΔSVibENCoM sonucunda da A220V mutasyonunun, N proteinin esnekliğini azalttığı tespit edildi (Şekil 4.10). Bu sonuçların proteinin RNA bağlanmasını ne şekilde etkileyeceği ileri bir analizin konusu olmalıdır.

Nükleokapsid proteinindeki D3L ve S235F mutasyonları, RNA ile etkileşen domainlerin dışında bulunmaktadır. Yapılan çalışmada, RNA etkileşimi için S235F mutasyonunun N proteinine stabilizasyon etkisi sağladığı öngörüldü (84). DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde D3L mutasyonu sonucunda bir değişim tespit edilmedi (Şekil 4.7). S23F mutasyonunun ΔΔSVib ENCoM değeri sonucunda N proteinin 3B yapısının esnekliğini azalttığı tespit edildi (Şekil 4.12). S235F mutasyonu sonucunda doğal tip yapıdaki hidrojen bağının bozulduğu ve mutant yapıda zayıf hidrojen oluşumu tespit edildi (Şekil 4.5). Bizim çalışmamızın sonuçları bu çalışma sonuçları ile örtüşmemektedir.

Yapılan bir çalışmada T362I mutasyonunun CTD'deki ikincil yapıyı değiştirdiğini ve dimerizasyon potansiyelini etkileyebileceğini bildirmişlerdir (73). DynaMut ağ sunucusu ile yapılan kimyasal bağ ve etkileşim analizinde T362I mutasyonu sonucunda doğal tip yapıdaki zayıf hidrojen bağının bozulduğu ve mutant yapıda hidrojen bağı ile hidrofobik etkileşim oluşumu tespit edildi (Şekil 4.8). T362I mutasyonunun ΔΔSVibENCoM değerine göre N proteinin esnekliğini azalttığı tespit edildi (Şekil 4.14). N proteininde dimerizasyonu etkileyeceği tahmin edilmektedir ancak ileri analizlerin yapılması gerekmektedir.

M234I ve D128N mutasyonlarının kimyasal bağ ve etkileşim analiz sonuçlarını desteleyecek ya da reddecek bir çalışma literatürde bulunmamaktadır.

Çalışmamız SARS-CoV-2 virüsündeki N proteininde oluşan mutasyonların, replikasyondaki ve patogenezdeki kesin rolünü araştırmak ve sonrasında geliştirilecek olan aşı ile ilaç için in-siliko ve ıslak laboratuvar çalışmalarına daha fazla ihtiyaç olduğunu göstermektedir.

6. SONUÇLAR VE ÖNERİLER

Genel olarak bulgularımız, SARS-CoV-2 N proteininde meydana gelen mutasyonların yapısını ve bağlanma afinitesini değiştirebileceğini ortaya koymaktadır. Bu yapısal değişikliklerin SARS-CoV-2 virüsüsün bulaşması, teşhisi ve COVID-19 tedavisi üzerinde etkileri analiz edilmiştir.

Kullandığımız bütünleştirici biyoinformatik yöntemler, ıslak laboratuvar yaklaşımlarını ve klinik çalışmaları kullanan araştırma toplulukları için, zamandan ve maliyetten önemli ölçüde tasarruf sağlamıştır. Bulgularımız COVID-19 terapötiklerine duyulan ihtiyacın karşılanması için yapılan çalışmalar için önemlidir. COVID pandemisini yavaşlatmayı ve bitirmeyi amaçlayan aşı ve ilaç geliştirme çalışamalarına katkısı olduğu düşünülen bulgularımız in vitro ve in vivo çalışamları ile valide edilmesine ihtiyaç duyulmaktadır.

.

KAYNAKLAR

1. Ortega, M. A., Fraile-Martínez, O., García-Montero, C., García-Gallego, S., Sánchez-Trujillo, L., Torres-Carranza, D., Álvarez-Mon, M. Á., Pekarek, L., García-Honduvilla, N., Bujan, J., Álvarez-Mon, M., Asúnsolo, Á., de la Torre, B. (2021). An integrative look at SARS-CoV-2 (Review). *International Journal of Molecular Medicine*, *47*(2), 415–434.
2. Sofi, M. S., Hamid, A., Bhat, S. U. (2020). SARS-CoV-2: A critical review of its history, pathogenesis, transmission, diagnosis and treatment. *Biosafety and Health*, *2*(4), 217–225.
3. Hu, B., Guo, H., Zhou, P., Shi, Z. L. (2021). Characteristics of SARS-CoV-2 and COVID-19. *Nature Reviews Microbiology*, *19*(3), 141–154.
4. Arya, R., Kumari, S., Pandey, B., Mistry, H., Bihani, S. C., Das, A., Prashar, V., Gupta, G. D., Panicker, L., Kumar, M. (2021). Structural insights into SARS-CoV-2 proteins. *Journal of Molecular Biology*, *433*(2), 166725.
5. Haque, S. K. M., Ashwaq, O., Sarief, A., Azad John Mohamed, A. K. (2020). A comprehensive review about SARS-CoV-2. *Future Virology*, *15*(9), 625–648.
6. Rahimi, A., Mirzazadeh, A., Tavakolpour, S. (2020). Genetics and genomics of SARS-CoV-2: A review of the literature with the special focus on genetic diversity and SARS-CoV-2 genome detection. *Genomics* 113, 1221–1232
7. Phan, T. (2020). Genetic diversity and evolution of SARS-CoV-2. *Infection, Genetics and Evolution*, *81*(February), 104260.
8. Wang, C., Liu, Z., Chen, Z., Huang, X., Xu, M., He, T., Zhang, Z. (2020). The establishment of reference sequence for SARS-CoV-2 and variation analysis. *Journal of Medical Virology*, *92*(6), 667–674.
9. Wen, F., Yu, H., Guo, J., Li, Y., Luo, K., Huang, S. (2020). Identification of the hyper-variable genomic hotspot for the novel coronavirus SARS-CoV-2. *Journal of Infection*, *80*(6), 671–693.
10. Liu, S. J., Leng, C. H., Lien, S. P., Chi, H. Y., Huang, C. Y., Lin, C. L., Lian, W. C., Chen, C. J., Hsieh, S. L., Chong, P. (2006). Immunological characterizations of the nucleocapsid protein based SARS vaccine candidates. *Vaccine*, *24*(16), 3100–3108.
11. He, R., Dobie, F., Ballantine, M., Leeson, A., Li, Y., Bastien, N., Cutts, T., Andonov, A., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., Baker, L., Li, X. (2004). Analysis of multimerization of the SARS coronavirus nucleocapsid protein. *Biochemical and Biophysical Research Communications*, *316*(2), 476–483.
12. Huang, Q., Yu, L., Petros, A. M., Gunasekera, A., Liu, Z., Xu, N., Hajduk, P., Mack, J., Fesik, S. W., Olejniczak, E. T. (2004). Structure of the N-terminal RNA-binding domain of the SARS CoV nucleocapsid protein. *Biochemistry*, *43*(20), 6059–6063.
13. Laha, S., Chakraborty, J., Das, S., Kanti, S. (2020). Characterizations of SARS-CoV-2 mutational profile, spike protein stability and viral transmission. *Infection, Genetics and Evolution* 85, 10444
14. Duffy, S. (2018). Why are RNA virus mutation rates so damn high? *PLoS Biology*, *16*(8), 1–6. https://doi.org/10.1371/journal.pbio.3000003
15. Teng, S., Sobitan, A., Rhoades, R., Liu, D., Tang, Q. (2021). Systemic effects of missense mutations on SARS-CoV-2 spike glycoprotein stability and receptor-binding affinity. *Briefings in Bioinformatics*, *22*(2), 1239–1253.
16. Mohammadi, E., Shafiee, F., Shahzamani, K., & Mehdi, M. R., Alibakhsi, A., Ahangarzadeh, S., Beikmohammadi, L., Shariati, L., Hooshmandi, S., Ataei, B., Javanmard, S. H.(2020). Novel and emerging mutations of SARS-CoV-2: Biomedical implications. *Biomedicine & Pharmacotherapy* 139, 111599
17. Pachetti, M., Marini, B., Benedetti, F., Giudici, F., Mauro, E., Storici, P., Masciovecchio, C., Angeletti, S., Ciccozzi, M., Gallo, R. C., Zella, D., Ippodrino, R. (2020). Emerging SARS-CoV-2 mutation hot spots include a novel RNA-dependent-RNA polymerase variant. *Journal of Translational Medicine*, *18*(1), 1–9.
18. Chellapandi, P., Saranya, S. (2020). Genomics insights of SARS-CoV-2 (COVID-19) into target-based drug discovery. *Medicinal Chemistry Research*, *29*(10), 1777–1791.
19. Asselah T., Durantel D., Pasmant E., Lau G., Schinazi R.F. (2020). COVID-19: Discovery, diagnostics and drug development. Journal of Hepatology, 74, 168–184.

20. Avaible from: https://www.gisaid.org/ 01/03/2021
21. Dehghanpoor, R., Ricks, E., Hursh, K., Gunderson, S., Farhoodi, R., Haspel, N., Hutchinson, B., Jagodzinski, F. (2018). Predicting the effect of single and multiple mutations on protein structural stability. *Molecules*, *23*(2), 1–18. h
22. Ortega, J. T., Serrano, M. L., Pujol, F. H., Rangel, H. R. (2020). Role of changes in SARS-CoV-2 spike protein in the interaction with the human ACE2 receptor: An in silico analysis. *EXCLI Journal*, *19*, 410–417.
23. Malik, Y. A. (2020). Properties of coronavirus and SARS-CoV-2. *Malaysian Journal of Pathology*, *42*(1), 3–11.
24. Wang, L., Wang, Y., Ye, D., Liu, Q. (2020). Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents*, *55*(6), 105948.
25. https://www.nicd.ac.za/nicd-capacity-for-isolation-and-culturing-of-severe-acute-respiratory-syndrome-coronavirus-2-sars-cov-2-from-clinical-specimens/
26. Uğurel O. M., Ata O., Turgut Balık D. (2020). An updated analysis of variations in SARS-CoV-2 genome. Turkish Journal of Biology, 44, 157-167.
27. Petersen, E., Gökengin, D. (2020). Sars-cov-2 epidemiology and control, different scenarios for Turkey. *Turkish Journal of Medical Sciences*, *50*(SI-1), 509–514.
28. https://covid19.saglik.gov.tr/
29. Singhal T. (2020). A Review of Coronavirus Disease-2019 (COVID-19). The Indian Journal of Pediatrics, 87, 281–286.
30. Prajapa M., Sarma P., Shekkar N., Avti P., Sinha S., Kaur H., Kumar S., Bhattacharyya A., Kumar H., Bansal S., Medhi B. (2020). Drug targets for corona virus: A systematic review. Indian Journal of Pharmacology, 52,56-65.
31. Yüce, M., Filiztekin, E., Özkaya, K. G. (2021).COVID-19 diagnosis —A review of current methods. *Biosensors and Bioelectronics*, 172, 12752.
32. Tarighi, P., Eftekhari, S., Chizari, M., Sabernavaei, M., Jafari, D., Mirzabeigi, P. (2021). A review of potential suggested drugs for coronavirus disease (COVID-19) treatment. *European Journal of Pharmacology*, *895*(December 2020), 173890.
33. Karpiński T. M., Ożarowski M., Seremak-Mrozikiewicz A., Wolski H., Wlodkowic D. (2021). The 2020 race towards SARS-CoV-2 specific vaccines. Theranostics, 11, 1690-1702
34. https://covid19asi.saglik.gov.tr/TR-77709/covid-19-asisi-uretim-teknolojileri.html
35. Adil M. T., Rahman R., Whitelaw D., Jain V., Al-Taan O., Rashid F., Munasinghe A., Jambulingam P. (2021). SARS-CoV-2 and the pandemic of COVID-19. Postgrad Med J, 97, 110–116.
36. Kang, S., Yang, M., Hong, Z., Zhang, L., Huang, Z., Chen, X., He, S., Zhou, Z., Zhou, Z., Chen, Q., Yan, Y., Zhang, C., Shan, H., & Chen, S. (2020). Crystal structure of SARS-CoV-2 nucleocapsid protein RNA binding domain reveals potential unique drug targeting sites. *Acta Pharmaceutica Sinica B*, *10*(7), 1228–1238. https://doi.org/10.1016/j.apsb.2020.04.009
37. Khan M. T., Irfan M., Ahsan M., Ahmed A., Kaushik A. C., Khan A. S., Chinnasamy S., Ali A., Wei D. (2021). Structures of SARS-CoV-2 RNA-Binding Proteins and Therapeutic Targets. Intervirology, 64, 55-68.
38. Peng, Y., Du, N., Lei, Y., Dorje, S., Qi, J., Luo, T., Gao, G. F., Song, H. (2020). Structures of the SARS -CoV-2 nucleocapsid and their perspectives for drug design . *The EMBO Journal*, *39*(20), 1–12.
39. Dinesh, D. C., Chalupska, D., Silhan, J., Koutna, E., Nencka, R., Veverka, V., Boura, E. (2020). Structural basis of RNA recognition by the SARS-CoV-2 nucleocapsid phosphoprotein. *PLoS Pathogens*, *16*(12).
40. Zeng, W., Liu, G., Ma, H., Zhao, D., Yang, Y., Liu, M., Mohammed, A., Zhao, C., Yang, Y., Xie, J., Ding, C., Ma, X., Weng, J., Gao, Y., He, H., Jin, T. (2020). Biochemical characterization of SARS-CoV-2 nucleocapsid protein. *Biochemical and Biophysical Research Communications*, *527*(3), 618–623.
41. Zinzula, L., Basquin, J., Bohn,S., Beck, F., Klumpe, S., Pfeifer, G., Nagy , I., Bracher, A., Hartl, U., Baumeister, W. (2021). High-resolution structure and biophysical characterization of the nucleocapsid phosphoprotein dimerization domain from the Covid-19 severe acute respiratory syndrome coronavirus 2. *Biochemical and Biophysical Research Communications*, 538,54e62.
42. Wu, S., Tian, C., Liu, P., Guo, D., Zheng, W., Huang, X., Zhang, Y., Liu, L. (2021). Effects of SARS-CoV-2 mutations on protein structures and intraviral protein–protein interactions. *Journal of Medical Virology*, *93*(4), 2132–2140.

43. Bendl, J., Stourac, J., Salanda, O., Pavelka, A., Wieben, E. D., Zendulka, J., Brezovsky, J., Damborsky, J. (2014). PredictSNP: Robust and Accurate Consensus Classifier for Prediction of Disease-Related Mutations. *PLoS Computational Biology*, *10*(1).
44. Pandurangan, A. P., & Blundell, T. L. (2020). Prediction of impacts of mutations on protein structure and interactions: SDM, a statistical approach, and mCSM, using machine learning. *Protein Science*, *29*(1), 247–257.
45. Pires, D. E. V., Ascher, D. B., Blundell, T. L. (2014a). DUET: A server for predicting effects of mutations on protein stability using an integrated computational approach. *Nucleic Acids Research*, *42*(W1), 314–319.
46. Rodrigues, C. H. M., Pires, D. E. V.,Ascher, D. B. (2018). DynaMut: Predicting the impact of mutations on protein conformation, flexibility and stability. *Nucleic Acids Research*, *46*(W1), W350–W355.
47. Frappier, V., Chartier, M., Najmanovich, R. J. (2015). ENCoM server: Exploring protein conformational space and the effect of mutations on protein function and stability. *Nucleic Acids Research*, *43*(W1), W395–W400.
48. Pires, D. E. V., Ascher, D. B., Blundell, T. L. (2014b). MCSM: Predicting the effects of mutations in proteins using graph-based signatures. *Bioinformatics*, *30*(3), 335–342.
49. Pandurangan, A. P., Ochoa-Montaño, B., Ascher, D. B., Blundell, T. L. (2017). SDM: A server for predicting effects of mutations on protein stability. *Nucleic Acids Research*, *45*(W1), W229–W235. https://doi.org/10.1093/nar/gkx439
50. Beg, M. A., Athar, F. (2020). Computational method in COVID-19: Revelation of preliminary mutations of RdRp of SARS CoV-2 that build new horizons for therapeutic development. *J Hum Virol Retrovirolog*, *8*(3), 62–72.
51. Vilar, S., Isom, D. G. (2021). One year of SARS-CoV-2: How much has the virus changed? *Biology*, *10*(2), 1–18.
52. Sanavia, T., Birolo, G., Montanucci, L., Turina, P., Capriotti, E., Fariselli, P. (2020). Limitations and challenges in protein stability prediction upon genome variations: towards future applications in precision medicine. *Computational and Structural Biotechnology Journal*, *18*, 1968–1979.
53. Li, G., Panday, S. K., Alexov, E. (2021). Saafec-seq: A sequence-based method for predicting the effect of single point mutations on protein thermodynamic stability. *International Journal of Molecular Sciences*, *22*(2), 1–11.
54. Bromberg, Y., Rost, B. (2007). SNAP: Predict effect of non-synonymous polymorphisms on function. *Nucleic Acids Research*, *35*(11), 3823–3835.
55. Vaser, R., Adusumalli, S., Leng, S. N., Sikic, M., Ng, P. C. (2016). SIFT missense predictions for genomes. *Nature Protocols*, *11*(1), 1–9.
56. Choi, Y., Chan, A. P. (2015). PROVEAN web server: A tool to predict the functional effect of amino acid substitutions and indels. *Bioinformatics*, *31*(16), 2745–2747.
57. Stone, E. A., Sidow, A. (2005). Physicochemical constraint violation by missense substitutions mediates impairment of protein function and disease severity. *Genome Research*, *15*(7), 978–986.
58. Avaible from https://clades.nextstrain.org/ 16/03/2021
59. Avaible from https://zhanglab.ccmb.med.umich.edu/COVID-19/ 01/03/2021
60. Avaible from https://www.uniprot.org/uniprot/P0DTC9 01 16/03/2021
61. Avaible from https://www.gisaid.org/hcov19-variants/ 16/03/2021
62. Avaible from: http://biosig.unimelb.edu.au/dynamut/ 20/03/2021
63. Tseng, Y. T., Wang, S. M., Huang, K. J., Wang, C. T. (2014). SARS-CoV envelope protein palmitoylation or nucleocapid association is not required for promoting virus-like particle production. *Journal of Biomedical Science*, *21*(1), 1–11.
64. He, R., Leeson, A., Ballantine, M., Andonov, A., Baker, L., Dobie, F., Li, Y., Bastien, N., Feldmann, H., Strocher, U., Theriault, S., Cutts, T., Cao, J., Booth, T. F., Plummer, F. A., Tyler, S., & Li, X. (2004). Characterization of protein-protein interactions between the nucleocapsid protein and membrane protein of the SARS coronavirus. *Virus Research*, *105*(2), 121–125.
65. Boscarino, J. A., Logan, H. L., Lacny, J. J., Gallagher, T. M. (2008). Envelope Protein Palmitoylations Are Crucial for Murine Coronavirus Assembly. *Journal of Virology*, *82*(6), 2989–2999.
66. Lopez, S., Yao, J. S., Kuhn, R. J., Strauss, E. G., Strauss, J. H. (1994). Nucleocapsid-glycoprotein interactions required for assembly of alphaviruses. *Journal of Virology*, *68*(3), 1316–1323.
67. Azad, G. K. (2021). Identification and molecular characterization of mutations in nucleocapsid phosphoprotein of SARS-CoV-2. *PeerJ*, *9*.

68. Wootton, S. K., Rowland, R. R. R., Yoo, D. (2002). Phosphorylation of the Porcine Reproductive and Respiratory Syndrome Virus Nucleocapsid Protein. *Journal of Virology*, *76*(20), 10569–10576.
69. McBride, R., van Zyl, M., Fielding, B. C. (2014). The coronavirus nucleocapsid is a multifunctional protein. *Viruses*, *6*(8), 2991–3018.
70. Zeng, W., Liu G., Ma, H., Zhao, D., Yang, Y., Liu, M., Mohammed, A., Zhao, C., Yang, Y., Xie, J., Ding, C., Ma, X., Weng, J., Gao, Y., He, H.,Jin, T.(2020). Biochemical characterization of SARS-CoV-2 nucleocapsid protein. *Biochemical and Biophysical Research Communications*, 527, 618e623
71. Das, J. K., Roy, S. (2021). A study on non-synonymous mutational patterns in structural proteins of SARS-CoV-2. *Genome*, *14*(September 2020), 1–14.
72. Chang, C.-K., Hsu, Y.-L., Chang, Y.-H., Chao, F.-A., Wu, M.-C., Huang, Y.-S., Hu, C.-K., Huang, T.-H. (2009). Multiple Nucleic Acid Binding Sites and Intrinsic Disorder of Severe Acute Respiratory Syndrome Coronavirus Nucleocapsid Protein: Implications for Ribonucleocapsid Protein Packaging. *Journal of Virology*, *83*(5), 2255–2264.
73. Azad, G. K. (2021). The molecular assessment of SARS-CoV-2 Nucleocapsid Phosphoprotein variants among Indian isolates. *Heliyon*, *7*(2), 06167.
74. Zúñiga, S., Cruz, J. L. G., Sola, I., Mateos-Gómez, P. A., Palacio, L., Enjuanes, L. (2010). Coronavirus Nucleocapsid Protein Facilitates Template Switching and Is Required for Efficient Transcription. *Journal of Virology*, *84*(4), 2169–2175.
75. Zúñiga, S., Sola, I., Moreno, J. L., Sabella, P., Plana-Durán, J., Enjuanes, L. (2007). Coronavirus nucleocapsid protein is an RNA chaperone. *Virology*, *357*(2), 215–227.
76. Lo, Y. S., Lin, S. Y., Wang, S. M., Wang, C. T., Chiu, Y. L., Huang, T. H., Hou, M. H. (2013). Oligomerization of the carboxyl terminal domain of the human coronavirus 229E nucleocapsid protein. *FEBS Letters*, *587*(2), 120–127.
77. Kadioglu, O., Saeed, M., Greten, H. J., Efferth, T. (2021). Identification of novel compounds against three targets of SARS CoV-2 coronavirus by combined virtual screening and supervised machine learning. *Computers in Biology and Medicine*, *133*(November 2020), 104359.
78. Hu, X., Zhou, Z., Li, F., Xiao, Y., Wang, Z., Xu, J., Dong, F., Zheng, H., Yu, R. (2021). The study of antiviral drugs targeting SARS-CoV-2 nucleocapsid and spike proteins through large-scale compound repurposing. *Heliyon*, *7*(3), e06387.
79. Ahamad, S., Gupta, D., & Kumar, V. (2020). Targeting SARS-CoV-2 nucleocapsid oligomerization: Insights from molecular docking and molecular dynamics simulations. *Journal of Biomolecular Structure and Dynamics*, *0*(0), 1–14.
80. Patro, P. P., Sathyaseelan, C., Uttamrao, P. P., Rathinavelan, T. (2020). Global variation in the SARS-CoV-2 proteome reveals the mutational hotspots in the drug and vaccine candidates. https://doi.org/10.1101/2020.07.31.230987
81. Eskier, D., Akalp, E., Dalan, Ö., Karakülah, G., Oktay, Y. (2021). Current mutatome of sars-cov-2 in turkey reveals mutations of interest. *Turkish Journal of Biology*, *45*(1), 104–113.
82. Rahman, M. S., Islam, M. R., Alam, A. S. M. R. U., Islam, I., Hoque, M. N., Akter, S., Rahaman, M. M., Sultana, M., Hossain, M. A. (2021). Evolutionary dynamics of SARS-CoV-2 nucleocapsid protein and its consequences. *Journal of Medical Virology*, *93*(4), 2177–2195.
83. M Rubayet Ul Alam, A. S., Kibria Islam, O., Shazid Hasan, M., Raihanul Islam, M., Mahmud, S., AlEmran, H. M., Kabir Jahid, I., Crandall, K. A., Anwar, M., & Anwar Hossain, M. (2021). Evolving Infection Paradox of SARS-CoV-2: Fitness Costs Virulence?. *medRxiv*, 2021.02.21.21252137.
84. Singh, J. (2021). Structure-function investigation of a new VUI-202012 / 01 SARS-CoV-2 variant *Keywords*. 1–8.