



T.C.  
MERSİN ÜNİVERSİTESİ  
Sağlık Bilimleri Enstitüsü

**İLAC ARAŞTIRMA VE GELİŞTİRME ÇALIŞMALARINDA  
MAKİNE ÖĞRENİMİ**

**YÜKSEK LİSANS TEZİ**

**MEHMET ALİ YÜCEL**

**FARMASÖTİK KİMYA ANABİLİM DALI**

**MERSİN**  
**ARALIK-2022**

T.C.  
MERSİN ÜNİVERSİTESİ  
Sağlık Bilimleri Enstitüsü

# İLAÇ ARAŞTIRMA VE GELİŞTİRME ÇALIŞMALARINDA MAKİNE ÖĞRENİMİ

YÜKSEK LİSANS TEZİ

MEHMET ALİ YÜCEL

ORCID ID: 0000-0003-2880-7992

DANIŞMAN  
Prof. Dr. ÖZTEKİN ALGÜL

FARMASÖTİK KİMYA ANABİLİM DALI

MERSİN  
ARALIK- 2022

## ETİK BEYAN

Mersin Üniversitesi Lisansüstü Eğitim-Öğretim Yönetmeliğinde belirtilen kurallara uygun olarak hazırladığım bu tez çalışmada,

- Tez içindeki bütün bilgi ve belgeleri akademik kurallar çerçevesinde elde ettiğimi,
- Görsel, işitsel ve yazılı tüm bilgi ve sonuçları bilimsel ahlâk kurallarına uygun olarak sunduğumu,
- Başkalarının eserlerinden yararlanılması durumunda ilgili eserlere bilimsel normlara uygun olarak atıfta bulunduğumu,
- Atıfta bulunduğum eserlerin tümünü kaynak olarak kullandığımı,
- Kullanılan verilerde herhangi bir tahrifat yapmadığımı,
- Bu tezin herhangi bir bölümünü Mersin Üniversitesi veya başka bir üniversitede başka bir tez çalışması olarak sunmadığımı,
- Tezin tüm telif haklarını Mersin Üniversitesi'ne devrettiğimi

beyan ederim.

## ETHICAL DECLARATION

This thesis is prepared in accordance with the rules specified in Mersin University Graduate Education Regulation and I declare to comply with the following conditions:

- I have obtained all the information and the documents of the thesis in accordance with the academic rules.
- I presented all the visual, auditory and written informations and results in accordance with scientific ethics.
- I refer in accordance with the norms of scientific Works about the case of exploitation of others' works.
- I used all of the referred works as the references.
- I did not do any tampering in the used data.
- I did not present any part of this thesis as an another thesis at Mersin University or another university.
- I transfer all copyrights of this thesis to the Mersin University.

22 /12/ 2022

İmza / Signature

Mehmet Ali YÜCEL

## ÖZET

### İLAÇ ARAŞTIRMA VE GELİŞTİRME ÇALIŞMALARINDA MAKİNE ÖĞRENİMİ

İlaç geliştirme çalışmaları insan yaşamının kalitesinin arttırması ve yaşam süresinin uzatılması açısından oldukça büyük bir öneme sahiptir. Yeni hastalıkların ortaya çıkması, ilaçlara direnç gelişmesi gibi nedenlerden ötürü yeni ilaç moleküllerin geliştirilmesinin sürekliliğine ihtiyaç vardır. Ancak bu durum uzun süreli ve yüksek maliyetli bir süreçtir. Hastalıkların temelinde yatan mekanizmaların keşfedilmesi ve bu mekanizmalara yönelik moleküllerin geliştirilmesinin hızlandırılması ve maliyeti azaltmak için yeni yöntemlerin geliştirilmektedir. Bu yöntemlerden olan yapay zeka ve alt disiplinleri olan makine öğrenmesi ve derin öğrenme son yıllarda ilaç keşfinde önemli başarılar elde etmişlerdir. Bu çalışmada ChEMBL veri tabanından *S. aureus* ve *C. albicans* için daha önce sentezlenen ve minimum inhibitör konsantrasyon değerine sahip olan moleküllerden veri seti oluşturulmuştur. Veri seti düzenlenerek moleküller aktif veya inaktif olarak etiketlenmiştir. Ardından veri setlerindeki moleküllerin moleküller tanımlayıcıları "ECFP" yöntemi ile elde edilmiştir. Veri setinin hazırlanmasından sonra, Keras temelli derin sinir ağıları mimarisi kullanılarak makine öğrenmesi modelleri geliştirilmiştir. Modellerin eğitilmesinin ardından, çeşitli metrikler ile değerlendirilmişlerdir. *C. albicans* için; doğruluk: 0.84, F1 skoru: 0.81, MKK: 0.68, ve AUC: 0.83 olarak bulunurken, *S. aureus* için; doğruluk: 0.84, F1 skoru: 0.87, MKK: 0.67, AUC: 0.83 olarak bulunmuştur. Ardından 19 adet 1,2,3,4-tetrasüstitüepirol-kalkon türevi bileşikler tasarlanmıştır. Bu moleküllerin oluşturulan modeller ile antimikrobiyal aktiviteleri tahmin edilmiştir. Modellerin çıktılarına göre moleküllerin antifungal aktivitelerinin olabileceği ancak antibiyotik etkilerinin olmayacağı tespit edildikten sonra moleküller sentezlenmiştir. Öncelikle 1,2,3,4-tetrasüstitü pirol türevleri sentezlenmiştir, sonrasında ticari olarak satın alınan benzaldehit, 4-florobenzaldehit ve 4-metoksibenzaldehit türevleri ile kalkon türevleri elde edilmiştir. Sentezlenen moleküllerin yapı tayinleri <sup>1</sup>H NMR, <sup>13</sup>C NMR ve IR ile gerçekleştirilmiştir. Sentezlenen moleküllerin, *S. aureus* ve *C. albicans* dahil antimikrobiyal etkileri *in vitro* olarak incelenmiştir. Antimikrobiyal aktivite sonuçları ile makine öğrenmesi sonuçları birbirini destekler nitelikte bulunmuştur. Biyolojik sonuçlara göre, **MAY-6** molekülünün *C. albicans* dahil çeşitli fungus türleri üzerinde en etkili molekül olduğu belirlenmiştir.

**Anahtar Kelimeler:** Makine Öğrenimi; İlaç Geliştirme, Antimikrobiyal; Yapay Zeka; Derin Öğrenme

## ABSTRACT

### MACHINE LEARNING IN DRUG RESEARCH AND DEVELOPMENT

Drug development studies are of great importance in terms of increasing the quality of human life and prolonging life expectancy. Due to the emergence of new diseases and the development of resistance to drugs, there is a need for the continuity of the development of new drug molecules. However, this is a long-term and costly process. New methods are being developed to discover the underlying mechanisms of diseases and accelerate the development of molecules for these mechanisms and to reduce costs. Artificial intelligence, which is one of these methods, and machine learning and deep learning, which are its sub-disciplines, have achieved significant success in drug discovery in recent years. In this study, a dataset was created from the ChEMBL database for *S. aureus* and *C. albicans* from previously synthesized molecules with a minimum inhibitory concentration value. The data set was edited, and the molecules were labeled as active or inactive. Then, the molecular identifiers of the molecules in the data sets were obtained with "ECFP". After the preparation of the data set, machine learning models were developed using the Keras-based deep neural networks architecture. After training the models, they were evaluated with various metrics. For *C. albicans*, the accuracy was found as 0.84, F1 score: 0.81, the MCC:0.68, and AUC:0.83, while for *S. aureus*; the accuracy was found as 0:84, F1 score: 0.87, MCC: 0.67, AUC: 0.83. Then 19 1,2,3,4-tetrasubstituteepirol-cachalcone derivative compounds were designed. The antimicrobial activities of these molecules have been predicted by the models created. According to the outputs of the models, the molecules were synthesized after it was determined that the molecules may have antimicrobial activities. First, 1,2,3,4-tetrtasubstituted pyrrole derivatives were synthesized, then commercially wound benzaldehyde, 4-fluorobenzaldehyde and 4-methoxybenzaldehyde derivatives and chalcone derivatives were obtained. The structure determination of the synthesized molecules was carried out with <sup>1</sup>H NMR, <sup>13</sup>C NMR, and IR. The antimicrobial effects of the synthesized molecules, including *S. aureus* and *C. albicans*, were examined in vitro. Antimicrobial activity results and machine learning results were found to support each other. According to the biological results, **MAY-6** compounds were the most effective molecules on fungus strains including *C.albicans*.

**Keywords:** Machine Learning; Drug Development, Antimicrobial; Artificial intelligence; Deep Learning

## TEŞEKKÜR

Lisans eğitimim ve yüksek lisans sürecimde çalışkanlığı ve kişiliğiyle her zaman örnek aldığım, çalışmamın her aşamasında bilgi ve deneyimlerini benden esirgemeyen, eşsiz sabrı ve özverisiyle büyük emeği geçen, geniş vizyonu ile her zaman yol gösterici olan, iyi bir çiftçi gibi yeşermemiz için önden giderek saban ile toprağı yumuşatıp tohumlarımızın yeşermesi için uygun şartları oluşturan, hayatımdaki varlığının önemini tarif edemediğim, kendisine çok şey borçlu olduğum değerli danışman hocam Sayın Prof. Dr. Öztekin ALGÜL'e,

Yaklaşık 7 yıl önce sıra arkadaşlığı ile başlayan zamanla en iyi anlaştığım insana dönüşen, her zaman yanımda olan, birlikte yaptığımız tartışmalar ile hayal gücümü her zaman diri tutan, kendisi ile o kadar uzun süredir aynı yolda yürüyoruz ki kendimden farkını göremediğim, iyi niyetinden hiç şüphe etmediğim değerli yol arkadaşım Ercan ADAL'a

Hayatımın her anında sevgi ve desteklerini benden hiç esirgemeyen, bana her zaman güvenen, bugünlere gelebilmemdeki en büyük pay sahipleri, varlıklarıyla hayatıma anlam katan sevgili aileme,

Teşekkürlerimi sunarım.

## İÇİNDEKİLER

	Sayfa
İÇ KAPAK	i
ONAY	ii
ETİK BEYAN	iii
ÖZET	iv
ABSTRACT	v
TEŞEKKÜR	vi
İÇİNDEKİLER	vii
TABLolar DİZİNİ	viii
ŞEKİLLER DİZİNİ	ix
KISALTMALAR ve SİMGELER	x
<b>1. GİRİŞ</b>	<b>1</b>
<b>2. KAYNAK ARAŞTIRMALARI</b>	<b>3</b>
2.1. Yapay Zeka ve Makine Öğrenimi	4
2.2. Kimyada Hesaplamalı Yöntemlerin Kısa Tarihi	7
2.3. Yapay Zeka ve İlaç Keşfi	8
2.3.1. Makine Öğrenimi ve İlaç Araştırma-Geliştirme Çalışmaları	8
2.4. Makine Öğrenimi Algoritmaları	11
2.4.1. Ekstrem Gradyan Arttırma (eXtreme Gradient Boost ,XGBoost)	11
2.4.2. Destekli Vektör Makinesi (Support Vector Machines, SVM)	11
2.4.3. Derin Sinir Ağları (Deep Neural Network, DNN)	12
2.4.4. Rastgele Orman (Random Forest, RF)	14
2.4.5. K-En Yakın Komşu (K-Nearest Neighbour, KNN)	14
2.5. Moleküler Tanımlayıcılar	15
2.5.1. 2.4.1. Genişletilmiş Bağlantılı Parmak İzi (Extended-Connectivity Fingerprint, ECFP)	15
2.5.2. Moleküler Erişim Sistemi (The Molecular Access System, MACCS)	17
2.5.3. Moleküler Grafikler	17
2.5. Makine Öğrenmesinin İlaç Araştırma ve Geliştirmede Uygulamaları	18
2.5.1. Makine Öğrenmesinin Yapı Temelli Sanal Taramada Kullanımları	18
2.5.2. Makine Öğrenmesinin Ligand Temelli Sanal Taramada (LTST) Kullanımları	21
2.5.3. Fizikokimyasal ve ADMET Özelliklerin Tahmininde Makine Öğrenimi Uygulamaları	22
2.5.3.1. Fizikokimyasal Özelliklerin Tahmini	23
2.5.3.2. ADMET Tahminleri	26

	<b>Sayfa</b>
2.5.4. İlaç Yeniden Konumlandırma Çalışmalarında Makine Öğrenimi Uygulamaları	31
2.5.5. Organik Sentez Planlamada Makine Öğrenimi Uygulamaları	32
2.6. Antimikrobiyal İlaç Geliştirmede Makine Öğrenimi	32
<b>3. MATERYAL ve YÖNTEM</b>	<b>33</b>
3.1. Makine Öğrenmesi Modelinin Oluşturulması	33
3.1.1. Veri Setinin Hazırlanması	33
3.1.2. Makine Öğrenmesi Modelinin Eğitilmesi	33
3.2. Sentez Çalışmaları	34
3.2.1. Materyal	34
3.2.2. Yöntem	34
3.2.2.1. Genel Sentez Yöntemi-I	34
3.2.2.2. Genel Sentez Yöntemi-II	34
3.3. Analitik Çalışmalar	35
3.3.1. Erime Noktası Tayini	35
3.3.2. İnce Tabaka Kromatografisi ile Yapılan Kontroller	35
3.3.3. Spektroskopik Çalışmalar	35
3.3.3.1. FT-IR Spektrumları	35
3.3.3.2. <sup>1</sup> H-NMR Spektrumları	36
3.3.3.3. <sup>13</sup> C-NMR Spektrumları	36
3.4. Antimikrobiyal Aktivite Çalışmaları	36
3.4.1. Materyal	36
3.4.2. Kullanılan Araç ve Gereçler	36
3.4.3. Bakteriyolojik Tanı İçin Kullanılan Besiyerleri	37
3.4.3.1. Mueller Hinton Broth (Fluka 70192)	37
3.4.3.2. Tryptic Soy Broth (TSB) (Merck 1.05459)	38
3.4.4. Antibakteriyel ve Antifungal Aktivite Çalışmaları	38
<b>4. BULGULAR</b>	<b>40</b>
4.1. Makine Öğrenmesi Modelinin Değerlendirilmesi	40
4.2. Kimyasal Bulgular	41
4.2.1. 1-(2-Metil-1,4-difenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-1)	41
4.2.2. 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-2)	42



	<b>Sayfa</b>
4.2.3. 1-(1-(4-Klorofenil)-4-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-3)	43
4.2.4. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-4)	44
4.2.5. 1-(4-(4-Klorofenil)-1-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-5)	45
4.2.6. 1-(1,4-bis(4-florofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-6)	46
4.2.7. 1-(4-(4-Florofenil)-2-metil-1-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-8)	47
4.2.8. 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-9)	48
4.2.9. 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-10)	49
4.2.10. 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-11)	50
4.2.11. 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-13)	51
4.2.12. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-15)	52
4.2.13. 3-(4-metoksifenil)-1-(2-metil-1,4-difenil-1H-pirol-3-il)prop-2-en-1-on (MAY-16)	53
4.2.14. 1-(1,4-bis(4-Klorofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on (MAY-17)	54
4.2.15. 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on (MAY-18)	55
4.2.16. 3-(4-Florofenil)-1-(4-(4-florofenil)-2-metil-1-fenil-1H-pirol-3-il)prop-2-en-1-on (MAY-19)	56
4.2.17. 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on (MAY-21)	57
4.2.18. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on (MAY-23)	58
4.2.19. 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (MAY-24)	59
4.3. Antimikrobiyal Bulgular	60
<b>5. TARTIŞMA ve SONUÇ</b>	<b>64</b>
KAYNAKLAR	68
EKLER (Var ise)	76
ÖZGEÇMİŞ	95

## TABLolar DİZİNİ

	<b>Sayfa</b>
Tablo 1.1. Sentezlenen 1,2,3,4-Tetrasübstitüe pirol-kalkon türevi bileşikler	3
Tablo 4.1. Makine öğrenmesi modellerinin değerlendirilme sonuçları	40
Tablo 4.2. Pirol-kalkon türevi bileşiklerin antimikrobiyal aktivite sonuçları (µg/mL)	63
Tablo 5.1. Sentezlenen 1,2,3,4-Tetrasübstitüe pirol-kalkon türevi bileşikler	66



## ŞEKİLLER DİZİNİ

	<b>Sayfa</b>
Şekil 2.1. Yapay zeka, makine öğrenmesi ve derin öğrenme	4
Şekil 2.2. Makine öğrenmesi ve klasik programlama arasındaki fark	5
Şekil 2.3. Makine öğrenmesi türleri ve örnekleri	6
Şekil 2.4. Erken aşama ilaç keşfi çalışmalarında yapay zeka ve makine öğrenimi uygulamaları. YTST: Yapı temelli sanal tarama. LTST: Ligand temelli sanal tarama	9
Şekil 2.5. Tahmin modellerinde kullanılan makine öğrenmesi metodlarının değişimi ve veri gereklilikleri	10
Şekil 2.6. İlaç geliştirmede makine öğrenmesinin kullanıldığı çeşitli alanlar	10
Şekil 2.7. Bu görselde SVM algoritmasının 2 boyutlu bir veri setinde linear ve non-linear şematize edilmiştir.	12
Şekil 2.8. Yapay sinir ağlarının mimarisi	13
Şekil 2.9. Rastgele orman algoritması	14
Şekil 2.10. Benzamid molekülü ve atomların numaralandırılması	15
Şekil 2.11. Parmak izi oluşturmanın ilk aşaması	16
Şekil 2.12. Bir yinelemeden sonra, tanımlayıcı artık 6 numaralı atomun en yakın komşuları hakkında bilgi içerir	16
Şekil 2.13. İki yinelemeden sonra, tanımlayıcı artık 6 numaralı atomun komşularının komşuları (2 bağ uzunluğundaki atomlar) hakkında da bilgi içerir	16
Şekil 2.14. Hipotetik olarak 10-bitlik bir temel yapı parmak izinin temsili, 3 tane bit 1 olarak işaretlenmiştir çünkü bunlar molekülde olan yapıları göstermektedir	17
Şekil 2.15. Moleküler grafiklerin gösterimi	18
Şekil 2.16. Farmokokinetik süreç ve önemli olan moleküler özellikler	23
Şekil 2.17. DeepTox ağının farklı katmanları tarafından öğrenilen toksikofor yapılar. İlk iki satırda belirtilen yapılar ağın ilk gizli katmanında öğrenilmiştir, 3. Satırda öğrenilen yapılar ise ağın daha derin bir katmanından öğrenilmiştir.	31
Şekil 3.1. 1,2,3,4-Tetrasübsitüe pirol bileşiklerin sentez şeması	34
Şekil 3.2. 1,2,3,4-Tetrasübsitüe Pirol-Kalkon Hibrit Bileşiklerinin Sentez Şeması	35
Şekil 4.1. Modellerin eğitim sürecinde kayıp fonksiyonun ve doğruluğunun değişimi	40
Şekil 4.2. 1-(2-Metil-1,4-difenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-1)	41
Şekil 4.3. 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-2)	42
Şekil 4.4. 1-(1-(4-Klorofenil)-4-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-3)	43
Şekil 4.5. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-4)	44

	<b>Sayfa</b>
Şekil 4.6. 1-(4-(4-Klorofenil)-1-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-5)	45
Şekil 4.7. 1-(1,4-bis(4-florofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-6)	46
Şekil 4.8. 1-(4-(4-Florofenil)-2-metil-1-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı (MAY-8)	47
Şekil 4.9. 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı (MAY-9)	48
Şekil 4.10. 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı (MAY-10)	49
Şekil 4.11. 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı (MAY-11)	50
Şekil 4.13. 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı (MAY-13)	51
Şekil 4.14. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı (MAY-15)	52
Şekil 4.15. 3-(4-metoksifenil)-1-(2-metil-1,4-difenil-1H-pirol-3-il)prop-2-en-1-on bileşiğinin yapısı (MAY-16)	53
Şekil 4.16. 1-(1,4-bis(4-Klorofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı (MAY-17)	54
Şekil 4.17. 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı (MAY-18)	55
Şekil 4.18. 3-(4-Florofenil)-1-(4-(4-florofenil)-2-metil-1-fenil-1H-pirol-3-il)prop-2-en-1-on bileşiğinin yapısı (MAY-19)	56
Şekil 4.19. 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı (MAY-21)	57
Şekil 4.21. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı (MAY-23)	58
Şekil 4.22. 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil) prop-2-en-1-on bileşiğinin yapısı (MAY-24)	59
Şekil 5.1. Genel sentez yöntemi I'de verilen reaksiyon ait önerilen mekanizma	64
Şekil 5.2. Genel sentez yöntemi II'de verilen reaksiyon ait önerilen mekanizma	64

## KISALTMALAR ve SİMGELER

---

Kısaltma/Simgesi	Tanım
<i>B. subtilis</i>	<i>Bacillus subtilis</i>
CADD	Computer-aided drug desing
<i>C. albicans</i>	<i>Candida albicans</i>
<i>C. parapsilosis</i>	<i>Candida parapsilosis</i>
<sup>13</sup> C-NMR	Karbon 13 Nükleer Manyetik Rezonans
CNN	Convolutional Neural Networks
DNN	Deep Neural Nwtwork
XGBoost	eXtreme Gradient Boost
<i>E. faecalis</i>	<i>Enterococcus faecalis</i>
<i>E. Coli</i>	<i>Escherichia Coli</i>
ECFP	Extended-Connectivity Fingerprint
FT-IR	Fourier Transformlu Kızılötesi (Fourier Transform Infrared)
<sup>1</sup> H-NMR	Proton Nükleer Manyetik Rezonans
İTK	İnce Tabaka Kromatografisi
KNN	K-Nearest Neighbour
LSTMs	Long Short Term Memory Networks
LTST	Ligand Temelli Sanal Tarama
ML	Machine Learning
MKK	Matthew korelasyon katsayısı
MACCS	Molecular Access System
PPB	Plazma Protein Bağlanması
<i>P. aeuroginosa</i>	<i>Pseudomonas aeuroginosa</i>
RF	Random Forest
RNNs	Recurrent Neural Networks
<i>S. aureus</i>	<i>Staphylococcus aureus</i>
<i>S. pneumoniae</i>	<i>Streptococcus pneumoniae</i>
SVM	Support Vector Machine
YTST	Yapı Temelli Sanal Tarama
YZ	Yapay Zeka
QSAR	Kantitatif Yapı Aktivite İlişkisi

---

## 1. GİRİŞ

İlaç moleküllerinin keşfi ve geliştirilmesi süreci; yüksek maliyet klinik denemelerde sınırlı başarı ve uzun süreler nedeniyle zorluklar barındırmaktadır. Ayrıca yeni moleküllerin, mevcut ilaçlardan daha etkili, daha az yan etkili ve ucuz olması gerekmesi de ilaç endüstrinin üretkenliğini azaltan bir husustur. Tüm bu zorluklara rağmen ilaç araştırma ve geliştirme çalışmalarının devamı bireyin hayat kalitesi, toplum sağlığı ve devletler için gerekli ve oldukça önemlidir.

İlaç geliştirme zorluklarının azaltılması için çeşitli teknikler ve yaklaşımlar geliştirilmiştir ve geliştirilmeye devam etmektedir. Bunların birçoğu bilgisayar temelli tekniklerdir. Özellikle moleküler docking ve kantitatif yapı-etki ilişkileri çalışmaları, bu alanda önemli hale gelen tekniklerdir. Bu tekniklerin ilaç geliştirme sürecine olumlu katkılarına rağmen hatasız veya kusursuz değildir. Örneğin; moleküler docking'de skor fonksiyonlarının göreceli olması ve kantitatif yapı-etki çalışmalarında büyük molekül setleri üstünde çalışılmaması, bu tekniklerin eksikliklerine örnek olarak verilebilir.

Günümüz dünyası, bir enformasyon çağı (veya büyük veri çağı) olarak ifade edilebilir. Bunun temel sebebi, insanlık tarihi boyunca üretilmiş verinin günümüzde çok kısa sürede üretilebilmesidir. Eski dönemlerde veri veya bilginin elde edilebilmesi oldukça uzun zaman ve emek gerektirmekteyken, günümüzde veri oluşturulması oldukça kolaydır. Bunun sonucunda birçok farklı alanda veriler birikmektedir. Ne var ki, bu oldukça büyük boyutlarda olan verilerin, işlenmesi ve bilgiye dönüştürülmesi gerekmektedir. Büyük veri-bilgi dönüşümünü sağlamak yüksek işlemci gücü gerektirdiğinden, bunun insan zihni tarafından efektif bir şekilde yapılabilmesi mümkün görünmemektedir. Bunun üstesinden gelebilmek için, her alandaki verileri hızlı bir şekilde bilgiye dönüştürebilecek bir bilgisayar programının fikri oluşmuştur ve bu fikir makine zekası veya yapay zeka (YZ) olarak isimlendirilmiştir.

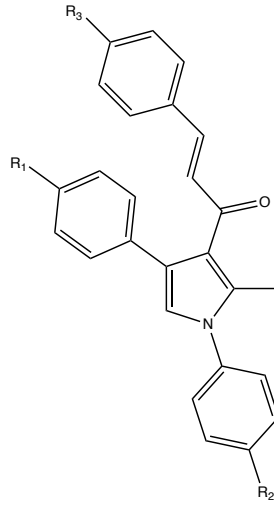
Yapay zeka, birçok farklı alandan veriyi işleyebilen ve genelleştirebilen bir ideal programı ifade etmektedir ancak günümüzde henüz bu ideal yapısına kavuşamamıştır. Bu nedenle zaman içerisinde yapay zekanın, göreve özgü geliştirilen ve başarılı sonuçlar veren makine öğrenmesi (machine learning, ML) ve derin öğrenme (deep learning, DL) alt alanları oluşmuştur.

İlaç araştırma ve geliştirme, birçok süreci kapsamaktadır ve bir molekülün bu süreçlerin her birinde başarılı olması gerekmektedir. Bu nedenle, moleküllerin bu aşamalarındaki başarılarını tahmin edebilmek oldukça önemlidir. Bu tahminlerin yüksek doğrulukla ve hızlı olabilmesi için YZ, ML ve DL teknikleri son yıllarda kullanılmaya başlanmıştır.

Bu tez çalışmasında, ilaç geliştirme çalışmalarında makine öğreniminin önemi ortaya konulmaya çalışmıştır. Bu kapsamda makine öğreniminin, ilaç geliştirmenin hangi aşamalarında

nasıl kullanıldığı, literatürde nasıl çalışmaların yapıldığı belirtilerek ve makine öğrenmesi teknikleri anlatılmıştır. Bu tez çalışmasında makine öğrenme teknikleri ile antimikrobiyal etki gösterebilecek moleküllerin tahminini yapabilen bir makine öğrenmesi modeli geliştirilmiş, (E)-1-(2-metil-1,4-difenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on genel yapısına sahip moleküllerin (Şekil 1.1) sentezi, yapı tayinlerinin ardından aktiviteleri belirlenmiştir. Son aşamada ise makine öğrenmesi modelinin aktif olarak tahmin ettiği moleküllerin *in vitro* aktivite sonuçları ile karşılaştırılarak modelin değerlendirilmesi gerçekleştirilmiştir.





**Tablo 1.1.** Sentezlenen 1,2,3,4-Tetrasüstitüe pirol-kalkon türevi bileşikler

R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Molekül
-H	-H	-H	MAY-1
-Cl	-Cl	-H	MAY-2
-Cl	-F	-H	MAY-3
-Cl	-H	-H	MAY-4
-F	-Cl	-H	MAY-5
-F	-F	-H	MAY-6
-H	-F	-H	MAY-8
-Cl	-Cl	-OCH <sub>3</sub>	MAY-9
-Cl	-F	-OCH <sub>3</sub>	MAY-10
-F	-Cl	-OCH <sub>3</sub>	MAY-11
-F	-F	- OCH <sub>3</sub>	MAY-13
-Cl	-H	- OCH <sub>3</sub>	MAY-15
-H	-H	- OCH <sub>3</sub>	MAY-16
-Cl	-Cl	- F	MAY-17
-F	-F	-F	MAY-18
-H	-F	-F	MAY-19
-Cl	-F	-F	MAY-21
-Cl	-H	-F	MAY-23
-F	-Cl	-F	MAY-24

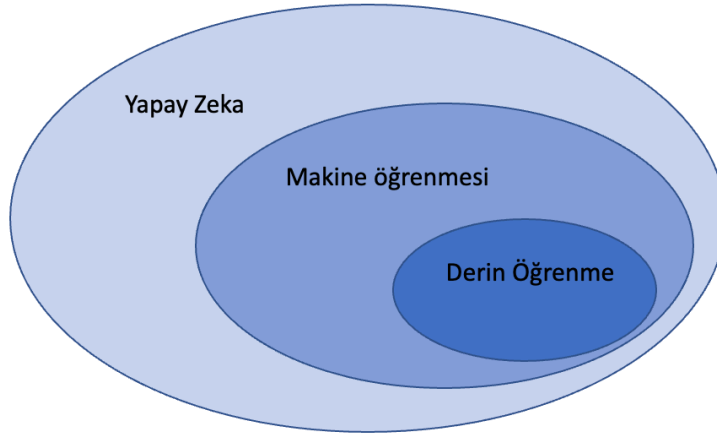


## 2. KAYNAK ARAŞTIRMALARI

### 2.1. Yapay Zeka ve Makine Öğrenimi

Yapay zeka, 1950'lerde bilgisayar biliminin öncüleri tarafından bilgisayarların düşünüp düşünemeyeceğinin sorgulanmasıyla başlamıştır. Yapay zeka, insanlar tarafından gerçekleştirilen entelektüel eylemleri otomatize etme çabası şeklinde tanımlanabilir. Ancak bu tanımlama yeterli değildir. Çünkü yapay zeka, makine öğrenimi ve derin öğrenmeyi kapsayan ancak aynı zaman herhangi bir öğrenmeyi içermeyen genel bir alandır. [1](Şekil 2.1.)

Örneğin satranç için geliştirilen ilk oyun taşlarının hareket kurallarının net olması ve bu kuralların programlarında, satranç kolaylıkla kodlanabilmesi nedeniyle bu programlama sadece bir otomasyon ve hızlandırma ile sonuçlanan bir durum oluşturmuştur. Bu durum klasik bir yapay zeka çalışmasıdır, ancak bir makine öğrenimi olarak nitelendirilemez. Çünkü klasik yapay zeka kuralları belli olan ve herhangi bir bilgisayar öğreniminin gerçekleşmediği bir bilgisayar programlamasıdır. Dolayısıyla her yapay zeka uygulaması makine öğrenimi olarak değerlendirilemez.

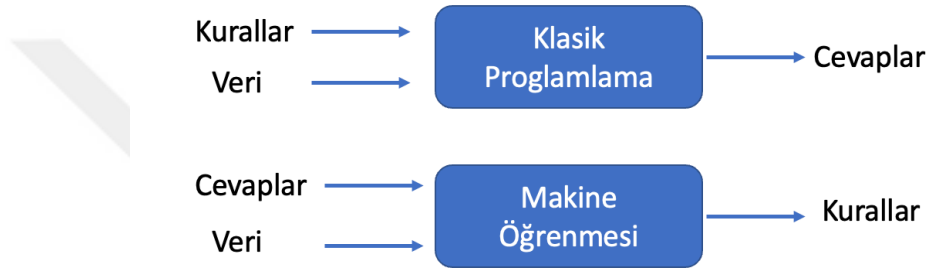


**Şekil 2.1.** Yapay zeka, makine öğrenmesi ve derin öğrenme

Bilgisayar programcıları satranç oyununda olduğu gibi kuralların tamamının bilinmesine bağlı olarak yapay zeka kodlamalarını kolaylıkla yazabileceklerini düşünüyorlardı. Bu durum oldukça uzun bir süre, 1950'lerden 1980'lerin sonlarına kadar, klasik yapay zekanın kabul gören görüş olarak sürmesine neden olmuştur. Ancak her alanda satranç da olduğu gibi kuralların net ve sonuçların hesaplanabilir olması mümkün değildir. İlave olarak herhangi bir konuda kurallar belli dahi olsa elde edilecek sonuçların hesaplanabilmesi mümkün olamayabilir. Klasik yapay zeka, satranç gibi iyi tanımlanmış mantıksal problemleri çözmek için uygun olsa da, görüntü sınıflandırması, konuşma tanıma, dil çevirisi ve ilaç keşfi gibi daha karmaşık alanlarda problemleri çözmek için kuralların net belirlenememesinden dolayı yüksek başarı elde

edilememiştir. Klasik yapay zekanın çözemediği problemler için yeni bir yaklaşım olarak makine öğrenimi ortaya çıkmıştır. [1]

Makine öğrenimi, bir bilgisayarın belirli bir görevi nasıl gerçekleştireceğini kendisinin öğrenmesini hedefler. Bir diğer ifadeyle kuralların programcı tarafından belirlendiği değil, verilere bakarak bilgisayarın otomatik olarak öğrenmesidir. Klasik programlamada, klasik yapay zeka paradigması, girdi kuralları ve bu kurallara göre işlenecek veriler sunulur bunun sonucunda yanıtlar bilgisayar tarafından ortaya çıkarılır. Makine öğreniminde ise sadece veriler değil verilerin yanıtları da verilir ve kurallar bilgisayar tarafından öğrenilmesi hedeflenir. Bu öğrenim becerisi daha sonra başka verilere de uygulanabilir.[1] (Şekil 2.2.).



Şekil 2.2. Makine öğrenmesi ve klasik programlama arasındaki fark

Makine öğrenimi sisteminde temel amaç, bilgisayarı programlamak yerine bilgisayara öğrenim becerisi kazandırmaktır. Bu becerinin oluşması için çok sayıda veri sunulur ve bu veriler algoritmalar tarafından değerlendirilir ve hedeflenen becerinin edinimini sağlayan kurallar matematiksel olarak öğrenilir. Örneğin, bir kişiye yaz ve kış fotoğrafları gösterildiğinde fotoğraflardaki belli özelliklere bağlı olarak fotoğrafları yaz ve kış mevsimi olarak sınıflandırabilir. Bu beceri makineye öğretilme yöntemi ise şu şekilde olur: Çok sayıda kış ve yaz fotoğraflarını içeren veri seti makine öğrenme sistemine tanımlanabilmesi için uygun matematik ifadeye dönüştürülür. Daha sonra sistemin yaz ve kış ile matematiksel ifadeler arasında var olan ilişkiyi algılaması öğrenme olarak tanımlanır. Ardından herhangi farklı bir fotoğraf yani veri setinde olmayan fotoğraf makine öğrenmesi sistemine tanımlandığı da bu fotoğrafın yaz veya kış fotoğrafı olduğu sistem tarafından tahmin edilir.

Makine öğrenimi, zaman içerisinde yapay zekanın en popüler ve en başarılı alt alanı haline gelmiştir. Makine öğrenimi, matematik ve istatistik ile yakından ilişkilidir. İstatistikten farklı olarak makine öğrenimi klasik istatistik yöntemlerin pratik olamayacağı karmaşık, büyük veri kümelerinde de anlamlı sonuçlar ortaya çıkarabilmektedir.

Günümüzde öğrenme alışkanlıklarına göre farklılaşan makine öğrenim yöntemleri bulunmaktadır (Şekil 2.3.). En başarılı makine öğrenimi algoritmaları, bilinen örneklerden genelleme yapılarak karar verme süreçlerini otomatikleştirenlerdir. Bu anlamda en sık kullanılan öğrenme tipi denetimli (gözetimli) öğrenme olarak bilinir. Kullanıcı algoritmaya girdileri ve istenen çıktıları verir ve algoritma bu girdilerden istenen çıktıyı üretmenin bir yolunu bulur. Böylece algoritma daha önce veri setinde olmayan bir girdi verildiğinde kolaylıkla bir çıktı oluşturabilir hale gelir. [2,3]

El yazısından zarf üstündeki posta kodunu tanımlama, medikal bir görüntüden tümörün malign veya benign olduğunu tanımlama veya bileşiklerin aktif ve inaktif olarak sınıflandırıldığı bir molekül kütüphanesinden, moleküler özellikler ve biyoaktivite arasındaki ilişkiyi öğrenerek yeni molekülleri aktif ve inaktif olarak tanımlama denetimli makine öğrenmesi için verilebilecek örneklerdendir. Denetimli makine öğrenmesi için en sık kullanılan algoritma örnekleri Naive bayes, karar ağacı (decision tree), destek vektör makineleri (support vector machine,SVM), sinir ağları ve lojistik regresyondur.

Bir diğer makine öğrenimi yöntemi ise denetimsiz (gözetimsiz) öğrenmedir. Denetimsiz öğrenmede, yalnızca girdi verileri algoritmaya verilir, bir çıktı verisi bulunmaz. Bu metodun başarılı uygulamaları olmasına karşın genellikle anlaşılması ve değerlendirilmesi oldukça zordur. [2,3]

Denetimsiz makine öğrenmesi için;

- Bir metinden konuyu tanımlama
- Müşterileri benzer tercihlerine göre sınıflandırma örnekleri verilebilir.

Denetimsiz makine öğrenmesi algoritmaları arasında k-ortalamlar kümelenme, hiyerarşik kümeleme, temel bileşen analiz örnek olarak verilebilir.

Diğer öğrenim metotları ise yarı denetimli öğrenme ve pekiştirmeli öğrenmedir.



Şekil 2.3. Makine öğrenmesi türleri ve örnekleri

Makine öğrenmesi sürecinin en önemli bölümü üstünde çalışılan veriyi anlamak ve bu verinin çözümlenmesi istenen problem ile ilişkisini ortaya koymaktır. Rastgele bir algoritma seçilip verilerin yerleştirilmesi verimli bir yöntem değildir. Model oluşturmaya başlamadan önce veri setinin özelliklerini belirlemiş olmak ve bunları anlamak gerekir. [2,3]

## 2.2. Kimyada Hesaplamalı Yöntemlerin Kısa Tarihi

Moleküllerin kimyasal yapıları ile biyolojik aktiviteleri arasındaki ilişki olabileceğine dair çalışmalar 1800'lü yıllara dayanmaktadır. Bu alandaki ilk önemli çalışmalardan biri ise 1863 yılında Croas, alkollerin sudaki çözünürlüğünün azalmasıyla toksik etkilerin arttığına dair bulduğu sonuçlardır. İlerleyen yıllarda da çeşitli bilim insanları tarafından düşük su çözünürlüğünün toksisiteleri ile ilişkili olabileceği bildirilmiştir. 1869 yılında Richardson, çeşitli primer alkollerin narkotik etkilerinin moleküllerin ağırlığı ile arttığını gözlemlemiştir. 1899 yılında Meyer ve Overton yaptıkları bağımsız çalışmalarda moleküllerin kullandıkları zeytin yağı-su sistemindeki partisyona katsayısının narkotik bileşiklerin toksisiteleri ile yakından ilişkili olduğu gösterilmiştir. Bir diğer önemli adım ise ileride de Hansch'ın yapacağı çalışmalara ışık tutacak Hammett'in 1935 ve 1937 yıllarında yaptığı çalışmalarıdır. Hammett bir kimyasal içindeki reaksiyon ve denge sabitleri ile ilişkilendirilebilecek elektronik sübstüent sabitlerinin geliştirmesidir. [4]

Bilgisayarın ilaç geliştirmede ilk kullanımı 1960'larda başlayan çalışmaları ile "QSAR'ın babası" olarak kabul edilen Hansch tarafından biyoaktif moleküllerin fizikokimyasal özelliklerinin bilgisayar destekli tanımlanması ve matematiksel olarak hesaplanmasıyla gerçekleştirilmiştir. Hansch ve çalışma arkadaşları, bir seri fenoksiasetik asit türevi bileşiklerin bitki büyüme düzenleyicileri üzerinde etkilerinin Hammett sabiti ve hidrofobikliği ile korelasyonunu gösterdi. Hansch, bu çalışmasında oktanol/su karışımını kullanarak moleküllerin partisyona katsayısını hesaplamıştır. Bu çalışma ile daha önceki yıllarda yapılan çalışmalar gibi sadece narkotik etkilerin değil kompleks biyolojik etkilerinde fizikokimyasal tanımlayıcılar ile modellenilebileceği gösterilmiştir. Ayrıca bu çalışma iyi bir biyolojik aktivite modelinin geliştirilebilmesi için birden fazla tanımlayıcının gerekebileceğini göstermiştir. [5]

Geçtiğimiz yıllarda, yeni aktif bileşiklerin geliştirme maliyetini azaltmak ve hızlı keşfi için bilgisayar destekli ilaç tasarım yöntemlerine (Computer Aided Drug Design, CADD) sıklıkla başvurulmuştur. CADD, reseptörler, enzimler ve taşıyıcılar dahil olmak üzere aktif bileşikler ve hedefleri arasındaki etkileşimleri simüle etmek için kullanılır. Simülasyon sonuçları, hedefler için tamamlayıcı yapılar tasarlamak ve daha sonra geniş kimyasal veri tabanları koleksiyonlarını tarayarak yeni biyoaktif bileşikler belirlemek için uygulanır. Bütün iyi yönlerine rağmen CADD metodları hala zaman alıcı ve maliyetlidir. Bu nedenle son zamanlarda büyük veri boyutlarında da

dahi hızlı ve başarılı sonuçlar veren makine öğrenimi ve alt alanı olan derin öğrenmenin kullanımı yaygınlaşmaktadır.

## **2.2. Yapay Zeka ve İlaç Keşfi**

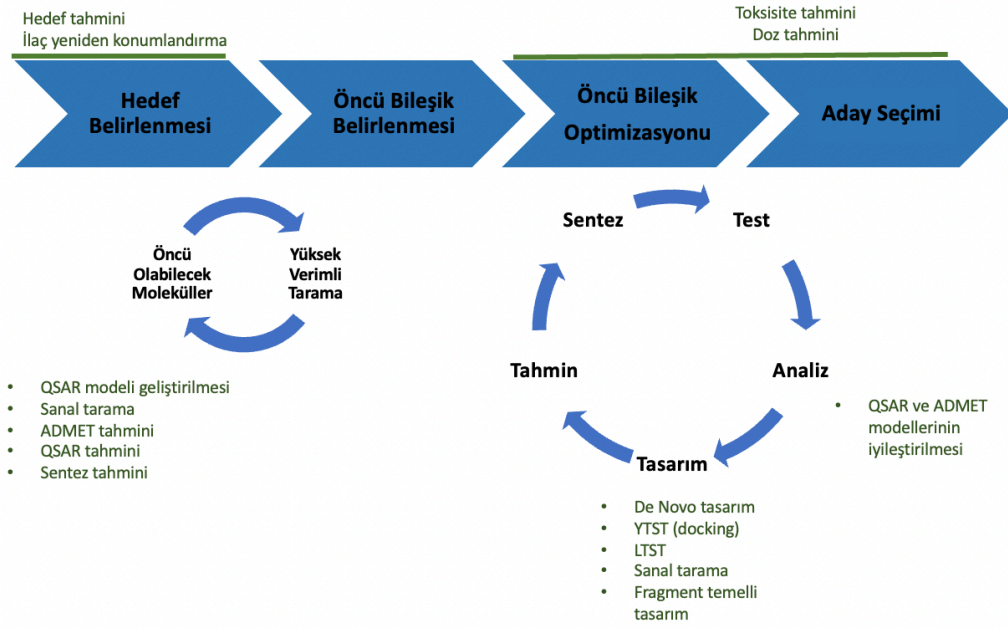
Yapay zeka bir çok alanda kullanımı yaygınlaşan bir uygulama haline gelmektedir. Özellikle doğal dil işleme alanında sağlanan gelişmeler yapay zekanın ilaç keşfinde aşamalarında da kullanılabilmesine ışık tutmuştur. Yapay zekanın, zaman içerisinde makine öğrenimi ve derin öğrenme olarak adlandırılan alt alanları gelişmiştir. İlaç araştırma ve geliştirme çalışmalarında da makine öğrenimi ve derin öğrenmenin kullanımı her geçen gün artmakta ve yaygınlaşmaktadır. Bu gelişmeler bu alanda yapılan çalışmaların hızla artmasına neden olmaktadır.

### **2.2.1. Makine Öğrenimi ve İlaç Araştırma Geliştirme Çalışmaları**

İlaç moleküllerinin yapıları ile organizmadaki etkileri arasında bir ilişki olduğu uzun zamandır bilinmektedir. İlaçların benzer yapısal özelliklerinden hareket edilerek aktivitesi daha yüksek, yan etkileri daha düşük ilaçlar geliştirilebileceği öngörülmektedir.

Moleküllerin benzerliklerin ortaya konulması manuel olarak oluşturulmaya çalışılması zaman ve maddi açıdan efektif bir yöntem değildir. Örneğin “ChEMBL” isimli internet sayfasında yaklaşık 2 milyon farklı molekülün 17.276.334 adet aktivite sonucu vardır. Bu büyük veri karşısında bilgisayar yöntemlerine başvurmak en doğru yol olacaktır.

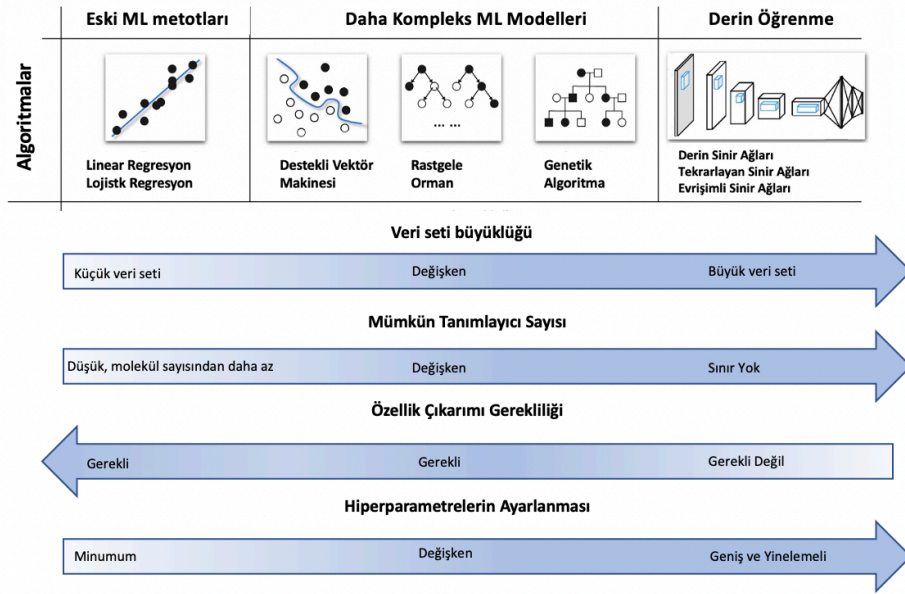
Bilgisayar yöntemleri içerisinde de makine öğrenimi günümüzde oldukça popüler hale gelmektedir. Bunun en önemli sebebi makine öğreniminin yüksek boyuttaki veriler ile işlem kapasitesinin diğer yöntemlere göre daha fazla olmasıdır. Örneğin, bir veri setindeki moleküllerin, moleküler tanımlayıcılar ile benzerliklerinin bulunması daha sonrasında ise bu benzerliklerden yola çıkarak moleküllerin aktiviteleri tahmin edilebilir. Buradaki amaç aslında yapı-etki ilişkisini bilgisayara öğreterek moleküllerin aktiviteleri hakkında bilgi edinmektir. Makine öğrenimi, sadece yapı-etki ilişkisi bağlamında değil ilaç geliştirme süreçlerinin birçokunda kullanılabilir. [6] (Şekil 2.4.)



**Şekil 2.4.** Erken aşama ilaç keşfi çalışmalarında yapay zeka ve makine öğrenimi uygulamaları. YTST: Yapı temelli sanal tarama. LTST: Ligand temelli sanal tarama

Bu alanda yapılan ilk çalışmalarda daha çok SVM, Random Forest gibi klasik makine öğrenimi algoritmaları kullanılmıştır. Son zamanlar da ise derin öğrenme metotlarının kullanımının arttığı görülmektedir. Bu algoritmalar arasında bazı farklılıklar şekilde gösterilmiştir. [6](Şekil 2.5)

Yeni molekülleri keşfetmek oldukça maliyetli ve zaman alıcı bir süreçtir. Son yıllarda makine öğrenimi yöntemleri bu sürecin daha verimli olabileceğini göstermektedir. Makine öğrenmesinin son zamanlarda çok popüler hale gelmesinin nedenlerinden biri de günümüzde her alanda çok fazla veri üretilmiş ve üreilmeye devam ediyor olmasıdır. Makine öğrenimi, bu verileri bilgiye dönüştürmenin etkili bir yoludur. İlaç araştırma ve geliştirme çalışmaları insanların yaşam kalitesi açısından elzem olduğu için bu alanda elde edilen veri sayısı fazladır. Bilgisayarlı görü ve doğal dil işleme makine öğrenimi ve derin öğrenme yöntemlerinin başarısı, ilaç geliştirmeye de ışık tutmuştur. Dikkate değer ilk örneklerden biri, 2012 yılında Merck tarafından düzenlenen Kaggle yarışmasında derin sinir ağlarının başarısıdır. 2019'da yapılan bir başka örnek çalışmada, In silico Medicine araştırmacıları tarafından "diskoidin alan reseptörü 1" için güçlü inhibitörler 21 gün gibi kısa bir sürede keşfedildi. [7]



**Şekil 2.5.** Tahmin modellerinde kullanılan makine öğrenmesi metotlarının değişimi ve veri gereklilikleri

ML, molekülün kimyasal özelliklerinin tahmini ve toksisite profilleri gibi ilaç geliştirmenin farklı aşamalarında kullanılabilir. (Şekil 2.6.) ML ile daha akılcı ilaç tasarımı yapılabilir ve bunun sonucunda yüksek aktivite, düşük toksisite ve daha az yan etki profiline sahip moleküller daha kısa sürede, daha az maliyetle elde edilebilir.



**Şekil 2.6.** İlaç geliştirmede makine öğrenmesinin kullanıldığı çeşitli alanlar

Örneğin Wang ve arkadaşları, 2019 yılında JAK2 inhibitörleri ile ilgili bir çalışma yayınlamışlardır. Bu çalışmada PubChem ve BindingDatabase kütüphanelerinden JAK2 inhibisyonu değerlerine sahip moleküllerden veri seti oluşturmuşlardır. Bu moleküllerin 3 farklı moleküler tanımlayıcıları (MACCS, ECFP2 ve Mol2Vec) kullanarak makine öğrenmesi modelleri (XGBoost) geliştirmişlerdir. Ardından ZINC veri tabanında yer alan moleküllere önce Lipinski'nin beşler kuralı uygulanarak bir kısım molekülleri eledikten sonra geliştirdikleri makine öğrenmesi modelleri ile geriye kalan moleküllerin JAK2 inhibisyon aktiviteleri tahmin edilmiştir. Modellerin aktif olarak tahmin ettiği moleküllerin, docking ve in vitro çalışmaları yapılmış ve sonuçların büyük oranda tutarlı olduğu belirlenmiştir. [8]

Başka bir çalışmada ise Cheng ve arkadaşları, blokajı kardiyotoksositeye neden olan hERG kanalları üzerine bir çalışma yapmışlardır. Bu yan etkiden dolayı piyasan çekilen veya kullanımı sınırlandırılan moleküller (astemizol, verdenafil) mevcuttur. Bu çalışmada ChEMBL molekül kütüphanesinden 7889 molekül veri seti olarak kullanılmıştır. Bu moleküller Mol2Vec ve MOE yöntemleriyle tanımlayıcıları elde edilmiştir. Ardından bir derin öğrenme modeli geliştirilerek (deepHERG) FDA tarafından ruhsatlanan 1824 molekülün hERG için aktivitesi tahmin edilmiştir ve 539 molekülün hERG aktivitesi olabileceğini göstermiştir. Bu 539 molekülün bazılarını literatür kaynakları ile doğrulamışlardır. [9]

### **2.3. Makine Öğrenimi Algoritmaları**

Bu bölümde makine öğrenimi algoritmaları ve özellikleri anlatılacaktır. Her proje için her zaman en iyi sonuç veren bir algoritma yoktur. Veri setinin özelliklerine algoritma seçimi yapılmalıdır.

#### **2.3.1. Ekstrem Gradyan Arttırma (eXtreme Gradient Boost, XGBoost)**

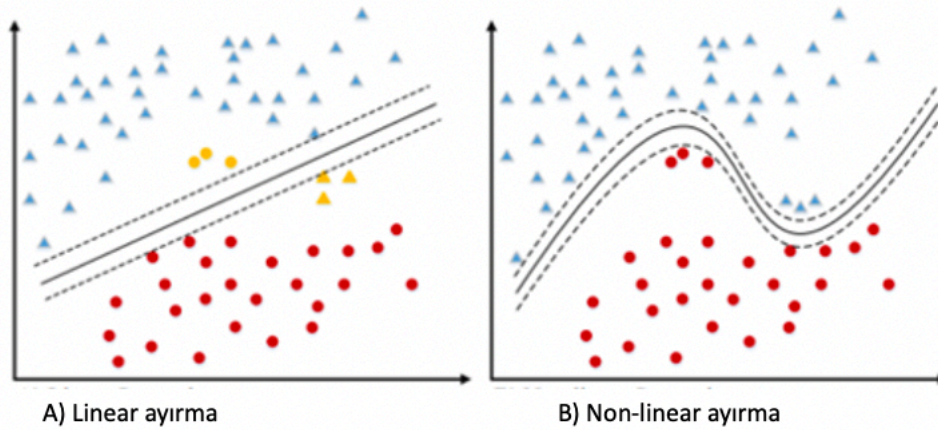
XGBoost, Gradient Boosting Karar Ağacı'na (GBDT) dayalı regresyon ve sınıflandırma problemleri için bir makine öğrenimi tekniğidir. Kayıp fonksiyonu, düzenleme ve segmentasyon noktası arama algoritması için geleneksel GBDT algoritmasını optimize eder. XGBoost modellerindeki temel parametrelerden biri, her ağacın derinliğini temsil eden "max\_depth" dir. Genel olarak, bu parametrenin değeri çok büyükse, yinelemedeki eğitim seti hatası hızla azalarak modelde aşırı öğrenmeye neden olur. Birçok farklı alanda başarılı sonuçlar göstermiştir. [10]

#### **2.3.2. Destekli Vektör Makinesi (Support Vector Machines, SVM)**

SVM, hiper düzlem ile sınıfların en yakın örnekleri (destek vektörü) arasındaki mesafeyi maksimize ederek onu doğrusal olarak sınıflandırmaya çalışır. SVM ayırmayı lineer yapmak zorunda değildir, veri setine göre polinom, radyal tabanlı veya sigmoid çekirdek fonksiyonu



olarak seçilebilir. Uygun temel işlevlerin seçilmesi kritik bir süreçtir çünkü öğrenme sürecinde farklı temel işlevler farklı şekilde çalışır. Verilerin karmaşıklığı seçim için önemli bir kriter olmasına rağmen, her bir veri tipi için belirlenmiş bir çekirdek fonksiyon yoktur. C ve gama, SVM'nin önemli parametreleridir. C parametresi, yanlış tahminlerin cezası ile ilgilidir ve bu parametrenin farklı değerleri, ilişkili hatasına bağlı olarak her bir destek vektörünü etkileyebilir. Bu anlamda, küçük C değerleri, destek vektörleri için küçük cezayı veya büyük marjları (mesafeyi) gösterir. Doğrusal olmayan çekirdeklerden gelen gama parametresi, değişkenlerin varyansı ve yanlılığı ile ilgilidir; küçük gama değerleri, büyük varyans ve düşük sapma modellerine yol açabilir ve bunun tersi de geçerlidir.[11]



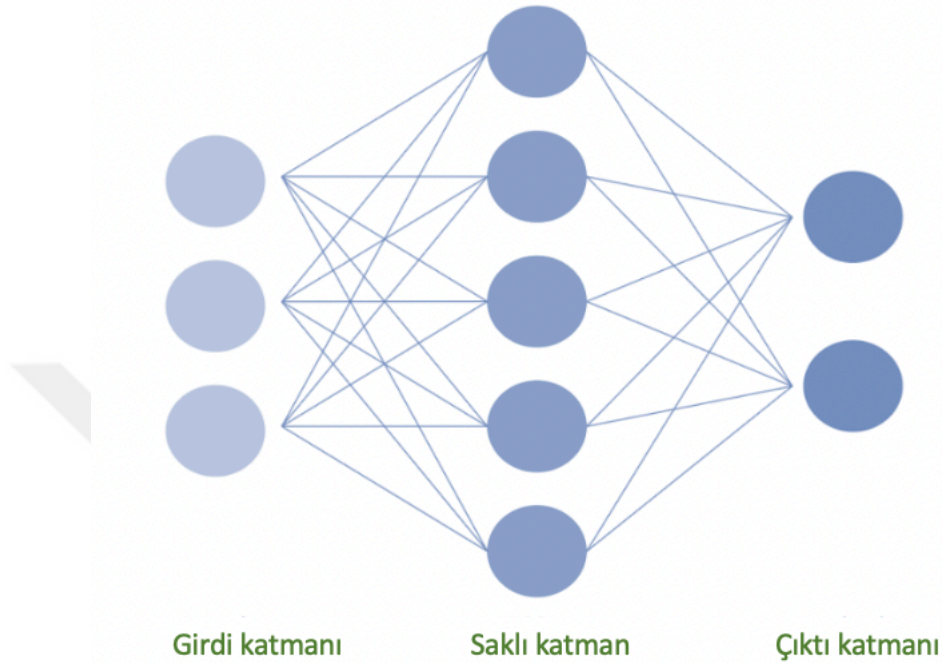
**Şekil 2.7.** Bu görselde SVM algoritmasının 2 boyutlu bir veri setinde linear ve non-linear şematize edilmişidir.

### 2.3.3. Derin Sinir Ağları (Deep Neural Network, DNN)

Derin sinir ağları (DNN), en az 2 veya daha fazla gizli katmana sahip sinir ağlarıdır. Bir DNN'de bir girdi katmanı, iki veya daha fazla gizli katman ve bir çıktı katmanı bulunur. (Şekil 2.8.) Bir DNN'deki her gizli katman, bir önceki katmandan girdi alır, verileri bir ağırlıkla çarpar, bir aktivasyon fonksiyonu uygular ve bir sonraki katmana iletir. Gizli katman düğümünün ağırlığı, girdinin çıktı için önemini gösterir. Bir DNN'nin çalışmasında iki adım yer alır. İlk adım ileri yayılma olarak adlandırılır. Bu adımda, DNN modeli, sınıflandırma veya regresyon problemleri için çıktı veya çıktılar verir. İkinci adım geri yayılımdır. Geri yayılım adımında, modelin içsel parametrelerinin türevleri kayıp fonksiyonuna göre hesaplanır ve tekrar düzenlenir. [12]

DNN'nin gizli katmanlarında bulunan düğümlerin ağırlıkları, geri yayılım adımında ayarlanır. Giriş katmanı, problem için belirlenen özelliklerden oluşur. Çıktı katmanında, çok etiketli veya

çok sınıflı sınıflandırmalar için her bir sınıf değerini temsil etmek için düğüm sayısı, sınıf sayısına eşit olmalıdır. İkili sınıflandırmalar için çıktı katmanının yalnızca tek bir düğümü olabilir. Gizli katmanlar için sigmoid veya RELU, genellikle aktivasyon fonksiyonu olarak kullanılır. [12]



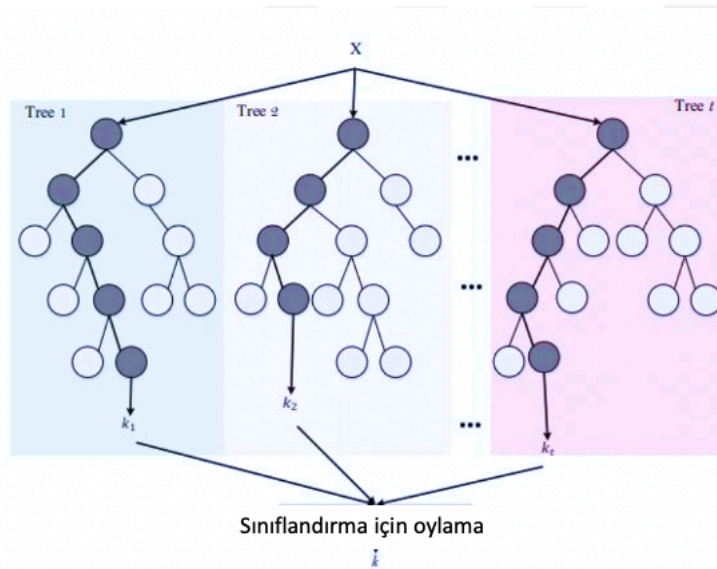
**Şekil 2.8.** Yapay sinir ağlarının mimarisi

Farklı yapıda derin öğrenme algoritmaları vardır (Burada algoritmaların isimleri İngilizce verilerek çeviriden kaynaklanabilecek hataların önüne geçilmek istenmiştir):

- a. Convolutional Neural Networks (CNNs)
- b. Long Short Term Memory Networks (LSTMs)
- c. Recurrent Neural Networks (RNNs)
- d. Generative Adversarial Networks (GANs)
- e. Radial Basis Function Networks (RBFNs)
- f. Multilayer Perceptrons (MLPs)
- g. Self Organizing Maps (SOMs)
- h. Deep Belief Networks (DBNs)
- i. Restricted Boltzmann Machines (RBMs)
- j. Autoencoders

### 2.3.4. Rastgele Orman (Random Forest, RF)

Rastgele orman algoritması, eğitim veri setinin alt kümelerinden eğitilen karar ağaçlarından oluşur. Bu karar ağaçları rastgele orman algoritması tarafından rastgele seçilir. Rastgele orman kararları, karar ağaçlarının ortalaması alınarak oluşturulur. Rastgele orman algoritması, sonuçların ortalamasını alarak karar ağaçlarının aşırı öğrenme sorununu da azaltabilir. Bu algoritmayı bir veri kümesine uygularsak, ilk adım eğitim veri kümesinden rastgele örnekler seçmek olacaktır. Ardından Rastgele orman her örnek için bir karar ağacı oluşturacaktır. Üçüncü adımda, Rastgele orman, karar ağaçlarının her birinden bir tahmin alacaktır. Sınıflandırma problemleri için rastgele orman, torbalama yöntemini (bagging) kullanacak ve en çok oylanan tahmini seçecektir. Eğitim aşamasından sonra rastgele orman, test veri seti tahminleri için eğitilmiş parametreleri kullanacaktır. Bu yöntem 1995 yılında T. Kam Ho tarafından tanıtıldı. [13]



Şekil 2.9. Rastgele orman algoritması

### 2.3.5. K-En Yakın Komşu (K-Nearest Neighbour, KNN)

KNN hem denetimli hem de denetimsiz öğrenme yöntemleri için temel ve temel makine öğrenmesi algoritmalarından biridir. Model, sınıflar açısından örnekler arasındaki benzerliği hesaplar. Ardından, tahmin sürecinde model, yeni verilerin her sınıfa benzerliğini hesaplar ve en benzer örneklere sahip olan sınıfı tahmin eder. K 1'e eşitse, veriler en yakın komşunun sınıflandırmasına atanır. Basitliğine rağmen, KNN güçlü başarılı sonuçlar verdiği çalışmalar mevcuttur. KNN'nin performansı makul bir uzaklık fonksiyonu bulmaya bağlıdır, bu da bu tekniği veriye çok bağımlı bir algoritma haline getirir. [14]

## 2.4. Moleküler Tanımlayıcılar

Bilgisayarların iletişim dili matematiktir. Yani bilgisayarların bir ifadeyi algılayabilmesi için bu ifadenin matematiksel olması gerekmektedir. Bir veri seti ile makine öğrenmesi modeli oluşturulmak istendiğinde, veri setinin matematiksel bir karşılığı olmalıdır. Bu amaç doğrultusunda, ilaç keşfinde makine öğrenmesinin kullanılması için çeşitli moleküler tanımlayıcılar geliştirilmiştir. Moleküler tanımlayıcılar, bir molekülün özelliklerini belirlemeyi ve bunları matematiksel bir ifadeye dönüştürmeyi sağlarlar. Böylece makine öğrenmesi modeli, molekülün çeşitli özelliklerini öğrenerek, başka özelliklerini tahmin edebilir.

### 2.4.1. Genişletilmiş Bağlantılı Parmak İzi (Extended-Connectivity Fingerprint, ECFP)

Genişletilmiş bağlantılı parmak izi, moleküler aktivite ile ilgili moleküler özellikleri arasındaki ilişkiyi yakalamak için tasarlanmış bir tür moleküler parmak izidir. İlaç keşfinde en popüler benzerlik arama araçları arasındadırlar ve çok çeşitli uygulamalarda etkin bir şekilde kullanılırlar. [15]

Bir molekülün ECFP özellikleri 4 basamaklı bir sürecin sonunda elde edilir. İlk olarak benzamid örneğini kullanarak bunları kısaca açıklanacaktır ve daha sonraki bir bölümde ayrıntılara anlatılacaktır. [15]

Aşağıda benzamid molekülü verilmiştir. (Şekil 2.10.) Benzamidin hidrojen hariç bütün atomlarına referans bir sayı ile gösterilmiştir. Bu örnekte 6 ile numaralandırılmış atom üzerinden parmak izi oluşturma açıklanacaktır.



Şekil 2.10. Benzamid molekülü ve atomların numaralandırılması

1. Adım: İlk olarak, moleküldeki her bir atoma ünük bir tamsayı tanımlayıcısı atanır. Bu tanımlayıcı, atom numaraları, atom kütlesi vb. gibi özelliklerin bir kombinasyonunun hash fonksiyonu ile (girdiye özel çıktı üreten bir fonksiyon) oluşturulur. Bu aşamada, atomun

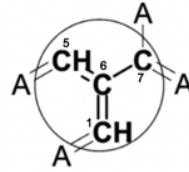
oluşturulan tanımlayıcısı yalnızca atomun kendisi ve bağları hakkındaki bilgileri temsil eder. (Şekil 2.11.)



Yineleme 0

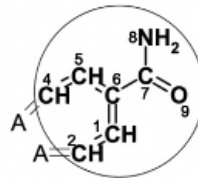
**Şekil 2.11.** Parmak izi oluşturmanın ilk aşaması

2. Adım: Bu aşamada, her atom tanımlayıcısını ve 1 bağ uzaklıktaki komşu atomlarının tanımlayıcılarını bir dizide toplar. Bu diziyi yeni, tek tamsayı bir tanımlayıcıya geri döndürmek için yine hash fonksiyonu kullanılır. Böylece, 6 numaralı atomun kendi özellikleri ve komşu atomlarının özellikleri elde edilmiş olur. Tüm atomlar yeni tanımlayıcılarını ürettikten sonra eski tanımlayıcılar yenileriyle değiştirilir. Bu güncelleme işlemi birçok kez yinelemeli olarak yapılır: İlk yinelemede, bireysel atomik özellikleri yakalama, sonraki yineleme, komşularını, bir sonrakini, komşuların komşularını vb. Yineleme sayısı seçimi kullanıcı belirleyebilir. (Şekil 2.12. ve Şekil 2.13.)



Yineleme 1

**Şekil 2.12.** Bir yinelemeden sonra, tanımlayıcı artık 6 numaralı atomun en yakın komşuları hakkında bilgi içerir



Yineleme 2

**Şekil 2.13.** İki yinelemeden sonra, tanımlayıcı artık 6 numaralı atomun komşularının komşuları (2 bağ uzunluğundaki atomlar) hakkında da bilgi içerir.

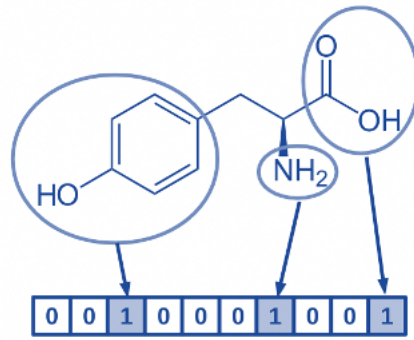
3. Adım: Bu aşamada ise önceki basamaklarda oluşturulmuş ve tekrar edilen özellikler silinir.

4. Adım: Tüm tanımlayıcılar belirli bir yineleme sayısı için hesaplandıktan sonra, son adım bu tanımlayıcıları bir bit vektörüne indirgemektir.

Bu süreç moleküldeki bütün atomlar için tekrarlanır (burada sadece 6 numaralı atom üzerinden gösterilmiştir). Her bir atom birer birer merkez olarak seçilir ve daha sonra aynı işlem tekrarlanır. Tüm süreç tamamlandığında, her atom, molekülün tüm bölümlerinden gelen temel yapıların (substructure) bilgilerini içerecek bir tanımlayıcıya sahip olacaktır.

#### 2.4.2. Moleküler Erişim Sistemi (The Molecular Access System, MACCS)

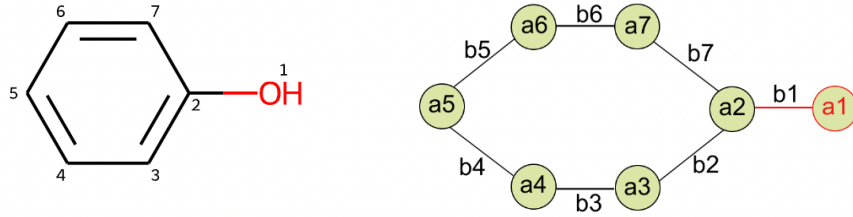
Temel yapılara dayalı parmak izleri, belirli bir temel yapının veya özelliklerin molekülde mevcudiyetine bağlı olarak bir bit dizisi oluşturur. Molekülün temel yapı listesindeki yapısal anahtar ile kapsanabildiği durumlarda başarılı sonuçlar verebilir. Bu tarz parmak izlerine örnek olarak MACCS verilebilir. SMARTS modellerine dayalı olarak biri 960 ve diğeri 166 yapısal anahtarlı olmak üzere iki çeşidi vardır. Daha kısa olanı en yaygın kullanılanıdır, ancak ilaç keşfi ve sanal tarama için önemli kimyasal özelliklerin çoğunu kapsar. (Şekil 2.14.) [16]



**Şekil 2.14.** Hipotetik olarak 10-bitlik bir temel yapı parmak izinin temsili, 3 tane bit 1 olarak işaretlenmiştir çünkü bunlar molekülde olan yapıları göstermektedir.

#### 2.4.3. Moleküler Grafikler

Herhangi bir makine öğrenimi yaklaşımında, verilerin özelliğini oluşturmak çok önemli bir adımdır. Bir grafik (graph) iki öğeden oluşur: düğümler ve kenarlar. Moleküller için atomlar düğüm, aralarındaki bağlar ise kenar olarak ifade edilir. Her düğüm ve kenar özelliklere sahip olabilir. Örneğin, bu durumda atomların özellikleri düğümlerin özellikleri ve bağların özellikleri kenarların özellikleri olacaktır. [17] (Şekil 2.15.)



**Şekil 2.15.** Moleküler grafiklerin gösterimi

## 2.5. Makine Öğrenmesinin İlaç Araştırma ve Geliştirmede Uygulamaları

### 2.5.1. Makine Öğrenmesinin Yapı Temelli Sanal Taramada Kullanımları

Yüksek verimli tarama maliyetlerini düşürmek, etkinliğini ve öngörülebilirliğini arttırmak amacıyla, öncü bileşik bulma iş akışında sanal tarama başarılı ile kullanılmaktadır.[18,19] Güçlü öğrenme ve genelleştirebilme kapasiteleri sayesinde makine öğrenmesi sanal tarama çalışmalarının çeşitli aşamalarında kullanılmıştır. Sanal tarama; ligand-hedef bağlanma bilgisinin kullanıldığı yapı temelli sanal tarama (YTST) ve hedef yapıdan bağımsız çalışmaların yapıldığı ligand temelli sanal tarama (LTST) olmak üzere ikiye ayrılır. Bu bölümde yapı öncelikli yapı temelli sanal tarama anlatılacaktır ardından ligand temelli arama ayrı bir bölümde anlatılacaktır.

YTST, hedeflenen makromolekülün (genellikle protein veya RNA) 3 boyutlu yapısal modeli ile potansiyel ligandların taramasını kapsar. Moleküler docking çalışmaları YTST çalışmalarında 1980'lerin başından beri kullanılmaktadır.[20] Birçok moleküler docking çalışmaları yüksek hesaplama maliyetleri nedeniyle serbest ligand ve rijit hedef arasında gerçekleştirilmiştir. [21] Modern donanımlar geliştirildikçe bu çalışmalar daha geniş boyutlarda da gerçekleştirilebilir hale gelmiştir (flexible ligand ve flexible target). Docking çalışmasının ilk adımı, önceden hazırlanmış bir kitaplardan bir molekülün, sterik ve fizikokimyasal özelliklere dayalı olarak bir hedefin bağlanma bölgesine sanal olarak yerleştirildiği yerleştirmedir. İkinci adımda, oluşturulan bağlanma pozunu, matematiksel bir model (skor fonksiyonu) kullanılarak afinitelerinin bir temsilcisi olarak hedef ve ligand arasındaki enerji etkileşimlerini tahmin ederek değerlendirilir. Bir pozun hangi yollarla puanlanması gerektiği çok tartışılan bir konudur. Bugüne kadar, bu yöntemler genellikle bir pozu değerlendirmek için enerji modellerini veya gözlemlenen özellikleri kullanan kuvvet alanı veya ampirik ölçümleri kullanmıştır. Son yıllarda ise çalışmalar, bilgi tabanlı ve makine öğrenimine dayalı puanlama işlevleriyle veri madenciliğine odaklandı.[22] Skorumlama işleminin ardından yüksek skorlu moleküller seçilerek deneysel veya daha pahalı hesaplamalı analizlere tabi tutulur.

Kullanılan skor fonksiyonun, docking çalışmalarının güvenilirliğinde önemli bir sınırlayıcı faktör olmaya devam etmektedir.[23] Onlarca yıllık araştırmaya ve çok sayıda puanlama işlevinin oluşturulmasına rağmen, her durumda diğerlerinden sürekli olarak daha iyi performans gösteren tek bir skor fonksiyonu yoktur ve bağlanma afinitelerini yüksek doğruluk ve hassasiyetle tahmin etmek, poz skorlanmasında en önemli sorunlardan biri olmaya devam etmektedir.[24]

Birçok skor fonksiyonu, konformasyonel entropinin (protein esnekliği gibi) yetersiz değerlendirilmesi ve çözücülerden gelen enerji katkısı ile ilgili sınırlamalarla karşı karşıyadır.[25]

Ayrıca, kuvvet alanına (force field) dayalı skorlama, bağlanma afinitesi ile bir protein-ligand kompleksini tanımlayan değişkenler arasındaki ilişkinin genellikle yetersiz, önceden belirlenmiş bir işlevsel biçimine dayanır.[26] Skor fonksiyonu çeşitli parametrelere göre ayarlanabilse de bu sorunlara tam bir çözüm değildir. Modelin bu katılığı, modelleme varsayımlarına uymayan kompleksler için yetersiz tahmin performansına yol açar. Bu nedenle, tahmin doğruluğunun iyileştirilmesi için esnek, doğrusal olmayan modellere dayalı yeni skor fonksiyonlarının geliştirilmesi gerekmektedir. [27]

Son yıllarda, yapay zeka algoritmaları YTST’da parametresiz skor fonksiyonu olarak kullanılmaya başlanmıştır. Özellik vektörleri ve bunların protein-ligand bağlayıcı serbest enerjiye katkıları arasındaki ilişki, mevcut deneysel verilerden, veriye dayalı bir şekilde örtük olarak öğrenilir; böylece, genellenebilir bir skor fonksiyonu elde etmek için yararlı doğrusal olmayan ilişkilerin çıkarılabilir. [28] SVM tabanlı ID-Score,[29] Rastgele orman temelli RF-Score[26] ve ANN-based NNScore [30] gibi YZ tabanlı önceden belirlenmemiş skor fonksiyonu, aktif ligandları gelişmiş doğrulukla tanımlamak için geliştirilmiştir.

YZ tabanlı önceden belirlenmemiş skor fonksiyonları, aktif araştırma konusudur ve çeşitli biyolojik sistemler için ligand bağlama afinitesini tahmin etmedeki uygulamalarıyla doğrulanmıştır. Bu tür modellerin en iyi nasıl oluşturulacağı, bu modellerin yeni kimyasal temsillerden yararlanıp yararlanamayacağı ve bu sistemlerin değişken özellik temsilleriyle nasıl davrandığı, akla gelen başlıca sorulardır. [26,31,32] Ayrıca oluşturulan modellerin doğrulanması gerekmektedir. Modellerin gerçek dünya performansı değerlendirilmelidir. Daha iyi skor fonksiyonlarına yönelik önemli bir adım, yanlış ligand poz tahmini riskini azaltmak için yenilikçi konformasyonel örnekleme yanı sıra bu tür yöntemlerin kullanıcı dostu docking yazılımlarına dahil edilmesi olacaktır. Birçok YZ tekniği skor fonksiyonu geliştirmek için kullanılmıştır, bunlar naive bayes, SVM, RF, yapay sinir ağları ve derin sinir ağlarıdır. YZ algoritmaları, özellikle derin öğrenme algoritmaları, gürültülü veriye karşı ve protein-ligand kompleksleri gibi yüksek boyutlu verilerde kullanışlıdır. [33–36]



Önceden belirlenmemiş skor fonksiyonu uygulayan yapay zeka modelleri, tahmin kıyaslamalarında klasik yaklaşımlara göre önemli iyileştirmeler göstermiştir.[37] Daha da önemlisi, ilgili doğrusal regresyon yönteminin doğrusal olmayan makine öğrenme yöntemleriyle, örneğin RF veya destek vektör regresyonu veya DNN'ler ve CNN'ler gibi daha karmaşık YZ modelleriyle değiştirilmesiyle, bir doğrusal skor fonksiyonu kolayca geliştirilebilir. [38–40] Her açıdan diğerlerinden daha iyi performans gösterebilecek en iyi YZ tabanlı, önceden belirlenmemiş puanlama işlevi yoktur. Sonuç olarak, skor fonksiyonlarının birleştirilmesi, sanal tarama projelerinin başarı oranını artırabilir.

Docking tabanlı YTST'de YZ uygulamaları, umut verici ve devam eden bir araştırma girişimidir. YZ ile geliştirilmiş YTST'nin geliştirilmesindeki eğilim, makine öğrenimi modelleri kullanılarak yerleştirme programlarıyla oluşturulan pozları yeniden puanlayarak iyileştirmeyi amaçlar. Örnek olarak, gelişmiş bir sanal tarama performansı elde etmek için Auto-Dock Vina RF-Score-VS ile geliştirilebilir ve genişletilebilir.[38] Otomatik ligand yerleştirme ve gelişmiş makine öğrenimi algoritmalarının birleşimi hem yanlış negatif hem de yanlış pozitif tahminlerin sayısını azaltmaya yardımcı olabilir. Bu alandaki gelecekteki çalışmaların, hedef proteinin yapısal bilgilerini ve fizikokimyasal özelliklerini (örneğin, sterik ve elektronik etkiler) içermesi beklenmektedir.

Veri kalitesi ve miktarı, seçilen moleküler temsil ve model mimarisi de dahil olmak üzere, yapay zeka modellerinin genel performansını etkileyen pek çok faktör göz önünde tutulmalıdır. YTST için faydalı skor fonksiyonlarının tasarlanmasına izin vermek için tüm bu faktörlerin iyi seçilmesi gerekir. Burada eğitim setine benzer ancak daha düşük kaliteli veya daha düşük güvene sahip yapı veya bağlanma verilerin dahil edilmesini ifade eden veri artırma (data augmentation) ile, veri dağılımından ve hacminin sağlayacağı faydalar ile artırılmış bir skor fonksiyonu performansı elde edilebilir. [41]

Molekülün reseptör etkileşimlerini yakalayan etkileşim parmak izleri, son kullanıcının belirli bir molekülün varsayılan veya kanıtlanmış bir anahtar kalıntı (key residue) ile etkileşip etkileşmediğine kolayca belirlenmesine izin verdiği için, docking işleminin yaygın olarak incelenen bir çıktısıdır; bu nedenle, genellikle bir araştırmacının en çok ilgisini çeken sabitlenmiş bir pozun özelliklerini tanımlarlar. Yakın tarihli bir çalışmada, daha kesin etkileşim parmak izlerinin belirlenmesi, mutlaka bağlanma afinitesinin daha doğru bir tahmini ile sonuçlanmadığını göstermiştir.[42] Ayrıca, seçilen bir temsilin istenen kesinliği, YZ tabanlı önceden belirlenmemiş skor fonksiyonu amaçlanan hedefe bağlıdır. Örneğin, belirli bir metal iyonunun varlığı gibi özellik bilgileri modele açıkça dahil edildiğinde, bir enzim veya reseptör ailesine özgü puanlama işlevlerinin tahmin performansının artması beklenir. Şimdilik, YZ tabanlı önceden belirlenmemiş skor fonksiyonlarının uygunluğu problem bazlıdır yani genelleştirilememiştir. Sürekli artan

erişilebilirlik ve küratörlü eğitim verilerinin bolluğu, gelecekte YZ tabanlı sanal tarama yaklaşımlarının mevcut modellerden daha iyi genelleştirilebileceğini işaret etmektedir. [43]

### 2.5.2. Makine Öğrenmesinin Ligand Temelli Sanal Taramada (LTST) Kullanımları

YZ'nın günümüzde sanal taramada en çok kullanımı ligand temellidir. LTST, kimyasal yapıları benzer bileşiklerin benzer biyolojik etkiye sahip olduğu hipotezine dayanmaktadır. [44] Genel yaklaşım şu şekildedir; belirli bir hedefe karşı aktif olduğu bilinen moleküllerin tanımlanması, molekül kütüphanesini taramasında kullanılacak bir benzerlik metodunun belirlenmesi ve ileri deneysel çalışmalar için aday olabilecek moleküllerin belirlenmesi. LTST çalışmaları için YZ uygulamaları, aktivite tahmini için regresyon modelleri ve sınıflandırma modelleri olarak ikiye ayrılabilir. Bu bölümde YZ algoritmalarının LTST çalışmalarında kullanımdan bahsedilecektir. Başlangıç olarak, başlıca YZ algoritma temelli QSAR çalışmalarının başarıları ve sınırları anlatılacaktır.

QSAR çalışmalarının ilaç geliştirmedeki hedefi, molekülün aktivitesi ile kimyasal özellikleri arasında matematiksel ilişkiyi bulmaktır. Özellikle, Random forest algoritması ile QSAR çalışmalarında ilgili bir veri seti üstünde; kolayca yorumlanabilen, tahmin performansı iyi güçlü olan modeller oluşturulmuştur, örneğin Profile-QSAR (pQSAR). [45] Bu RF algoritması, model tarafından önemli kabul edilen moleküler tanımlayıcılar ve yapısal motiflerle ilgili olarak uygun bir kimyasal ve biyolojik yorumlamaya olanak tanır. [46]

SVM, QSAR çalışmalarında kullanılan bir başka popüler tekniklerden biridir. SVM, Cathepsin L ve faktör XIIa inhibitörlerinin geliştirilmesinde kullanılmıştır.[47,48] QSAR çalışmalarında SVM'in performansında, veri setinin büyüklüğü ve parametre değerlerinin etkilerine odaklanan çalışmalar da mevcuttur.[49,50]

Çeşitli derin öğrenme algoritmaları çok hedefli çalışmalarda kullanılmıştır.[51,52] Örneğin, Dahl ve ark. eğitim setinin büyüklüğü ve heterojenliğinin çok hedefli QSAR çalışmalarında etkilerinin araştırmışlardır ve limitli sayıda veride gizli katmanlar aracılığıyla modelin kompleks hale getirilmesinin kullanışlı olmadığını gösterdiler.[53,54] Derin öğrenme yoluyla çoklu görev modellemeyi kullanan Unterthiner ve ark., biyolojik hedefler için iki milyon veri noktasından oluşan bir veri seti kullandı, eğitim performansını iyileştirmek için birkaç değişiklik ekledi ve modelin kNN ve SVM dahil olmak üzere çeşitli popüler yöntemlerden daha iyi performans gösterdiğini buldu.[55] Pande ve ark., çoklu hedeflerin sayısının sınıflandırma modelin performansına etkilerini incelemişlerdir.[56] 200'ün üzerinde hedef 1.6 milyon bileşiğin 37.8 milyon deneysel verisi ve bu moleküllerin genişletilmiş parmak izleri oluşturulmuştur. Çok hedefli olan sinir ağı, lojistik regresyon, RF ve tek hedefli sinir ağlarından daha iyi performans gösterdi. Çok hedefleri DNN'lerden kaynaklanan performans iyileştirmeleri birçok farklı grup

tarafından bildirilmiş olsa da, bu etkinin temeline odaklanan az sayıda çalışma vardır. Buna yönelik çalışmalardan birinde Xu ve ark., çok görevli bir DNN'ye gömülü bir görevin, eğitim süreci sırasında diğer QSAR görevlerinden bilgi "ödünç alabileceğini" buldu.[57]

DNN modellerinin geleneksel makine öğrenmesi modellerine göre QSAR çalışmalarında daha iyi sonuçlar gösterdiği birçok çalışma tarafından gösterilmiştir. DNN modelinin hiperparametrelerinin ayarlanmasında ise optimal bir yol henüz yoktur. Yapılan bir çalışmada, aktivasyon fonksiyonu, dropout, gizli katmanların sayısı ve sahip olduğu nöron sayısı dahil belirli hiperparametrelerin ayarlanmasının DNN modelinin performansını önemli ölçüde etkilediği gösterilmiştir. [58–60]

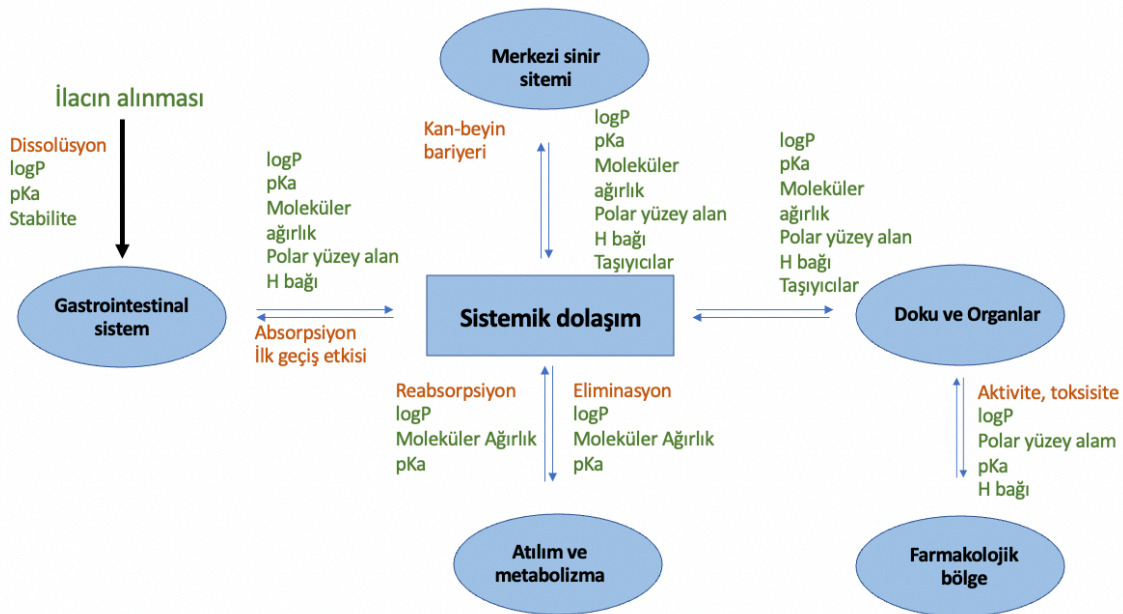
Bu çalışmalar dışında kümeleme ve benzerlik temelli molekül kütüphanelerinin taranmasında kullanılan YZ yöntemleri de mevcuttur. [43] Sonuç olarak, sofistike YZ modellerin seçimi ve tasarımının artması, bunların geniş veri setleri üstünde uygulanmasının artması LTST çalışmalarının gelişmesine yol açmıştır. Farklı makine öğrenme modellerinin karşılaştırılmasının zorluğu ise çalışmaların standardize değerlendirme eksiliğinden kaynaklanmaktadır. Yapay zekadaki kimya ile ilgili sorunlar için bunu daha sağlam bir temele oturtma girişimleri devam etmektedir. [61] Genelleştirilmiş ve sağlam YZ modeller oluşturmak amacıyla, geliştirilmiş bir model için uygulanabilirlik alanının resmi bir tanımını sağlamak önemlidir. Bir diğer önemli konu ise yapılacak çalışma için hangi moleküler temsil metodunun kullanılacağıdır. Son yıllarda, karmaşık moleküler temsil yöntemleri geliştirilmiş ve tanıtılmıştır. Örneğin, küçük molekülleri sonlu boyutlu özellik vektörleri olarak temsil etmek için CNN'leri kullanan "nöral parmak izi" yöntemleri geliştirilmiştir. [62] Farklı bir soyutlama düzeyinde, moleküler kuantum mekaniksel hesaplamalar son yıllarda çok daha rutin hale geldi ve kabul edilebilir bir zaman çerçevesi içinde makul ölçüde yüksek doğrulukta moleküler temsiller oluşturabilir. Bu nedenle bu yöntemler, prensipte moleküllerin elektronik ve geometrik özelliklerini ve etkileşimlerini daha basit temsillerden daha eksiksiz olarak ifade edebilen önemli moleküler tanımlayıcı kaynaklardır. [63,64]Yakın gelecekte, alana özgü moleküler temsillerin YZ yöntemleriyle kombinasyonlarının ana LTST teknikleri haline gelmesini bekleniyor.

### **2.5.3. Fizikokimyasal ve ADMET Özelliklerin Tahmininde Makine Öğrenimi Uygulamaları**

Uygun olmayan ADMET özellikleri, ilaç adayı moleküllerin geç aşamalarda başarısızlığa ve onaylanmış ilaçların geri çekilmesine neden olabilmektedir. [65,66] Bir bileşiğin ADMET özelliklerinin ilaç geliştirme çalışmalarının başında belirlenmesi önemlidir.[67,68] Bu nedenle, bu sorunları azaltabilecek *in silico* modellerinin geliştirilmesi endüstri ve akademik çalışmaların odak noktalarından biri haline gelmiştir. Son yıllarda YZ teknikleri ile, moleküller özelliklerden hareketle ADMET özelliklerini tahmin eden modeller geliştirilmiştir. Moleküllerin özellikle

farmakokinetik özelliklerine etki eden, fizikokimyasal parametreleri (LogP, sudaki çözünürlüğü gibi) hızlıca tahmin eden modeller geliştirilmiştir.

İlaç adayı bir bileşiğin fizikokimyasal özellikleri, Absorpsiyon,-Dağılım-Metabolizma-Eliminasyon-Toksisite (ADMET) özelliklerini, potensi ve selektivitesini etkileyerek tedaviye girme sürecini önemli ölçüde etkilemektedir. [69,70] Örneğin, moleküllerin benzer fizikokimyasal özellikleri, hedef protein ailelerine göre sınıflandırmak için kullanılmıştır (nükleer reseptör ligandlarının ortalama log P değeri ve moleküler ağırlığı sırasıyla 4.8 ve 573 Da'dır).[71] Küçük moleküllü ilaç adayları, etki bölgelerine ulaşmak ve hedeflerine bağlanmak için yeterli çözünürlük ve geçirgenlik göstermelidir. Bu nedenle, istenen farmakokinetik ve farmakodinamik profillere sahip bileşiklerin tasarımı için fizikokimyasal özelliklerin daha iyi anlaşılması ve doğru bir şekilde tahmin edilmesi oldukça önemlidir. (Şekil 2.16.)



Şekil 2.16. Farmokinetik süreç ve önemli olan moleküler özellikler

### 2.5.3.1. Fizikokimyasal Özelliklerin Tahmini

Lipofilite, birçok farmakokinetik sürecin düzenlenmesinde görev alır ve en önemli fizikokimyasal özelliklerden biridir. [72] Özellikle, molekülün hücre membranından geçmesinde ve dolayısıyla molekülün taşınmasını, dağılmasını ve atılımını etkiler. Aynı zamanda, ilaç moleküllerinin makromoleküllere (albümin gibi) bağlanmasını güçlü bir şekilde etkiler, metabolik ve toksik süreçleri ve ayrıca vücuttaki hemen hemen tüm ilaca bağlı biyolojik süreçleri etkiler. Lipofilikliğin nicel karakterizasyonu için altın standart, oktanol-su dağılım katsayısının

(log P) veya pH'a bağlı dağılım katsayısının (log D) logaritmasıdır; alternatifler arasında lipozom/su partiyonu ve hareketsiz yapay membran kromatografisi bulunur.

ADMET Predictor ve ALOGPS gibi programlarda oktanol-su bölüm katsayısını tahmin etmek için sinir ağları kullanılmıştır. Özellikle, ileri beslemeli sinir ağı unsurlarını ve kNN yaklaşımını birleştiren birleştirici bir sinir ağına dayanan ALOGPS programının, log P değerlerini veya düşük moleküler ağırlıklı bileşikler için güvenilir bir şekilde tahmin ettiği kanıtlanmıştır ve birkaç grup tarafından uygulanmıştır. [73-75]

Geniş bir çeşitlilikteki veri setleriyle log P değerini güvenilir bir şekilde tahmin etmek için oktanol-su sistemini referans kabul eden YZ temelli protokoller oluşturulmuştur. Bu YZ tabanlı protokoller doğruluk ve verimlilik açısından farklılık gösterir ancak tümü, gerçek çözümler veya moleküler simülasyonlar için kondüktör (COSMO) benzeri tarama modelleri gibi diğer fiziksel tabanlı yaklaşımlara kıyasla uygulanabilirlik alanını sınırlayabilen deneysel veriler üzerinde eğitilmiştir.[43]

Yeterli su çözünürlüğü (log S), oral yoldan verilen ilaçlar için çok önemlidir, çünkü gastrointestinal sistemden emilmeden önce gastrointestinal sıvılarda çözümleri gerekir.[76] Sudaki çözünürlüğü tahmin etmek için çeşitli in silico yöntemler geliştirilmiştir. Bu çalışmalar temellerine göre 3'e ayrılabilir; kuantum kimyasal özelliklerine, moleküler mekanik simülasyonlarına ve tanımlayıcı temelli makine öğrenimi algoritmaları.[77,78] Bu bölümde sadece log S değerinin makine öğrenmesi ile tahminine yer verilecektir.

Suda çözünürlüğün tahmini, genellikle (i) fizikokimyasal özellikler, (ii) atomik veya grup etkileri, (iii) kuantum kimyasal özellikler ve (iv) yapı hakkında bilgileri içeren dört ana tanımlayıcı grubuna dayanır. Fizikokimyasal özellik tabanlı tanımlayıcılar, deneysel olarak ölçülebilen veya hesaplama yaklaşımlarından elde edilebilen özelliklerle ilgilidir. (log P veya Erime noktası gibi)[79,80] Atomik veya grup etki tanımlayıcılarını kullanan yaklaşımlar, bir moleküldeki atomların veya grupların varlığını saptar ve mevcut deneysel verilerle korelasyon yoluyla her birinin log S'ye katkısını tahmin eder.[81] Yapıya dayalı yöntemler, moleküler topoloji, bağlantı veya fragment bilgisi gibi kimyasal yapısal bilgilere dayalı olarak çözünürlüğü tahmin eder.[82] Fühner'in çalışmasına dayanan temel bir özellik olan moleküler boyuta ek olarak, hidrojen bağlarının sayısı ve çeşitli moleküler bağlantı indeksleri dahil olmak üzere, log S ile ilgili diğer moleküler özellikler de yıllar içinde keşfedilmiştir. [83,84]

Moleküler yapılar, basitlik ve tahmin gücü arasındaki iyi dengeden ötürü klasik ikili parmak izi çeşitli bakış açılarıyla temsil edilmektedir. Örnek olarak genişletilmiş bağlantılı parmak izi (daha önceki bölümde bahsedilmişti) ve moleküler graph'lardan elde edilen nöral parmak izleri verilebilir. [85]

Birçok tanımlayıcı geliştirilmiş ve YZ modellerine uygulanmış olsa da hiçbiri mükemmel bir doğrulukla çözünürlüğü tahmin edememektedir. Farklı tahmin yöntemleri kullanan 4 metotun (Çoklu lineer regresyon, Yapay sinir ağları, kategori oluşumu ve 4 adet ticari olarak bulunan model) kaliteleri araştırılmış ve mükemmel tahmin yapabilen bir modelin olmadığını göstermişlerdir ancak çoklu lineer regresyonun biraz daha iyi performans gösterdiği ve kompleks modellerde aşırı öğrenme olduğu gösterilmiştir. Mevcut modellerin sınırlı tahmin yeteneği için sıklıkla belirtilen bir başka neden ise, literatürde birden çok kaynaktan toplanan çözünürlük verilerinin tahmini kök ortalama kare hataları (RMSE'ler). 0,6–0,7 log S birimleri arasında değişen deneysel veriler olmasıdır.[86]

Bununla birlikte, Palmer ve Mitchell tarafından yakın zamanda yapılan bir araştırma, deneysel verilerin kalitesinin, ilaca benzer moleküllerin su çözünürlüğünü tahmin etmede aşırı sınırlayıcı bir faktör olmadığını göstermiştir. Yazarlar, yüksek tekrarlanabilirlik ve yaklaşık 0.05 log S birim standart hata ile 132 ilaç molekülünün içsel sudaki çözünürlüklerini içeren bir CheqSol çözünürlük veri seti kullanmışlardır. Karşılaştırmacı olarak, yayınlanmış literatürden birkaç farklı kaynaktan 0,6–0,7 log S belirsizlikle çıkarılan çok daha gürültülü bir veri seti de kullandılar. Beklentilerinin aksine, CheqSol deneysel verilerinden türetilen modeller, yayınlanan verilerden türetilenlerden daha doğru değildi. Bu nedenle, ortaya çıkan RMSE'yi, yaygın olarak önerildiği gibi deneysel ölçümlerdeki bir belirsizlik ürünü yerine QSPR yöntemlerindeki (algoritmalar ve/veya tanımlayıcı kümeler) bir eksikliğin göstergesi olarak kabul ettiler. Bu çalışmada veri kümesi boyutunun katkısı hesaba katılmıştır, ancak yazarlar alt kümedeki çeşitliliğin önemine ilişkin tahminin dikkate alınmadığını belirtmektedir.[87,88]

Genel olarak, su çözünürlüğü, küçük moleküllü ilaçların ve ilaç adaylarının önemli bir fizikokimyasal özelliği olmasına rağmen, doğru bir şekilde tahmin etmek zor olmaya devam etmektedir. Her tahmin yaklaşımının kendi yararları ve sınırlamaları vardır ve en iyi seçim büyük ölçüde çalışmaya bağlıdır. Yüksek verimli moleküler keşif için verimli yapay zeka protokolleri, uygun ve yeterince güvenilir bir yaklaşım olarak kabul görmektedir. Bununla birlikte, birkaç spesifik ilaç molekülünün (moleküllerinin) altında yatan prensiplerin spesifik bir fiziksel ve mekanik anlayışı istendiğinde, araştırmacılar tipik olarak kuantum mekaniğine ve/veya moleküler mekaniğe dayalı fiziksel olarak daha anlamlı yöntemleri kullanmayı tercih etmektedir. [43]

In vitro membran geçirgenlik verilerinin değerlendirilmesi, farmasötik olarak ilgili biyolojik işlemlerin çoğunda pasif transmembran difüzyon ve/veya aktif taşıma mekanizmaları yoluyla moleküllerin biyolojik membranlardan geçirgenliğinin önemine dayanan ilaç keşfi ve geliştirme çabaları için çok önemlidir. [89] Moleküllerin hücre zarından geçiş karakterini tahmin etmek için birçok hesaplamalı yöntemler geliştirilmiştir. Bu amaçla, geniş bir kimyasal çeşitlilikte moleküller

ve deneysel verilerin yer aldığı veri setleri ile YZ modelleri oluşturulmuştur. Bu veriler çeşitli in vitro geçirgenlik hücre dizilerinden (MDCK, PAMPA, Caco-2 gibi) elde edilmiştir. [90–92]

Veri set(ler)i toplanırken dikkatli olunmalıdır. Model eğitim sürecine koşul bilgisi entegre edilmediği sürece, yalnızca tutarlı koşullar altında ölçülen geçirgenlik verileri toplanabilir. Aynı zamanda, hücre hatlarında ifade edilen farklı taşıyıcıların etkisinin yanı sıra çeşitli bileşik kütüphanelerinin kullanımı da dikkate alınmalıdır.

### 2.5.3.2. ADMET Tahminleri

Erken ADMET tahmini, araştırmacıların iyi ilaç adayları seçmesine ve geliştirme süreci boyunca molekülün ilaca benzerliği konusunda yardımcı olur. Son yıllarda, YZ destekli ADMET tahminleri alanında birçok ilerlemeye gerçekleşmiştir ve derin öğrenme ve büyük verilere dayalı yeni nesil tahmin modelleri, laboratuvarından kliniğe ilaç keşif ve geliştirme sürecine fayda sağlamaktadır.

İlaç absorpsiyonu, ilaç moleküllerinin bir veya daha fazla biyolojik membran bariyeri yoluyla absorpsiyon bölgesinden kan dolaşımına absorbe edildiği adımdır. Çeşitli fizikokimyasal özelliklerin, ilaç uygulama yolu, formülasyon tercihi ve fizyolojik faktörler tarafından etkilenir. [93] Absorpsiyon için tahmin modelleri fizikokimyasal ve fizyolojik olmak üzere 2 alana ayrılır.

İnsan bağırsak absorpsiyonu (İBA), en etkili ADMET özelliklerinden biridir ve oral biyoyararlanımı doğrudan etkiler. İn vivo ve in vitro deneysel analizlerden elde edilen verilere dayalı olarak İBA'yı tahmin etmek için bir dizi hesaplamalı sınıflandırma ve regresyon modelleri oluşturulmuştur. Makine öğrenmesi de bu alanda oldukça sık kullanılmaktadır. [94,95] Genel olarak, yayınlanmış çalışmaların büyük çoğunluğunda, İBA için tahmin doğruluğu kabul edilebilir bir aralıktadır ancak en önemli sorun, mevcut sınırlı verilerin neden olduğu aşırı öğrenmeye yatkınlıktır. [43] Küçük veri seti problemini azaltmak ve daha güvenilir performans elde etmek için çeşitli teknikler uygulanmıştır. Örneğin, Basant ve ark., yakın zamanda, çeşitli kimyasalların absorpsiyonunu tahmin etmek ve güvenilir bir model elde etmek amacıyla topluluk öğrenme tekniklerinin kullanıldığını bildirmiştir (torbalı: bagging ve gradyan: gradient boosted trees). [96] Önerilen topluluk öğrenimine dayalı tahmin modelleri, birkaç katı büyüklüğünde dışarıdan bir veri seti ile değerlendirildi ve sağlam İBA modelleri üretme konusunda başarılı bir performans sergiledi. Alternatif bir yaklaşım kullanarak, Shin ve ark. yakın zamanda, moleküler tanımlayıcılardan kimyasal bileşiklerin absorpsiyon potansiyelini tahmin eden, özellik mühendisliğine (feature engineering) ihtiyaç duymadan mevcut yöntemlere göre mütevazı bir gelişme gösteren, DNN tabanlı bir *in silico* modeli sundu. [97]

Aşırı öğrenme sorununu ele almak için alternatif bir yöntem ise, büyük ve çeşitli bir veri seti kullanmaktır. Wang ve ark. Yaptıkları çalışma pratik uygulamalarda ilaç adayları için İBA tahmini

için faydalı bir veri seti olarak hizmet edebilir. 1272 molekülden oluşan nispeten büyük ve yapısal olarak çeşitli Caco-2 hücre geçirgenliği veri seti, birkaç filtreleme kuralı altında dikkatlice toplandı. Bu veri setinin faydası, yazarlar tarafından sağlam ve güvenilir tahmin modellerinin üretilmesiyle onaylanmıştır.[98]

Dağılım, ilaç moleküllerinin kan dolaşımından vücut dokularına, özellikle organlara, etkinin gerçekleşmesinin beklendiği yerde difüze olduğu veya aktif olarak taşındığı süreçtir. Absorpsiyonu etkileyen (lipofilite, molekül büyüklüğü, pH, iyonizasyon) birçok faktör dağılımında farklı dokularda ne oranda olacağını etkilemektedir. Ayrıca, insan serum albümini gibi plazma proteinlerine bağlanmada dağılım sürecinde önemlidir. Plazma protein bağlanması (PPB), yalnızca bağlanmamış ilaç molekülleri biyolojik hedeflerine ulaşabildiği ve amaçlanan terapötik etkileri gösterebildiği için ilacın dağılım hacmini, farmakolojisini ve farmakokinetiğini etkiler. Plazma proteinlerine bağlanma oranı dağılım modelleri oluşturmak için önemli bir parametredir. [99]

Son zamanlarda, Basant ve ark. 930 bileşikten oluşan büyük bir PPB veri seti için dört kategorili bir sınıflandırma ve çeşitli bileşiklerin plazma protein bağlanma afinitesinin tahmini için topluluk makine öğrenme modelleri geliştirdiler ve %93'lük bir sınıflandırma doğruluğu ve 0.92'lik RMSE elde ettiler. [100] Sun ve ark.,967 molekül içeren heterojen bir veri setinden plazma proteine bağlanma fraksiyonu tahmin eden altı tane makine öğrenmesi modelleri (RF, boosting tree, çoklu lineer regresyon, kNN, SVM, yapay sinir ağları) oluşturmuşlardır. Bu çalışmada 26 fiziokimyasal ve yapısal tanımlayıcılar kullanarak, bütün modellerde 0.126 ile 0.178 arası ortalama mutlak hataya ulaşmışlardır. [101]

İlaç dağılımında birçok faktör belirlenmesine rağmen, kan-beyin bariyeri, plasental bariyer gibi konular hakkında yapılan çalışmalar az seviyededir. Kan-beyin bariyeri, kanın içeriğindeki değişimlerden beyin homeostazını korumak için önemli bir rol oynar. Aynı zamanda bu bariyer, santral sistemi hastalıklarının tedavisi için ilaç geliştirmenin önemli zorluklarından da biridir. Bir molekülün kandan beyine geçişinin oranı çoğunlukla membran geçirgenliğiyle belirlenir. Son yıllarda, yüksek maliyetli, zaman alıcı ve deneysel olarak zorlu olan kan-beyin bariyeri geçişi sürecini minimize etmek için makine öğrenmesi teknikleri kullanılmaktadır. Kan-beyin bariyeri penetrasyonunu ölçmek için çok sayıda *in vivo* ve *in vitro* parametre mevcut olmasına rağmen, en popüler ve iyi bilinen parametre beyindeki ilaç konsantrasyonlarının kandaki konsantrasyona oranıdır. Bu oranın büyüklüğü, beyinde daha yüksek bir göreceli konsantrasyona karşılık gelir. Birkaç tanımlayıcının kan-beyin bariyeri geçirgenliği ile ilişkili olduğu gösterilmiştir ve fizikokimyasal özellikler (örneğin, lipofiliklik ve iyonizasyon durumu), yük tanımlayıcıları ve geometrik/topolojik özellikler (örneğin, hidrojen bağı donörleri ve alıcılarının sayısı, polar yüzey alanı ve dönebilen bağların sayısı) dahil olmak üzere tahmin modelleri oluşturmak için yaygın



olarak kullanılmıştır. Moleküler tanımlayıcılar ve ölçülen kan-beyin bariyeri geçirgenliği arasında güvenilir bir ilişkinin kurulması, mevcut kamuya açık veri setleri, verilerin hem miktarı hem de tutarlılığı açısından bir miktar eksik olduğundan, zorlu olmaya devam etmektedir. Ayrıca, kan-beyin bariyerinin geçişini sınırlayan karmaşık mekanizmalara ilişkin anlayışımız, örneğin bağırsak geçirgenliği bilgimizle karşılaştırıldığında hala nispeten sınırlıdır. Ökaryotik sistemlerdeki geçirgenlik mekanizmaları temel olarak pasif difüzyon, akış taşınması ve aktif taşımadan oluşur. Pasif difüzyon, kan-beyin bariyerinden ilaç geçişinin en yaygın mekanizması olarak kabul edilir ve bu nedenle, bugüne kadar oluşturulan penantrasyon tahmin modellerinin çoğu, diğerlerinden ziyade bu mekanizmaya odaklanmıştır. [43,99]

Metabolizma, ksenobiyotiklerin genellikle karaciğerde uğradığı kimyasal dönüşüm sürecinin ifade etmektedir. Metabolizma, genellikle enzim aracılı bir süreçtir ve metabolik enzimlerin çok yönlülüğü, substratların ve inhibitörlerin çeşitliliği ve metabolik süreçlerin karmaşıklığı göz önüne alındığında, bu süreci tahmin etmenin zorluğunu gözler önüne sermektedir. Bir ilacın metabolizma süreci, vücutta göstereceği toksik ve terapötik etkileri doğrudan etki edebilir. Birçok ilacın etkinliğinin kaybı, kendine özgü advers ilaç etkileşimleri, hepatotoksisite ve ilaç-ilac etkileşimleri gibi birçok süreçte kilit rol oynar.[102]

Yapay zekanın metabolizmayı öngörme alanındaki uygulamaları üç ana kategoriye ayrılır: (1) metabolizma bölgelerinin yerini tahmin etmek, (2) metabolizmadan sorumlu spesifik izoformları tahmin etmek ve (3) metabolik farmakokinetiği tahmin etmek.[43]

Metabolizma bölgelerinin tahmini, olası metabolitlerin türetilmesi için anahtar bilgiler sağladığı ve sonuç olarak daha güvenli ve daha etkili ilaçların geliştirilmesi için sağladığı faydalar sebebiyle bir ksenobiyotik için çok önemlidir. Şu anda, metabolizma bölgelerini tahmin etmek için birkaç popüler YZ tabanlı model geliştirildi ve tatmin edici doğruluk göstermektedirler. Örneğin, RS-predictor, hiyerarşik tanımlayıcılar (393 kuantum kimyasal ve atom bazlı tanımlayıcılar ve 148 topolojik tanımlayıcı) kullanarak ilaç benzeri moleküllerin CYP450 enzimi aracılığıyla, molekülün neresinden metabolizmaya uğrayacağı potansiyel bölgeleri tahmin etmektedir. [103-105]

Genel olarak, hızlı ve etkili ilaç metabolizması tahmin modellerine olan talep artmaya devam etmektedir. Metabolizma bölgelerini ve CYP450 aracılı metabolizmadan sorumlu spesifik izoformları tahmin etmeye yönelik metodolojiler önemli ölçüde iyileşmiştir. Bu arada, araştırmacılar, CYP450 dışında enzimlerin aracılık ettiği metabolizmayı tahmin etmeye yönelik çalışmalar az sayıdadır bunun nedeni ise, verinin yeteri büyüklükte olmamasıdır. Halihazırda mevcut modellerin çoğu (CYP450 için olanlar dahil), yalnızca oluşabilecek birincil metabolitler için metabolizma tahminleri yapmaktadır. Bununla birlikte, birçok birincil metabolit, bir daha faz 1 veya faz 2 metabolizasyonu sonucunda ikincil toksik metabolitlerine dönüşebilmektedir.

(Örneğin faz 1 aşamasında poliaromatik hidrokarbonların epoksit yapılarına dönüşmesi) Bu nedenle, bu reaksiyonların termodinamiğini ve kinetiğini ve ayrıca metabolik biyotransformasyonda yer alan ayrıntılı mekanizmayı taklit etmek için kuantum ve moleküler mekanik yöntemleri kullanılmıştır. [43,102]

Kuantum kimyası, ilaç optimizasyon kampanyalarında geriye dönük tahminler için sürekli doğruluk sağlayarak bilgilendirici moleküler temsiller oluşturmadaki faydasını kanıtlamıştır. Metabolitleri daha gerçekçi bir şekilde tahmin etmek ve ilaç geliştirmenin erken aşamalarında sorunlu bileşikleri belirlemek için, birleşik, kimyasal olarak bilinçli bir metabolik platform sağlamak için mevcut metabolomik ve toksisite bilgilerini kimyasal ve enzimatik reaksiyon mekanizmalarının moleküler bilgisi ile entegre etmek faydalı olacaktır. [43]

Eliminasyon, ilaç moleküllerinin vücuttan atılma sürecidir ve ilaçların ve metabolitlerinin plazma konsantrasyonları doğrudan etkiler. Eliminasyon, her biri bir dizi farklı süreçten oluşan biliyer ve renal atılım gibi birkaç eliminasyon yolunu içeren karmaşık bir süreçtir. Örneğin, renal atılım; glomerüler filtrasyon, aktif sekresyon ve reabsorpsiyon süreçlerini kapsar. [106] İlaç atılım süreçlerinin karmaşıklığı, bugüne kadar YZ tabanlı moleküllerin atılımını tahmin eden modellerin geliştirilmesini engelledi. Berellini ve ark. 754 bileşikten elde edilen verilere dayanarak insan plazma klirensini tahmin etmek için doğrusal bir kısmi en küçük kareler tahmincisi oluşturmak için fizikokimyasal tanımlayıcılar ve moleküler yapı parçaları kullandı. [107]Tahmin edici, 2.1'lik bir geometrik ortalama katlama hatası verdi. Kusama ve ark., insanlarda onaylanmış 141 ilacın ana klirens yollarını tahmin etmek için ilaçların kimyasal yapılarından hesaplanan bir dizi tanımlayıcıya sahip SVM tabanlı modeller geliştirmişlerdir ve modellerin yüksek tahmin performansı gösterdi belirlenmiştir.[108]

İlaç toksisitesi, bir bileşiğin etkisi veya metabolizması nedeniyle bir organizm üzerindeki olumsuz etkiyi ifade eder. Toksikite analizleri, ilaç keşif ve geliştirme döngüsünde yer alan en önemli ve zorlu adımlardan biridir. [109] Güvenilir, yüksek verimli testler pahalıdır; bu nedenle, büyük ölçekli in vivo ve in vitro çalışmalara hızlı, ucuz ve güvenilir bir alternatif sağlayan hesaplamalı yöntemler yüksek talep görmektedir. Birçok, LimTox, pkCSM, admetSAR, Toxtree ve DL-AOT gibi çevrimiçi toksisite tahmini yapan araçlar geliştirilmiştir. [110–114]Bu alanda geliştirilen YZ modelleri toksik bileşikler sınıflandırma veya toksolojik hasarı tahmin üstüne odaklanmıştır. Bu algoritmaların bir kısmı kimyasal benzerliği göz önünde bulundururken diğer kısmı ise moleküllerin kimyasal özellikleri üzerinden tahminleri gerçekleştirmişlerdir.

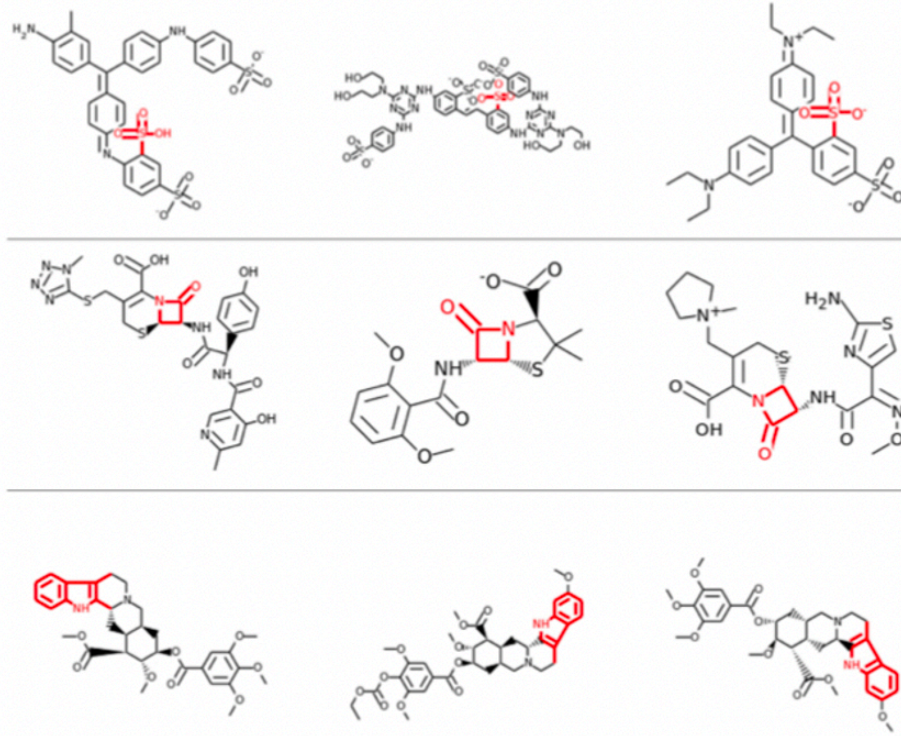
Kimyasal benzerliğe dayalı yaklaşımda, bileşikler arasındaki ikili benzerliklerin bir matrisini hesaplayarak bir molekülün toksisitesini tahmin eder; bu, daha sonra bir sınıflandırma atamak için en KNN ve SVM gibi makine öğrenmesi algoritmaları kullanılabilir. Çoğu sanal tarama çalışmasına benzer şekilde, altta yatan hipotez, benzer yapıların benzer biyolojik etkiler gösterme

potansiyelidir. Kimyasal özelliğe dayalı yaklaşımlarda ise, girdi özelliklerini seçerek veya ağırlıklandırarak toksisiteyi tahmin eder. Örnek olarak, RF ve naive Bayes modelleri verilebilir.[115–118]

Kimyasal özelliğe dayalı yaklaşımda, hedef ile ilişkili olabilecek özelliklerin belirlenmesi başarılı bir tahmin için gereklidir. Benzerlik yaklaşımında ise, iki molekülün benzerliğini özellik vektörlerinden veya 2 veya 3 boyutlu grafiklerden ifade edebilecek bir benzerlik ölçüsü gerekmektedir.

Amerika Birleşik Devletleri federal kurumları [Ulusal Sağlık Enstitüleri (NIH), Çevre Koruma Ajansı (EPA) ve Gıda ve İlaç İdaresi (FDA)] tarafından başlatılan Tox21 Veri Yarışması, bu alandaki en büyük girişim olmuştur. Amacı, toksisiteyi tahmin etmede çeşitli hesaplamalı yöntemlerinin performansını değerlendirmek ve bugüne kadar *in vitro* ve *in vivo* deneylerin sayısını azaltmada bu araçların potansiyel değerini değerlendirmek için tasarlanmış bilimsel bu topluluk kendi oluşturdukları veri setini açık kaynak haline getirmişlerdir. Mayr ve ark., toksisiteyi tahmin etmede çok görevli derin öğrenme yöntemleri ile oluşturdukları DeepTox modeli ile en başarılı sonucu elde etmişlerdir.[119] DeepTox, derin öğrenmenin toksik etkiye neden uzmanların daha önceden belirlediği özellikleri otomatik olarak öğrendiğini belirlemişlerdir. Şekil 2.17.'de gösterildiği gibi, modeldeki ilk katmanlar, sülfonik asit grupları gibi toksikoforların küçük bileşenlerini öğrenme eğilimindeyken (ilk gizli katmanda öğrenilen Şekil 2.17.'in 1. ve 2. satırlarında vurgulanan bölümlere bakın), oysa daha derin katmanlarda, bu bileşenlerin alt yapıya veya toksikofor yapının tamamına dönüşmüştür (daha derin bir katmandan öğrenilen Şekil 2.17.'deki 3. satıra bakın).

Genel olarak, uygulanabilir ve güvenilir bir ADMET tahmin modeli oluşturmak için çok daha fazla araştırmaya ihtiyaç vardır. Çok türlü toksisite verilerinin dahil edilmesinin önemi ve *in silico* ile *in vitro* (veya *in vivo*) korelasyonlarla ilgili bilinen sorunlar gibi, güvenilir YZ modelleri oluşturulurken çeşitli faktörler göz önünde bulundurulmalıdır. Mevcut YZ tabanlı ADMET modelleri, biyolojik olarak ilgili sistemlerde *in vitro* ölçümlerin yerini almak için henüz yetersiz doğrulukları olsa da ADMET ile ilgili sorunları çözmek için gereken sentez döngülerinin sayısında bir azalma sağlayarak, medisinal kimyayı doğru yöne itmek için faydalıdır. Bu yaklaşımın önemli bir ön koşulu, uygulanabilirlik alanlarının doğrulanması ve dikkatli bir şekilde tanımlanması yoluyla YZ modellerine güven artacaktır. [43]



**Şekil 2.17.** DeepTox ağının farklı katmanları tarafından öğrenilen toksikofor yapılar. İlk iki satırda belirtilen yapılar ağın ilk gizli katmanında öğrenilmiştir, 3. Satırda öğrenilen yapılar ise ağın daha derin bir katmanından öğrenilmiştir.

#### 2.5.4. İlaç Yeniden Konumlandırma Çalışmalarında Makine Öğrenimi Uygulamaları

İlaç geliştirme maliyetlerinin çok maliyetli hale gelmesi nedeniyle araştırmacılar tarafından önceden geliştirilen ilaçların farklı bir endikasyonda için de araştırma fikri vermiştir. Böylece belirli bir endikasyonda kullanılan ilaç, farklı bir endikasyonda tekrar konumlandırarak ilaç geliştirme maliyetinin ve süresinin azaltılabilecektir. [120]

İlaç yeniden konumlandırma çalışmalarında da kimyasal benzerlik ve moleküler docking gibi hesaplamalı yöntemler kullanılmaktadır. Kimyasal ve genomik veriler hızla biriktikçe ve farmakolojik ve fenotipik bilgiler giderek daha erişilebilir ve standart hale geldikçe, ilaç yeniden konumlandırma çalışmaları, ilaç endüstrisinde yapay zeka teknolojilerinin kullanım alanlarından birine dönüşmektedir. Normalde sorun oluşturabilecek unsurlar, (gürültülü, eksik ve yüksek boyutlu veriler) YZ'nın güçlü olduğu alanlardır. Bu alanda yapılan en iyi çalışmalardan biri DTINet'tir. Luo ve ark., FDA'nın ruhsatladığı 732 ilaç molekülünden, topoisomera inhibitörü olan topectan'ın, insan retinoik asit reseptör gamma içinde terapötik etkileri olabileceğini göstermişlerdir. [121]

### 2.5.5. Organik Sentez Çalışmalarında Makine Öğrenimi Uygulamaları

Bilinen organik bileşiklerin büyük çoğunluğu, sınırlı sayıdaki reaksiyonlarıyla sentezlenebilir. Güvenilir ve tamamen otomatize sentez planı oluşturmak henüz mümkün değildir. Sebeplerden biri, verimli ileri ve retrosentetik sentez planlama için gerekli olan kapsamlı kimya uzmanlığı gerektirmesidir. Artan hesaplamalı yöntemler, büyük veri ve yeni derin öğrenme algoritmalarının ve optimizasyonların geliştirilmesi YZ'nın organik sentezin planlanmasında kullanılmasıyla sonuçlanmıştır. Retrosentetik yaklaşımda sentezlenmek istenen molekülün tekrarlanabilir, verimli bir planını oluşturmak hedeflenir. Bu yöntemin ana zorluğu ise geniş kimyasal dönüşümlerin/reaksiyonlara bağlılığıdır. Bunlar genellikle manuel yapım ve iyileştirme gerektirir. Son yıllarda doğal dil işleme metodlarının gelişmiştir ve bu yöntemler organik sentez planlanmasına da yansımıştır. Bu araştırmalarda, organik moleküllerdeki parçaların, bir dildeki (Türkçe, İngilizce gibi) cümlelerde yer alan kelimelerinkine benzer olduğu düşüncesiyle hareket edilir. Bu yol için genellikle metin bazlı bir molekül temsil yolu olan SMILES kullanılır. [122]

### 2.6. Antimikrobiyal İlaç Geliştirmede Makine Öğrenimi

Günümüzde tedavilerde kullanılan antibiyotiklere karşı bakterilerin direnç geliştirmeleri nedeniyle yeni antibiyotiklerin keşfi oldukça önemlidir. Literatürde yeni antibiyotik keşfi için makine öğrenmesinin kullanıldığı çalışmalar mevcuttur. Örneğin; 2020 yılında yapılan bir çalışmada, derin sinir ağları eğitilerek bir moleküllerin antibiyotik etkilerini tahmin eden bir model oluşturulmuştur. Ardından, ilaç yeniden konumlandırma yaklaşımı ile halisin molekülünün *Mycobacterium tuberculosis* ve karbapenama dirençli *Enterobacteriaceae* gibi çeşitli patojenlerin üzerinde bakterisidal etkisinin olduğu belirlenmiştir. [123] Sean Ekins ve ark. ise makine öğrenmesi kullanarak *Mycobacterium tuberculosis* için çeşitli öncü bileşikler tespit etmişlerdir. [124–127]

Bu tez çalışmasında, *S. aureus* ve *C. albicans* inhibisyonu için makine öğrenmesi modelleri geliştirilmiştir. Bu modeller ile dışarıdan sunulan bir molekülün *S. aureus* ve *C. albicans* inhibe edip edemeyeceğini (aktif veya inaktif olarak sınıflandırarak) yüksek bir doğrulukla tahmine edebilmektedir. Böylece molekül tasarımını daha efektif hale getirilerek molekülün sentezlenmeden aktivitesinin hızlı ve doğru tespit edilebilmesi sayesinde zaman ve ekonomik faydalar sağlanması hedeflenmiştir. Bu çalışmada çeşitli pirol-kalkon türevi bileşikler tasarlanarak, moleküllerin *S. aureus* ve *C. albicans* aktiviteleri makine öğrenmesi ile tahmin edilmiştir. Bu bileşiklerin antifungal aktiviteye olabileceği makine öğrenmesi ile belirlendikten sonra sentez çalışması yapılmıştır. Sentezlenen moleküllerin *in vitro* *S. aureus* ve *C. albicans* dahil olmak üzere çeşitli antimikrobiyal aktiviteleri incelenmiştir.

### 3. MATERYAL ve YÖNTEM

#### 3.1. Makine Öğrenmesi Modelinin Oluşturulması

Makine öğrenmesi modelinin oluşturulması için yazılım dili olarak kullanımının kolay olması ve birçok kütüphaneye sahip olması nedeniyle “Python” tercih edilmiştir. Kullanılan kütüphaneler ise veri setinin işlenmesi ve düzenlenmesi işlemleri için pandas [128] ve numpy [129], moleküllerin görüntülenmesi, düzenlenmesi gibi işlemler için RDKit[130], derin öğrenme algoritması için ise keras ve tensorflow kullanılmıştır. Python kodu yazılmasını kolaylaştıran, internet tabanlı ve GPU desteği sağlayan Google Colaboratory ile model oluşturulmuştur.

##### 3.1.1. Veri Setinin Hazırlanması

*S. aureus* (ChEMBL 352) ve *C. albicans* (ChEMBL612868) için Minimum İnhibitör Konsantrasyonuna (MİK) sahip moleküller ChEMBL veri tabanından (versiyon 30, Şubat 2022) toplanmıştır. Veri setinin ön incelemesi yapılarak, aktivite değerlerin dağılımı ve standart üniteleri incelenmiştir. Sadece aktivite değerleri “µg/mL” cinsinden verilen moleküller veri setinde tutulmuştur, diğer moleküller silinmiştir. Molekülleri aktif veya inaktif olarak etiketlemek için 31.25 µg/mL eşik değeri olarak belirlenmiştir. Moleküller, MİK değerleri eşik değerinden daha düşükse aktif, yüksekse inaktif olarak etiketlendi. *S. aureus* veri setinde toplam 42216 molekül, *C. albicans* veri setinde ise toplamda 15310 molekül yer almaktadır. Moleküller tanımlayıcı olarak ise ECFP4 belirlenmiştir. Veri setlerinin bölünmesi için Scikit-learn kütüphanesinin “train\_test\_split” fonksiyonu kullanılmıştır. Veri setlerinin %90’ı eğitim seti, %10’u ise test seti olarak bölünmüştür.

##### 3.1.2. Makine Öğrenmesi Modelinin Eğitilmesi

Moleküllerin aktivitesinin tahmini için derin sinir ağları algoritması kullanılarak model eğitilmiştir. Oluşturulan sinir ağları; giriş katmanı, 2 tane saklı katman (256,256), dropout katmanı ve çıktı katmanından oluşmaktadır. Saklı katmanlar için aktivasyon fonksiyonu olarak ReLU, çıktı katmanı içinse sigmoid fonksiyonu kullanılmıştır. Dropout katmanların oranı ise 0.3 olarak belirlenmiştir. Optimizasyon fonksiyonu olarak ADAM kullanılmıştır. Modeller, Keras’ın erken durdurma fonksiyonları kullanılarak eğitilmiştir. Modellerin, mimarileri oluşturulduktan sonra eğitim setleri ile eğitimi gerçekleştirilmiştir. Modelleri değerlendirmek için test seti kullanılmıştır ve modeller doğruluk skoru, F1 skoru, Matthew korelasyon katsayısı (MKK) ve AUC olmak üzere 4 farklı metrik ile değerlendirilmiştir.

### 3.2. Sentez Çalışmaları

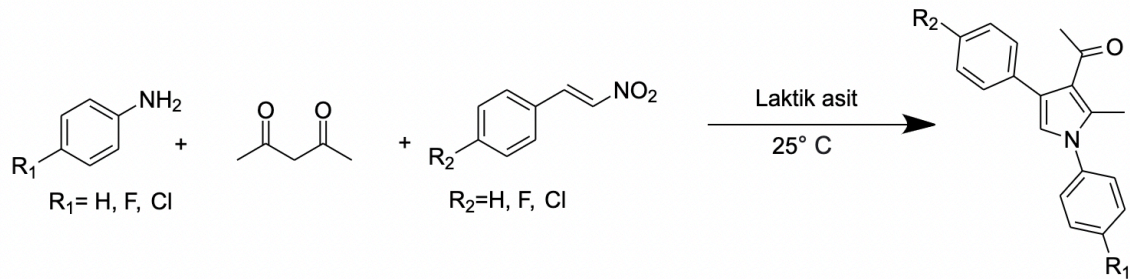
#### 3.2.1. Materyal

Bu yüksek lisans tezi kapsamında kullanılan tüm çözücü, reaktant ve katalizörler analitik saflıktadır. Sentez başlangıç maddeleri olarak; anilin, asetilaseton (Fluka),  $\beta$ -trans nitro stiren (BLDpharma), laktik asit (Abcr), benzaldehit (BLDpharma), 4-florobenzaldehit (BLDpharma), 4-metoksibenzaldehit (BLDpharma) kullanılmıştır. Çözücü, katalizör ve yardımcı reaktant olarak; metanol, etanol, kloroform, diklorometan (DCM), potasyum hidroksit (KOH), n-hekzan, etilasetat, petrol eteri, aseton, dietileter, sodyum sülfat ( $\text{NaSO}_4$ ) kullanılmıştır.

#### 3.2.2. Yöntem

##### 3.2.2.1. Genel Sentez Yöntemi-I: Üç Bileşenli Tek Kap Yöntemi ile 1,2,3,4-Tetrasüstitüe Pirol Bileşiklerinin Sentezi

Primer amin (5 mmol), asetilaseton (5 mmol) ve  $\beta$ -trans nitrostiren türevleri (5 mmol) 5 mL laktik asit ortamında oda sıcaklığında karıştırıldı. Tepkimenin tamamlanıp tamamlanmadığı İTK (80:20/petrol eteri:etilasetat) ile takip edildi. Reaksiyon sonunda oluşan katı, reaksiyon karışımından süzülerek ayrıldıktan sonra, uygun çözücü sisteminde (diklorometan/hekzan) kristallendirilerek ya da kolon kromatografisi ile saflaştırıldı. (Şekil 3.1.)

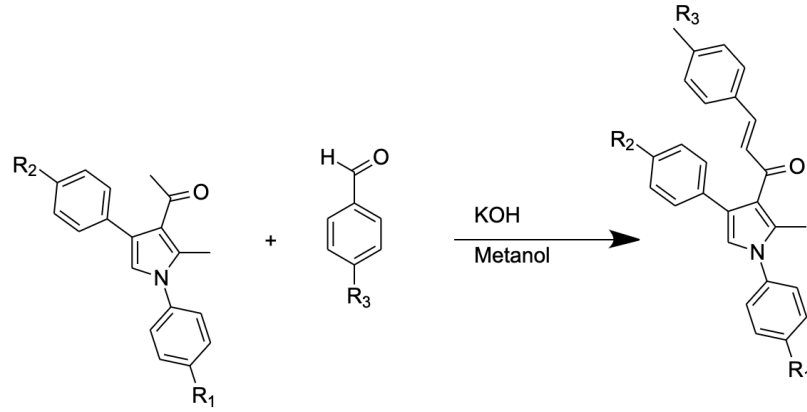


Şekil 3.1. 1,2,3,4-Tetrasüstitüe pirol bileşiklerin sentez şeması

##### 3.2.2.2. Genel Sentez Yöntemi-II: 1,2,3,4-Tetrasüstitüe Pirol-Kalkon Hibrit Bileşiklerinin Sentezi

1,2,3,4-tetrasüstitüe pirol-kalkon hibrit bileşikleri, Genel Sentez Yöntemi II'ye göre elde edilmiştir. Bu sentez yönteminde Batt ve çalışma grubunun 1993 yılında yaptıkları çalışmada kullandıkları metot modifiye edilerek kullanılmıştır. Benzaldehit türevleri ile 1,2,3,4-terasüstitüe pirol bileşikleri metanol içerisinde çözüldü ve eşdeğer miktarda %50 KOH çözeltisi ilave edilerek oda sıcaklığında karıştırıldı. Tepkimenin tamamlanıp tamamlanmadığı İTK (90:10/petrol eteri: etil asetat) ile takip edildi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından

etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı. (Şekil 3.2.)[131]



**Şekil 3.2.** 1,2,3,4-Tetrasübstitüe Pirol-Kalkon Hibrit Bileşiklerinin Sentez Şeması

### 3.3. Analitik Çalışmalar

#### 3.3.1. Erime Noktası Tayini

Sentezleri yapılan bileşiklerin erime dereceleri, Mettler Toledo MP90 erime derecesi tayin cihazı ile saptandı.

#### 3.3.2. İnce Tabaka Kromatografisi ile Yapılan Kontroller

İnce Tabaka Kromatografisi çalışmalarında Kieselgel 60 F254 ile 2 mm kalınlığında kaplanmış hazır alüminyum plaklar (Merck) kullanıldı. Sentezleri yapılan bileşiklerin kromatografik kontrollerinde aşağıdaki çözücü sistemleri kullanıldı.

- i) Petrol eteri-etil asetat (90:10)
- ii) Petrol eteri-etil asetat (80:20)

Kromotogramlarda sentez ürünleri ve başlangıç maddelerine ait lekeleri belirlemek için UV ışığı kullanıldı.

#### 3.3.3. Spektroskopik Kontroller

##### 3.3.3.1. FT-IR Spektrumları

Bileşiklerin FT-IR spektrumları, Perkin Elmer 400 (USA) FTIR/FIR Spectrometer Frontier Spektrofotometresi (32 scan, 4000-650 cm<sup>-1</sup> dalga sayısı aralığı) kullanılarak analiz edildi ve dalga sayısı (cm<sup>-1</sup>) cinsinden değerlendirildi. Spekturumlar, *Spectrum v2.0 Software* yazılımı kullanılarak yorumlandı.



### 3.3.3.2. <sup>1</sup>H-NMR Spektrumları

Bileşiklerin <sup>1</sup>H-NMR spektrumları, CDCl<sub>3</sub> (Merck) veya (CD<sub>3</sub>)<sub>2</sub>SO (Merck) içindeki çözeltileri ile Bruker 400 NMR spektrofotometrisi kullanılarak kaydedildi. Kimyasal kayma değerleri, TMS (tetrametilsilan) iç standardına göre ppm (milyonda bir) skalasında değerlendirildi. Elde edilen nmr piklerindeki etkileşimler, s (singlet), d (dublet), dd (dubletin dubleti), ddd (dubletin dubletin dubleti), t (triplet), m (multiplet) olarak Hertz (Hz) cinsinden yorumlandı. <sup>1</sup>H-NMR Spektrumları, *Mestronova* yazılımı kullanılarak kaydedildi.

### 3.3.3.3. <sup>13</sup>C-NMR Spektrumları

Bileşiklerin <sup>13</sup>C-NMR spektrumları (CD<sub>3</sub>)<sub>2</sub>SO (Merck) içindeki çözeltilerinde ile Bruker 100 NMR spektrofotometrisinde alınıp kimyasal kayma değerleri  $\delta$  skalasında değerlendirildi. <sup>13</sup>C - NMR Spektrumları, MestReC programı kullanılarak verildi.

## 3.4 Antimikrobiyal Aktivite Çalışmaları

### 3.4.1. Materyal

Antibakteriyel ve antifungal etki çalışmaları Mersin Üniversitesi Eczacılık Fakültesi Farmasötik Mikrobiyoloji Anabilim Dalı koleksiyonundan temin edilen aşağıdaki bakteri ve mantar suşları ile yapıldı:

*Escherichia Coli* (ATCC 25922),

*Enterococcus faecalis* (ATCC 29112),

*Staphylococcus aureus* (ATCC 29213),

*Streptococcus pneumoniae* (ATCC 49619),

*Klebsiella pneumoniae* (ATCC 700603),

*Candida albicans* (ATCC 10231),

*Candida metapsilosis* (CBS 2916),

*Candida parapsilosis* (ATCC 22019).

MİK değerinin saptanmasında bakteriler için Müller Hinton Broth, mayalar için Tryptic Soy Broth besiyerleri kullanıldı. *In vitro* antibakteriyel etki denemelerinde ampisilin, antifungal etki denemelerinde ise flukonazol kullanıldı.

### 3.4.2. Kullanılan Araç ve Gereçler

– Etüv (Mettler)

- Pastör Fırını(Memmert)
- Otoklav (Nüve- OT 020)
- Hassas Terazı (Scaltec)
- Distile su cihazı (Nüve NS 108)
- Öze
- Petri kapları
- Pipet ucu
- Balon joje
- Ependorf tüpler
- Deney tüpleri

### **3.4.3. Bakteriyolojik Tanı İçin Kullanılan Besiyerleri**

#### **3.4.3.1. Mueller Hinton Broth (Fluka 70192)**

*In vitro* (canlı hücre dışında) yapılan standart mikrobiyolojik analizlerde klinik olarak önemli patojenlerin, antibiyotik ve MIC değerlerini belirlemek için kullanılan sıvı besiyeridir.

#### **Bileşim;**

- Katı sığır infüzyonu 2,0 g/L
- Nişasta 1,5 g/L
- Kazein hidrolizatı 17,5 g/L
- pH 7,4 +/- 0.2 37 °C

#### **Etki şekli;**

Çok amaçlı kullanım alanı olan, inhibitör ve indikatör içermeyen genel sıvı besiyeridir.

#### **Hazırlanması;**

Dehidre besiyeri 23,0 g/L olacak şekilde amaca uygun kaplara (balon, erlen vb.) eklenip, 1 litre damıtık su içinde çözdürülür. Otoklavda 121°C'da 15 dakika sterilize edilir. Hazırlanmış besiyeri berrak, sarımsı renktedir ve 37°C'de pH'sı 7,4±0,2'dir. Soğuduktan sonra buzdolabında + 4 °C'de muhafaza edilir.

### Ürün Bilgileri;

Ürün Adı	Kat. No	Ambalaj
Mueller-Hinton Broth	70192.0500	500 g

### 3.4.3.2. Tryptic Soy Broth (TSB) (Merck 1.05459)

*In vitro* (canlı hücre dışında) yapılan standart mikrobiyolojik analizlerde genel sıvı besiyeri olarak kullanılır. Çok amaçlı kullanım alanı olan, inhibitör ve indikatör içermeyen genel sıvı besiyeridir.

#### Bileşim

- Kazein peptonu 17,0 g/L;
- Soya peptonu 3,0 g/L;
- D(+) Glukoz 2,5 g/L;
- NaCl 5,0 g/L;
- K<sub>2</sub>HPO<sub>4</sub> 2,5 g/L.

#### Etki şekli;

Çok amaçlı kullanım alanı olan, inhibitör ve indikatör içermeyen genel sıvı besiyeridir.

#### Hazırlanması;

Dehidre besiyeri, 30,0 g/L olacak şekilde amaca uygun kaplara (balon, erlen vb.) eklenip, 1 litre damıtık su içinde çözülür. Otoklavda 121°C'da 15 dakika sterilize edilir. Hazırlanmış besiyeri berrak sarımsı renkte olup, 25°C'da pH'sı 7,3±0,2'dir. Soğuduktan sonra buzdolabında +4 °C'de muhafaza edilir.

### Ürün Bilgileri;

Ürün Adı	Kat. No	Ambalaj
Tryptic Soy Broth (TSB)	1.05459.0500	500 g

### 3.4.4. Antibakteriyel ve Antifungal Aktivite Çalışmaları

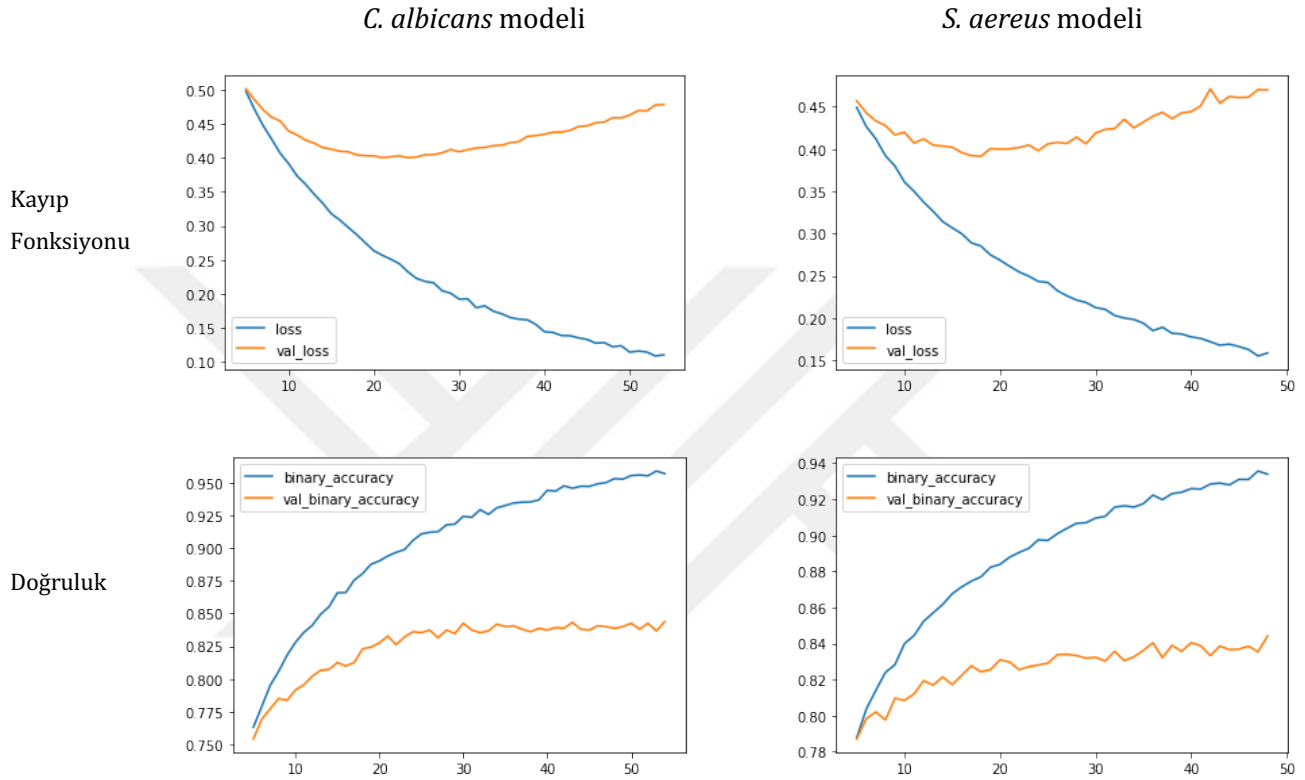
Bu çalışmada sentezlenen toplam 20 bileşiğin *in vitro* antibakteriyel ve antifungal etkilerinin belirlenmesi amacıyla Mikrodilüsyon Broth Yöntemi kullanılmıştır. Bu yöntemde antibakteriyel etki çalışmaları için ampisilin, antifungal etki çalışmaları için ise flukanazol

standart maddeler seçilmiştir. Sentezlenen pirol-kalkon türevi bileşiklerinden 6 µg alınıp, stok solüsyonları %98'lik dimetilsülfoksit (DMSO (Merck 1.02952)) ile çözündürülerek bir kısmı su ile bir kısmı da DMSO (1ml/1ml) ile 6mL'ye tamamlanmıştır. Antibakteriyel ve antifungal aktiviteleri incelenecek olan bileşiklerin DMSO'de çözündürüldükten sonra her bakteri ve mayalar için hazırlanmış serinin ilk tüpünde 500 µg/mL olacak şekilde tüplere ilave edildi. Sulandırma yöntemi ile 10 dilüsyon hazırlandı (500, 250, 125, 62.5, 31.25, 15.62, 7.81, 3.90 µg/mL). Bu tüplere bakterilerin ve mayaların 18–24 saatlik buyyon kültürlerinin yoğunluğu Mc Farland 0.5'e göre ayarlandıktan sonra 10 µL inoküle edildi. Bunlara ek olarak sadece bileşikleri içeren mikroorganizma içermeyen kontrol tüpleri hazırlandı. Bütün deney tüpleri 37 °C'de 24 saat inkübasyona bırakıldı. Bunun sonucunda bakteri üremesi sonucu oluşan bulanıklık dikkate alınarak hiç bulanıklık göstermeyen üreme olmayan tüpteki en düşük dilüsyon µg/mL cinsinden Minimum İnhibisyon Konsantrasyonu (MİK) olarak saptandı. [132,133]

## 4. BULGULAR

### 4.1. Makine Öğrenmesi Modelinin Değerlendirilmesi

Bu tez çalışması için *C. albicans* ve *S. aureus* için 2 tane makine öğrenmesi modeli geliştirilmiştir. Bu modellerin eğitimine ilişkin kayıp fonksiyonu ve doğruluk grafikleri verilmiştir. (Şekil 4.1.) Modellerin eğitilmesinin ardından, doğrulama seti kullanılarak 4 farklı metrik ile modeller değerlendirilmiştir. (Tablo 4.1.)



Şekil 4.1. Modellerin eğitim sürecinde kayıp fonksiyonunun ve doğruluğunun değişimi

Tablo 4.1. Makine öğrenmesi modellerinin değerlendirilme sonuçları

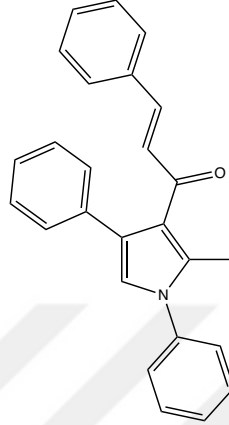
Modeller	Doğruluk	F1 skoru	MKK	AUC
<i>C. albicans</i>	0.84	0.81	0.68	0.83
<i>S. aureus</i>	0.84	0.87	0.67	0.83

Modellerin değerlendirilmesinin ardından, bu tez için tasarlanan moleküllerin antimikrobiyal etkileri tahmin edilmiştir. Tasarlanan bileşikler arasından, MAY-3, MAY-5, MAY-6, MAY-18, MAY-17, MAY-21, MAY-24 *C. albicans* türüne karşı aktif tahmin edilmiştir ancak *S. aureus*'a karşı aktif tahmin edilen bir molekül olmamıştır. Modellerin aktif ve inaktif hassasiyetini test etmek için moleküller *S. aureus* dahil olmak üzere çeşitli bakteri suşlarında da biyolojik aktivite sonuçları değerlendirilmiştir.

## 4.2. Kimyasal Bulgular

Tez çalışmasındaki bileşikler genel sentez yöntemleri uygulanarak elde edilmiştir.

### 4.2.1. 1-(2-Metil-1,4-difenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-1)

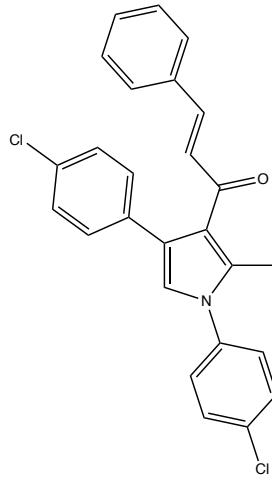


**Şekil 4.2.** 1-(2-Metil-1,4-difenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(2-Metil-1,4-difenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 29 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%68
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO <sub>4</sub> -d <sub>6</sub> ) ppm	:	δ=7.63-7.24 (m, 15H, Ar-H), 7.48 (d, J=15.8 Hz, 1H,-CH), 7.12 (s, 1H,Ar-H), 6.74 (d, J=15.9 Hz, 1H,-CH), 2.40 (s, 3H, -CH <sub>3</sub> ).
<sup>13</sup> C NMR spektrumu 100 MHz, DMSO <sub>4</sub> -d <sub>6</sub> ) ppm	:	δ= 187.55, 139.86, 138.28, 135.37, 134.80, 134.77, 129.90, 129.49, 128.79, 128.33, 128.09, 127.77, 127.66, 126.42, 126.01, 125.12, 121.96, 120.75, 12.53.
Elementel Analiz (C <sub>26</sub> H <sub>21</sub> N)	:	Hesaplanan; C, 85.93; H, 5.82; N, 3.85; Bulunan: C, 86.08; H, 5.78; N, 3.82.

#### 4.2.2. 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-2)

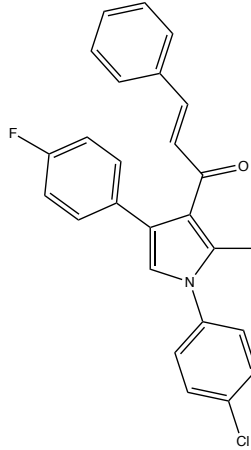


**Şekil 4.3.** 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntemi 1 ile elde edilen 1-(1,4-bis(4-klorofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 27 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%38
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.38 (3H, s, -CH <sub>3</sub> ), 6.76 (1H, d, J = 15.87 Hz, -CH), 7.22-7.67 (14H, m, Ar-H), 7.49 (1H, d, J = 15.87 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.45, 121.06, 122.02, 123.90, 127.61, 127.79, 127.89, 128.25, 128.85, 129.46, 130.10, 130.28, 131.18, 132.68, 134.04, 134.68, 137.05, 140.70, 187.82.
Elementel Analiz (C <sub>26</sub> H <sub>19</sub> Cl <sub>2</sub> NO)	:	Hesaplanan; C, 72.23; H, 4.43; Cl, 16.40; N, 3.24; O, 3.70 Bulunan; C, 72.23; H, 4.43; Cl, 16.38; N, 3.24; O, 3.68

#### 4.2.3. 1-(1-(4-Klorofenil)-4-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-3)



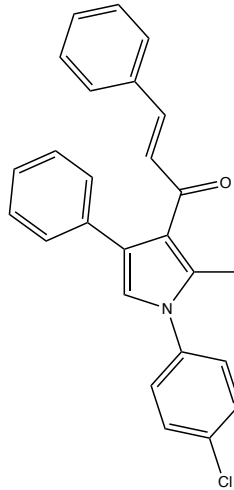
**Şekil 4.4.** 1-(1-(4-Klorofenil)-4-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1-(4-Klorofenil)-4-(4-fluorofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), yeterli miktarda metanol eklenerek çözüldü. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 25 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%69
Renk	:	Sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etil asetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.39 (3H, s, -CH <sub>3</sub> ), 6.75 (1H, d, J = 15.86 Hz, -CH), 7.16-7.67 (14H, m), 7.48 (1H, d, J = 15.86 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.45, 114.98, 115.20, 120.85, 122.12, 124.13, 127.53, 127.78, 127.81, 128.84, 129.45, 130.05, 130.48, 130.55, 131.54, 131.58, 132.64, 134.61, 134.71, 137.09, 140.46, 159.96, 162.37, 187.70.
Elementel Analiz (C <sub>26</sub> H <sub>19</sub> ClFNO)	:	Hesaplanan: C, 75.09; H, 4.61; Cl, 8.52; F, 4.57; N, 3.37; O, 3.85; Bulunan: C, 75.10; H, 4.60; Cl, 8.52; F, 4.57; N, 3.35; O, 3.84



#### 4.2.4. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-4)

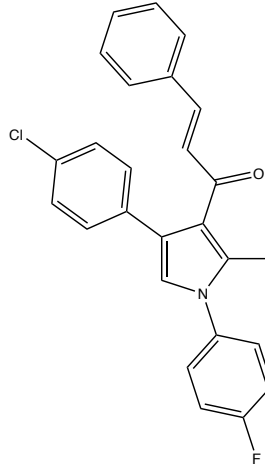


**Şekil 4.5.** 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 32 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%54
Renk	:	Sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.39 (3H, s, -CH <sub>3</sub> ), 6.73 (1H, d, J = 15.84 Hz, -CH), 7.16-7.67 (15H, m, Ar-H) 7.48 (1H, d, J = 15.85 Hz, -CH),).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.45, 120.70, 122.15, 125.28, 126.48, 127.60, 127.79, 128.34, 128.75, 128.79, 129.44, 129.94, 132.62, 134.71, 134.76, 135.21, 137.13, 140.08, 187.63.
Elementel Analiz (C <sub>26</sub> H <sub>20</sub> ClNO)	:	Hesaplanan; C, 78.48; H, 5.07; Cl, 8.91; N, 3.52; O, 4.02; Bulunan; C, 78.46; H, 5.05; Cl, 8.91; N, 3.50; O, 4.02.

#### 4.2.5. 1-(4-(4-Klorofenil)-1-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-5)

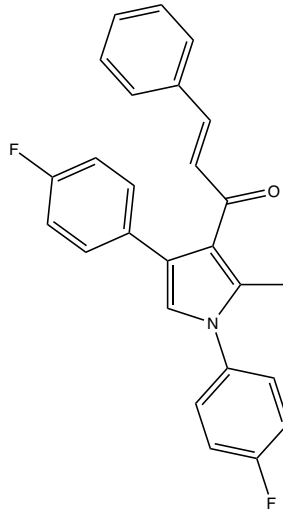


**Şekil 4.6.** 1-(4-(4-Klorofenil)-1-(4-fluorofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(4-(4-Klorofenil)-1-(4-fluorofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 32 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%52
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
R <sub>f</sub>	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.35 (3H, s, -CH <sub>3</sub> ), 6.75 (1H, d, J = 15.86 Hz, -CH), 7.18-7.63 (14H, m, Ar-H) 7.48 (d, J = 15.86 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.37, 116.19, 116.41, 121.21, 121.73, 123.70, 127.59, 127.86, 128.20, 128.24, 128.29, 128.85, 130.10, 130.31, 131.14, 134.11, 134.52, 134.67, 134.94, 140.58, 160.21, 162.65, 187.79.
Elementel Analiz (C <sub>26</sub> H <sub>19</sub> ClFNO)	:	Hesaplanan; C, 75.09; H, 4.61; Cl, 8.52; F, 4.57; N, 3.37; O, 3.85; Bulunan; C, 75.09; H, 4.60; Cl, 8.50; F, 4.57; N, 3.35; O, 3.84.

#### 4.2.6. 1-(1,4-bis(4-florofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-6)

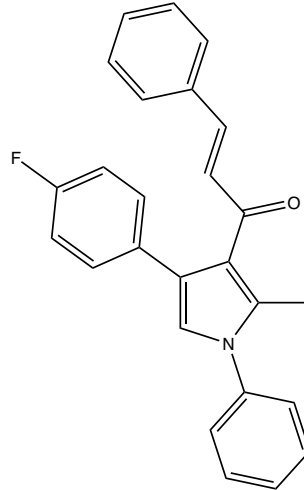


**Şekil 4.7.** 1-(1,4-bis(4-florofenil)-2-metil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntemi 1 ile elde edilen 1-(1,4-Bis(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 25 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%51
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
R <sub>f</sub>	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.37 (3H, s, -CH <sub>3</sub> ), 6.75 (1H, d, J = 15.85 Hz, -CH), 7.15-7.64 (14H,m), 7.48 (1H, d, J = 15.92 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ= 12.39, 114.97, 115.18, 116.17, 116.40, 121.04, 121.84, 123.91, 127.54, 127.79, 128.22, 128.30, 128.84, 130.03, 130.49, 130.57, 131.67, 134.59, 134.73, 134.84, 140.32, 159.94, 160.20, 162.36, 187.63.
Elementel Analiz (C <sub>26</sub> H <sub>19</sub> F <sub>2</sub> NO)	:	Hesaplanan; C, 78.18; H, 4.79; F, 9.51; N, 3.51; O, 4.01; Bulunan; C, 78.16; H, 4.75; F, 9.51; N, 3.50; O, 4.00

#### 4.2.7. 1-(4-(4-Florofenil)-2-metil-1-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on (MAY-8)

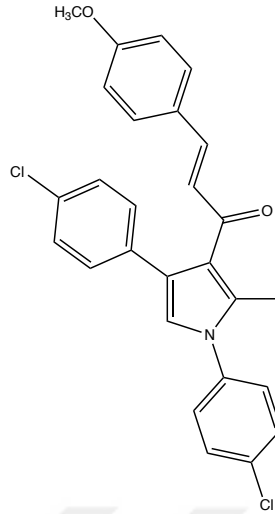


**Şekil 4.8.** 1-(4-(4-Florofenil)-2-metil-1-fenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(4-(4-Florofenil)-2-metil-1-fenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan benzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 24 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%38
Renk	:	Sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ =2.39 (3H, s, -CH <sub>3</sub> ), 6.76 (1H, d, J = 15.9 Hz, -CH), 7.16-7.63 (15H, m, Ar-H), 7.48 (1H, d, J = 16.08 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ= 12.54, 114.96, 115.17, 120.89, 121.94, 123.97, 126.00, 127.58, 127.80, 128.12, 128.84, 129.50, 130.02, 130.52, 130.60, 131.73, 134.67, 134.75, 138.24, 140.23, 187.61.
Elementel Analiz (C <sub>26</sub> H <sub>20</sub> FNO)	:	Hesaplanan; C, 81.87; H, 5.29; F, 4.98; N, 3.67; O, 4.19; Bulunan; C, 81.85; H, 5.27; F, 4.97; N, 3.68; O, 4.18.

#### 4.2.8. 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-9)

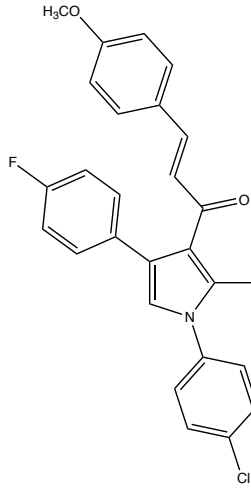


**Şekil 4.9.** 1-(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen (1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-metoksibenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 27 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%24
Renk	:	Beyaz sarı katı
Çözücü	:	Etil asetat
R <sub>f</sub>	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.36 (3H, s, -CH <sub>3</sub> ), 3.80 (3H, s, -OCH <sub>3</sub> ), 6.65(1H, d, J = 15.82 Hz, -CH), 6.92-7.67 (13H, m, Ar-H), 7.46 (1H, d, J = 16.63 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ= 12.38, 55.32, 114.39, 122.16, 123.78, 125.42, 127.22, 127.76, 128.21, 129.45, 129.69, 130.13, 131.04, 134.06, 134.14, 137.13, 140.88, 160.98, 188.09.
Elementel Analiz (C <sub>27</sub> H <sub>21</sub> Cl <sub>2</sub> NO <sub>2</sub> )	:	Hesaplanan; C, 70.14; H, 4.58; Cl, 15.33; N, 3.03; O, 6.92; Bulunan: C, 70.10; H, 4.54; Cl, 15.30; N, 3.07; O, 6.96.

**4.2.9. 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil) prop-2-en-1-on (MAY-10)**

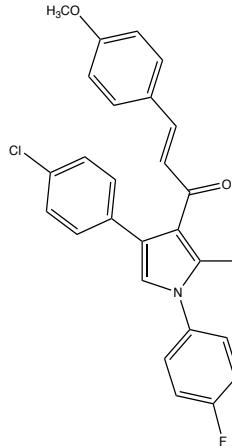


**Şekil 4.10.** 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil) prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-metoksibenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 29 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%41
Renk	:	Sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.37 (3H, s, -CH <sub>3</sub> ), 3.79 (3H, s, -OCH <sub>3</sub> ), 6.63 (1H, d, J = 15.8 Hz, -CH), 6.93-7.66 (13H, m, Ar-H), 7.45 (1H, d, J = 15.8 Hz, -CH)
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.39, 55.30, 114.39, 114.94, 115.15, 120.70, 122.26, 124.04, 125.33, 127.25, 127.74, 129.43, 129.59, 130.34, 130.42, 131.31, 131.57, 131.60, 132.57, 134.12, 137.16, 140.59, 159.89, 160.94, 162.31, 187.92.
Elementel Analiz (C <sub>27</sub> H <sub>21</sub> ClFNO <sub>2</sub> )	:	Hesaplanan; C, 72.73; H, 4.75; Cl, 7.95; F, 4.26; N, 3.14; O, 7.18; Bulunan; C, 72.75; H, 4.74; Cl, 7.96; F, 4.26; N, 3.12; O, 7.18.

**4.2.10. 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil) prop-2-en-1-on (MAY-11)**

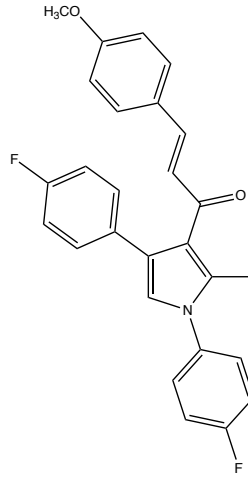


**Şekil 4.11.** 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil) prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-metoksibenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 32 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%15
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.34 (3H, s- CH <sub>3</sub> ), 3.80 (3H, s, -OCH <sub>3</sub> ), 6.65 (1H, d, J = 15.81 Hz), 6.92-7.63 (14H,m, Ar-H, -CH)
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ= 12.31, 55.31, 114.39, 116.16, 116.39, 121.12, 121.88, 123.56, 125.43, 127.24, 128.19, 128.28, 129.66, 130.15, 131.00, 134.15, 134.38, 134.62, 140.72, 160.18, 160.95, 162.62, 188.01.
Elementel Analiz (C <sub>27</sub> H <sub>21</sub> ClFNO <sub>2</sub> )	:	Hesaplanan; C, 72.73; H, 4.75; Cl, 7.95; F, 4.26; N, 3.14; O, 7.18; Bulunan: C, 72.74; H, 4.74; Cl, 7.96; F, 4.27; N, 3.12; O, 7.18.

**4.2.11. 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on  
(MAY-13)**



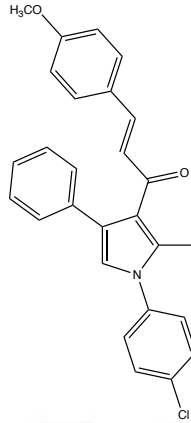
**Şekil 4.13.** 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1,4-Bis(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-metoksibenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 25 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%41
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.35 (3H, s, -CH <sub>3</sub> ), 3.79 (3H, s, -OCH <sub>3</sub> ), 6.63 (1H, d, J = 15.8 Hz, C-H), 7.13-7.61 (14H, m, Ar-H)
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.33, 55.30, 114.38, 114.92, 115.13, 116.15, 116.38, 120.89, 121.97, 123.81, 125.34, 127.26, 128.17, 128.26, 129.57, 130.36, 130.44, 131.66, 131.69, 134.37, 134.66, 134.69, 140.43, 159.87, 160.16, 160.91, 162.29, 162.60, 187.84.
Elementel Analiz (C <sub>27</sub> H <sub>21</sub> F <sub>2</sub> NO <sub>2</sub> )	:	Hesaplanan; C, 75.51; H, 4.93; F, 8.85; N, 3.26; O, 7.45; Bulunan; C, 75.49; H, 4.90; F, 8.82; N, 3.28; O, 7.48.



#### 4.2.12. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on (MAY-15)

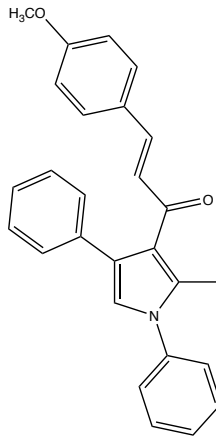


**Şekil 4.14.** 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-metoksifenil)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-metoksibenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 32 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%23
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
R <sub>f</sub>	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.38 (3H, s), 3.78 (3H, s), 6.61 (1H, d, J = 15.78 Hz), 7.22-7.66 (13H, m), 7.45 (1H, d, J = 15.79 Hz), 7.15 (1H, s, Ar-H)
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 187.83, 160.85, 140.16, 135.25, 134.24, 129.55, 129.43, 128.65, 128.29, 127.76, 127.30, 126.40, 125.42, 125.19, 122.29, 120.56, 114.34, 55.29, 12.40
Elementel Analiz (C <sub>27</sub> H <sub>22</sub> ClNO <sub>2</sub> )	:	Hesaplanan; C, 75.78; H, 5.18; Cl, 8.28; N, 3.27; O, 7.48; Bulunan; C, 75.76; H, 5.17; Cl, 8.27; N, 3.28; O, 7.49.

#### 4.2.13. 3-(4-metoksifenil)-1-(2-metil-1,4-difenil-1H-pirol-3-il)prop-2-en-1-on (MAY-16)

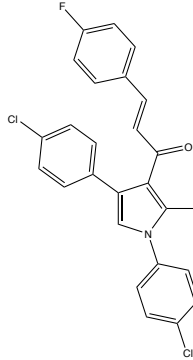


**Şekil 4.15.** 3-(4-metoksifenil)-1-(2-metil-1,4-difenil-1H-pirol-3-il)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(2-Metil-1,4-difenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-metoksibenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 29 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%27
Renk	:	Açık sarı katı
Çözücü	:	Etil asetat
Rf	:	0,81 Etilasetat/n-hekzan (1:1) ve UV aktif
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.38 (3H, s, -CH), 3.78 (3H, s, -OCH <sub>3</sub> ), 6.62 (1H, d, J = 15.78 Hz, -CH), 6.88-7.45 (15H, m, Ar-H), 7.44 (1H, d, J = 15.79 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ=12.48, 55.29, 114.33, 120.61, 122.09, 125.02, 125.49, 125.99, 126.34, 127.34, 128.02, 128.27, 128.70, 129.48, 129.52, 134.30, 139.93, 160.82, 187.76.
Elementel Analiz (C <sub>27</sub> H <sub>23</sub> NO <sub>2</sub> )	:	Hesaplanan; C, 82.42; H, 5.89; N, 3.56; O, 8.13; Bulunan: C, 82.40; H, 5.86; N, 3.57; O, 8.15.

**4.2.14. 1-(1,4-bis(4-Klorofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on  
(MAY-17)**

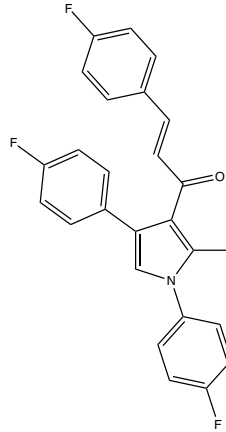


**Şekil 4.16.** 1-(1,4-bis(4-Klorofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen -(1,4-Bis(4-klorofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-florobenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 27 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%50
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.37 (3H, s, -CH <sub>3</sub> ), 6.73 (1H, d, J = 15.88 Hz, -CH), 7.16-7.80 (13H, m, Ar-H), 7.49 (1H, d, J = 15.62 Hz, -CH)
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.43, 115.76, 115.98, 121.07, 121.99, 123.87, 127.56, 127.77, 128.22, 129.46, 130.09, 130.21, 131.12, 131.30, 131.33, 132.68, 133.97, 134.55, 137.06, 139.63, 161.82, 164.29, 187.92.
Elementel Analiz (C <sub>26</sub> H <sub>18</sub> Cl <sub>2</sub> FNO)	:	Hesaplanan; C, 69.35; H, 4.03; Cl, 15.74; F, 4.22; N, 3.11; O, 3.55; Bulunan: C, 69.36; H, 4.01; Cl, 15.76; F, 4.22; N, 3.10; O, 3.54.

**4.2.15. 1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on  
(MAY-18)**

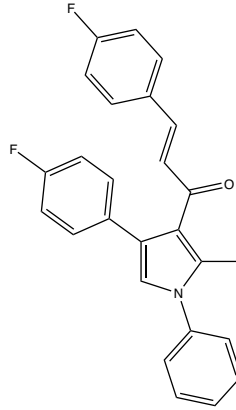


**Şekil 4.17.1-(1,4-bis(4-Florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı**

Genel sentez yöntem 1 ile elde edilen -(1,4-Bis(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-florobenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 25 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%53
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	$\delta$ = 2.36 (3H, s, -CH <sub>3</sub> ), 6.71 (1H, d, $J$ = 15.86 Hz, -CH), 7.14-7.63 (13H, m), 7.48 (1H, d, $J$ = 16.32 Hz, -CH)
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	$\delta$ = 12.37, 114.95, 115.16, 115.76, 115.98, 116.17, 116.40, 121.04, 121.80, 123.89, 127.50, 128.20, 128.29, 129.97, 130.06, 130.43, 130.51, 131.34, 131.37, 131.57, 131.60, 134.62, 134.75, 139.23, 159.90, 160.20, 161.78, 162.32, 164.25, 187.69.
Elementel Analiz (C <sub>26</sub> H <sub>18</sub> F <sub>3</sub> NO)	:	Hesaplanan; C, 74.81; H, 4.35; F, 13.65; N, 3.36; O, 3.83; Bulunan; C, 74.82; H, 4.36; F, 13.64; N, 3.37; O, 3.81.

**4.2.16. 3-(4-Florofenil)-1-(4-(4-florofenil)-2-metil-1-fenil-1H-pirol-3-il)prop-2-en-1-on  
(MAY-19)**

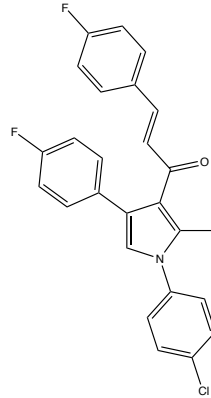


**Şekil 4.18.** 3-(4-Florofenil)-1-(4-(4-florofenil)-2-metil-1-fenil-1H-pirol-3-il)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(4-(4-Florofenil)-2-metil-1-fenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-florobenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 24 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%61
Renk	:	Turuncu katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.38 (3H, s, -CH <sub>3</sub> ), 6.71 (1H, d, J = 15.86 Hz, -CH), 7.10-7.80 (14H, m, Ar-H), 7.48 (1H, d, J = 15.89 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.51, 114.93, 115.14, 115.75, 115.97, 120.88, 121.90, 123.95, 125.98, 127.53, 128.10, 129.49, 129.97, 130.05, 130.45, 130.53, 131.36, 131.39, 131.63, 131.66, 134.57, 138.24, 139.13, 159.90, 161.77, 162.31, 164.24, 187.67.
Elementel Analiz (C <sub>26</sub> H <sub>19</sub> F <sub>2</sub> NO)	:	Hesaplanan; C, 78.18; H, 4.79; F, 9.51; N, 3.51; O, 4.01; Bulunan: C, 78.18; H, 4.80; F, 9.53; N, 3.50; O, 4.00.

**4.2.17. 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on (MAY-21)**

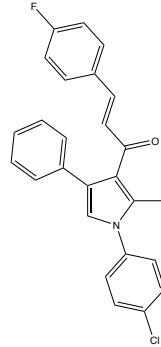


**Şekil 4.19.** 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1-(4-Klorofenil)-4-(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-florobenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 29 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%57
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.37 (3H, s, -CH <sub>3</sub> ), 6.71 (1H, d, J = 15.87 Hz, -CH), 7.10-7.80 (13H, m, Ar-H), 7.48 (1H, d, J = 15.87 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.44, 114.96, 115.18, 115.77, 115.98, 120.85, 122.08, 124.11, 127.47, 127.77, 129.45, 130.00, 130.09, 130.41, 130.49, 131.32, 131.51, 132.63, 134.51, 137.09, 139.38, 161.80, 164.27, 187.77..
Elementel Analiz (C <sub>26</sub> H <sub>18</sub> ClF <sub>2</sub> NO)	:	Hesaplanan; C, 71.97; H, 4.18; Cl, 8.17; F, 8.76; N, 3.23; O, 3.69; Bulunan: C, 71.98; H, 4.19; Cl, 8.16; F, 8.75; N, 3.22; O, 3.68.

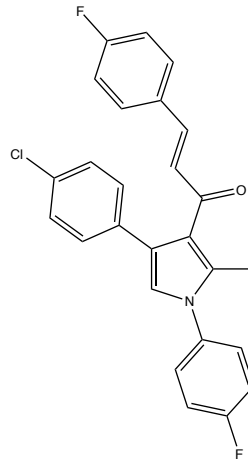
**4.2.18. 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on  
(MAY-23)**



**Şekil 4.21.** 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)-3-(4-florofenil)prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(1-(4-Klorofenil)-2-metil-4-fenil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-florobenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 32 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%63
Renk	:	Açık sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 2.53 (3H, s, -CH <sub>3</sub> ), 6.67 (1H, d, J = 15.85 Hz, -CH), 7.28-7.67 (14H, m, Ar-H), 7.47 (1H, d, J = 15.86 Hz, -CH).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 12.44, 115.71, 115.93, 120.70, 122.11, 125.26, 126.48, 127.55, 127.78, 128.32, 128.72, 129.45, 129.95, 130.03, 131.40, 132.61, 134.64, 135.17, 137.13, 138.94, 161.73, 164.20, 187.64.
Elementel Analiz (C <sub>26</sub> H <sub>19</sub> ClFNO)	:	Hesaplanan; C, 75.09; H, 4.61; Cl, 8.52; F, 4.57; N, 3.37; O, 3.85; Bulunan; C, 75.07; H, 4.60; Cl, 8.54; F, 4.56; N, 3.35; O, 3.87.

**4.2.19. 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (MAY-24)**

**Şekil 4.22.** 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)-3-(4-florofenil) prop-2-en-1-on bileşiğinin yapısı

Genel sentez yöntem 1 ile elde edilen 1-(4-(4-Klorofenil)-1-(4-florofenil)-2-metil-1H-pirol-3-il)etan-1-on (0.5 mmol), metanolde çözülmüştür. Ardından buz banyosunda %50 KOH (10 mmol) çözeltisi ilave edildi. Ticari olarak satın alınan 4-florobenzaldehit (0.75 mmol) damla damla eklenerek 1 saat buz banyosunda karıştırılmaya bırakıldı ve sonrasında oda sıcaklığında karıştırıldı. İTK (90:10/petrol eteri: etil asetat) ile yapılan kontroller ile 32 saatin sonunda pirol türevi ketonun tükendiği gözlemlendi. Reaksiyonun tamamlandığı belirlendikten sonra; balon içerisindeki metanol uçuruldu. Etil asetat/ su sistemi ile ekstraksiyon yapıldı. Ardından etil asetat fazına NaSO<sub>4</sub> eklendi. Etil asetat fazı uçuruldu, katı ürün etil asetat/petrol eteri karışımı ile kristallendirilerek saflaştırıldı.

Verim	:	%62
Renk	:	Beyaz-sarı katı
Çözücü	:	Etil asetat
Rf	:	0,53 Etilasetat/n-hekzan (10:90)
<sup>1</sup> H NMR spektrumu (300 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ = 7.64-7.18 (13H, m, Ar-H), 7.49 (1H, d, J= 16.27, -CH), 6.73 (1H, d, J = 15.87 Hz, -CH), 2.53 (3H, s, -CH <sub>3</sub> ).
<sup>13</sup> C NMR spektrumu (100 MHz, DMSO-d <sub>6</sub> ) ppm	:	δ= 12.37, 115.76, 115.98, 116.19, 116.41, 121.26, 121.71, 123.65, 127.57, 128.21, 128.29, 130.07, 130.16, 130.23, 131.08, 131.32, 131.35, 134.06, 134.55, 134.79, 139.48, 160.22, 161.81, 162.66, 164.28, 187.84.
Elementel Analiz (C <sub>26</sub> H <sub>18</sub> ClF <sub>2</sub> NO)	:	Hesaplanan; C, 71.97; H, 4.18; Cl, 8.17; F, 8.76; N, 3.23; O, 3.69; Bulunan; C, 71.98; H, 4.17; Cl, 8.18; F, 8.78; N, 3.23; O, 3.69.



### 4.3. Antimikrobiyal Bulgular

Bu tez çalışmasında 19 adet 1,2,3,4-tetrasübsititü pirol-kalkon türevi bileşiğin antimikrobiyal aktiviteleri incelenmiştir. Çalışmada ampisilin referans bileşik olarak kullanılarak, *E. coli*, *P. aeuroginosa*, *E. facealis*, *S. aureus*, *S. pneumoniae* ile *B. subtilis* bakterileri üzerine antibakteriyal aktiviteleri, flukonazol referans bileşik olarak kullanılarak, *C. metapsilosis*, *C. albicans* ve *C. parapsilosis* fungusları üzerine antifungal aktiviteleri belirlenmiştir. (Tablo 4.2.)



**Tablo 4.2.** Pirol-kalkon türevi bileşiklerin antimikrobiyal aktivite sonuçları ( $\mu\text{g/mL}$ )

<b>Madde</b>	<i>Enterococcus faecalis</i>	<i>Streptococcus pneumoniae</i>	<i>Staphylococcus aureus</i>	<i>Escherichia coli</i>	<i>Klebsiella pneumoniae</i>	<i>Candida parapsilosis</i>	<i>Candida metapsilosis</i>	<i>Candida albicans</i>
MAY-1	<b>62.5</b>	250	250	250	250	250	250	250
MAY-2	500	500	500	500	500	62.5	125	62.5
MAY-3	500	500	500	<b>62.5</b>	500	<b>31.25</b>	125	62.5
MAY-4	500	500	500	500	500	125	<b>31.25</b>	62.5
MAY-5	250	250	250	250	250	62.5	125	125
MAY-6	500	500	500	250	500	<b>31.25</b>	<b>31.25</b>	<b>31.25</b>
MAY-8	125	125	125	250	<b>62.5</b>	250	250	250
MAY-9	500	500	500	500	500	<b>31.25</b>	125	500
MAY-10	125	125	125	<b>62.5</b>	<b>62.5</b>	<b>31.25</b>	62.5	62.5
MAY-11	500	500	500	500	500	62.5	125	125
MAY-13	500	250	500	500	<b>62.5</b>	500	500	500
MAY-15	500	500	500	500	500	62.5	62.5	62.5
MAY-16	500	500	125	<b>62.5</b>	<b>62.5</b>	<b>31.25</b>	62.5	62.5
MAY-17	500	500	500	500	500	125	125	125

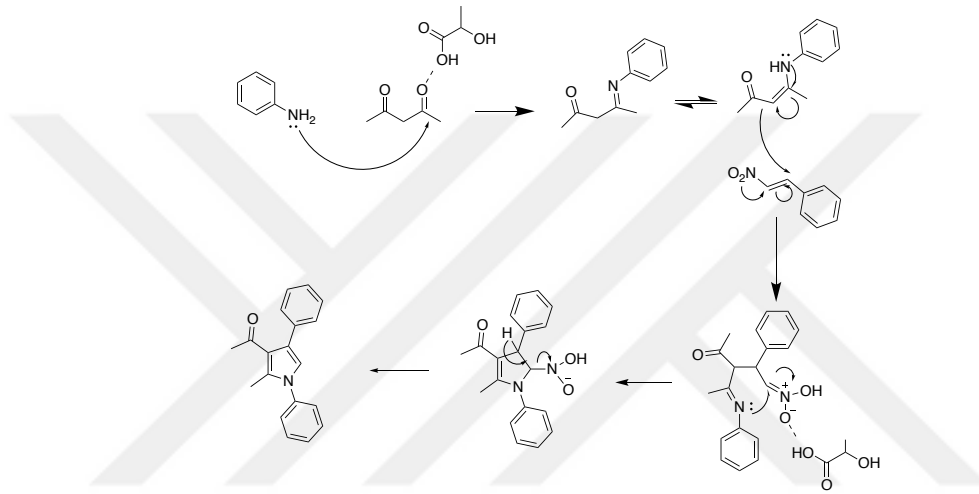
MAY-18	500	500	500	250	125	250	250	250
MAY-19	500	500	500	500	500	250	250	250
MAY-21	500	500	500	500	500	125	125	125
MAY-23	500	250	500	500	500	125	62.5	125
MAY-24	125	250	250	125	125	125	125	125
Fluconazole	*	*	*	*	*	-	-	-
Ampicillin	-	-	-	31.25	31.25	*	*	*

-: Yapılan bütün konsantrasyonlarda etkin

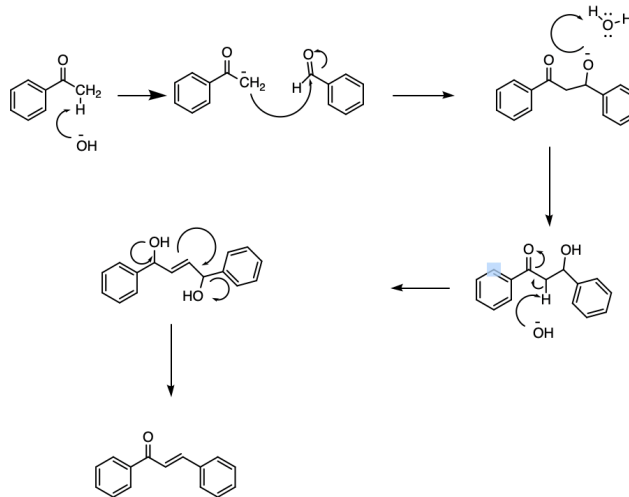
\* : Etkisiz

## 5. TARTIŞMA ve SONUÇ

Bu tez çalışması kapsamında, moleküllerin *S. aureus* ve *C. albicans* antimikrobiyal aktivitelerini tahmin edebilen 2 tane makine öğrenmesi modeli oluşturulmuştur. Tasarlanan pirol-kalkon türevi bileşiklerin antimikrobiyal aktiviteleri bu modeller ile tahmin edildiğinde hepsinin aktif olabileceği saptandı. Ardından tasarlanan 1-(2-metil-1,4-difenil-1H-pirol-3-il)-3-fenilprop-2-en-1-on genel yapısına sahip 19 molekülün (Tablo 5.1.), sentezi ve yapı tayinleri gerçekleştirildi. Genel sentez yöntemlerine ait önerilen mekanizmalar aşağıdaki şekillerde verilmiştir. (Şekil 5.1. ve Şekil 5.2.) Son olarak da moleküllerin, *S. aureus* ve *C. albicans* dahil olmak üzere çeşitli suşlar üstünde antimikrobiyal aktiviteleri incelendi.



Şekil 5.1. Genel sentez yöntemi I'de verilen reaksiyonun mekanizması

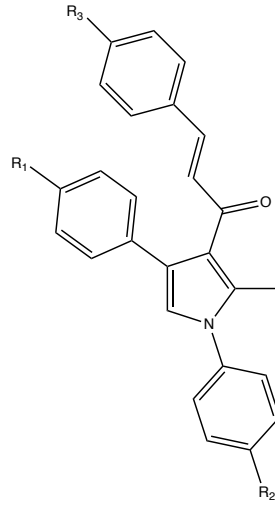


Şekil 5.2. Genel sentez yöntemi II'de verilen reaksiyonun mekanizması

Makine öğrenmesi modelleri sonucunda tasarlanan bileşiklerin antibiyotik etkiden ziyade antifungal etkisi olabileceği öngörülmüştür. Makine öğrenme modelleri verinin fazla olması nedeniyle *C. albicans* ve *S. aureus* için oluşturulmuştur. Makine öğrenmesi modelleri oluşturulurken molekülleri aktif veya inaktif olarak sınıflandırmak için eşik değeri 31.25 µg/mL olarak belirlenmiştir. *C. albicans* için aktif tahmin edilen moleküllerin (MAY-3, MAY-5, MAY-6, MAY-18, MAY-17, MAY-21, MAY-24) biyolojik aktivite sonuçları incelendiğinde en iyi aktiviteye sahip olan MAY-6 (31.25 µg/mL) molekülün doğru tahmin edildiği görülmüştür. Oluşturulan modellerde *S. aureus*'a karşı etkininin olmayacağı tahmin edilmiş ve biyolojik aktivite sonuçları incelendiğinde moleküllerin hiçbir molekülüm eşik değerinden daha iyi aktivite değerine sahip olmadığı görülmüştür.

Sentezlenen bileşikler ilk kez bu çalışma kapsamında sentezlenmiş olup yapılarının aydınlatılmasında FT-IR, <sup>1</sup>H-NMR, <sup>13</sup>C-NMR spektrumları ve erime dereceleri verilerinden yararlanılmıştır.

Bileşiklerin IR spektrumları FT-IR spektrofotometresi yardımıyla toz veya kristal örneklerden alınmıştır. Bileşiklerin aromatik halkalarına ait =C-H gerilim bandları (3200-3049 cm<sup>-1</sup>) bölgesinde değişen aralıklarda gözlenirken, aromatik halkalar üzerinde veya ara zincirde alifatik grupları taşıyan bileşiklerdeki alifatik C-H gerilim bandları (2962-2719 cm<sup>-1</sup>) bölgesinde görülmektedir. Bileşiklerin IR spektrumlarında incelenen ikinci bölge; 1700-1300 bölgesidir. Bu bölgede özellikle bileşiklerin C=C gerilim bandı (1574-1404 cm<sup>-1</sup>) ve C=O (1642-1720 cm<sup>-1</sup>) görülmektedir.



**Tablo 5.1.** Sentezlenen 1,2,3,4-Tetrasüstitüe pirol-kalkon türevi bileşikler

R <sub>1</sub>	R <sub>2</sub>	R <sub>3</sub>	Molekül
-H	-H	-H	MAY-1
-Cl	-Cl	-H	MAY-2
-Cl	-F	-H	MAY-3
-Cl	-H	-H	MAY-4
-F	-Cl	-H	MAY-5
-F	-F	-H	MAY-6
-H	-F	-H	MAY-8
-Cl	-Cl	-OCH <sub>3</sub>	MAY-9
-Cl	-F	-OCH <sub>3</sub>	MAY-10
-F	-Cl	-OCH <sub>3</sub>	MAY-11
-F	-F	- OCH <sub>3</sub>	MAY-13
-Cl	-H	- OCH <sub>3</sub>	MAY-15
-H	-H	- OCH <sub>3</sub>	MAY-16
-Cl	-Cl	- F	MAY-17
-F	-F	-F	MAY-18
-H	-F	-F	MAY-19
-Cl	-F	-F	MAY-21
-Cl	-H	-F	MAY-23
-F	-Cl	-F	MAY-24

Elde edilen bileşiklerin  $^1\text{H-NMR}$  spektrumları  $(\text{CD}_3)_2\text{SO}$  kullanılarak analiz edilmiştir. Pirol halkasında bulunan C-H protonu  $\delta$  7.10 ile 7.22 ppm değerleri arasında gözlenmiştir.  $\alpha$  ve  $\beta$ -doymamış karbonil sistemine ait  $\alpha$  konumundaki C-H protonu  $\delta$  7.45 ile 7.50 ppm değerleri arasında dublet,  $\beta$  konumundaki C-H protonu ise  $\delta$  6.61 ile 6.76 ppm arasında dublet olarak gözlenmiştir.  $\alpha$  ve  $\beta$ -doymamış karbonil sistemine sterik engelden dolayı ancak (E) geometrik izomer formunda bulunabilmektedir.  $^1\text{H-NMR}$  spektrumlarında bu durum J değerleri incelenerek gözlenebilir.  $\alpha$  ve  $\beta$ -doymamış karbonlara ait dublet piklerin J değeri incelendiğinde 15.80- 16.23 arasında olduğu saptanmıştır bu da yapının E izomeri halinde olduğunu göstermektedir. Aromatik yapılara ait hidrojenler ise beklenen aromatik bölge aralığında multipler olarak gözlenmiştir. Pirol halkasının 2.konumundan bağlı olan metil grubundaki C-H protonları ise  $\delta$  2.34 ile 2.53 ppm değerleri arasında singlet olarak gözlenmiştir. Metoksi grubu taşıyan bileşiklerin C-H protonları  $\delta$  3.78 ile 3.80 ppm değerleri arasında singlet olarak gözlenmiştir.

Elde edilen bileşiklerin  $^{13}\text{C-NMR}$  spektrumları  $(\text{CD}_3)_2\text{SO}$  kullanılarak analiz edilmiştir. Karbonil grubuna ait C atomu  $\delta$  187.55 ile 188.09. ppm değerleri arasında görülmüştür. Pirol halkasının 2.konumundan bağlı olan metil grubu ait C'na ait spektrumlar ise  $\delta$  12.31 ile 12.54 ppm arasında görülmüştür. Aromatik halkalara ait karbon atomları ile  $\alpha$  ve  $\beta$  konumlarındaki karbon atomları  $\delta$  114.33 ile 164.28 ppm değerleri arasında gözlenmiştir. Metoksi grubu taşıyan türevlerde bu gruba ait C ise  $\delta$  55.29 ile 55.32 ppm arasında gözlenmiştir.

Sentezlenen moleküllerin yapıları ve aktiviteleri incelendiğinde 4-metoksibenzaldehit ve benzaldehit türevlerinin 4-florobenzaldehit türevlerine göre daha etkili olduğu görülmüştür. 1 ve 4. konumlardaki benzen halkalarının 4. konumlarında halojen taşıması non-süstitüe türevlere göre daha aktif olduğu görülmüştür.

İlaç keşfinde makine öğrenmesi kullanılması günümüzde önemli bir araştırma alanı olarak görülmektedir. Türkiye'de ise bu alanda çalışan araştırmacı sayısı oldukça sınırlıdır. Makine öğrenimi; ilaç keşfini hızlandırarak maliyetleri düşürmeyi hedeflemektedir. Bu tez kapsamında geliştirilen makine öğrenmesi modelleri ile bileşiklerin *S. aureus* ve *C. albicans* aktivitelerini yüksek bir doğrulukla tahmin edilebilmektedir. Bu modellerin ileri çalışmalarda da kullanılabilecek olması da oldukça önemlidir. Etkileri belirlenen moleküllerin, hangi etki mekanizması ile etki ettiğinin tahmin eden bir makine öğrenmesi modeli oluşturulması da ileri hedefler aramızdadır.

## KAYNAKLAR

- [1] Chollet, F. Deep Learning with Python. **2021**.
- [2] Raschka, S.; Scikit-Learn, V.M.-; edition, and TensorFlow.S.; 2017, undefined. Python Machine Learning: Machine Learning and Deep Learning with Python. *igi-global.com*.
- [3] Pedregosa FABIANPEDREGOSA, F.; Michel, V.; Grisel OLIVIERGRISEL, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Vanderplas, J.; Cournapeau, D.; Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Thirion, B.; Grisel, O.; Dubourg, V.; Passos, A.; Brucher, M.; Perrot and Édouardand, M.; Duchesnay, and Édouard; Duchesnay EDOUARDDUCHESNAY, Fré. Scikit-Learn: Machine Learning in Python. *jmlr.org*, **2011**, *12*, 2825–2830.
- [4] Dearden, J.C. The History and Development of Quantitative Structure-Activity Relationships (QSARs). *International Journal of Quantitative Structure-Property Relationships*, **2016**, *1*, 1–44.
- [5] Hansch, C. The Advent and Evolution of QSAR at Pomona College. *J Comput Aided Mol Des*, **2011**, *25*, 495–507.
- [6] Panteleev, J.; Gao, H.; Jia, L. Recent Applications of Machine Learning in Medicinal Chemistry. *Bioorg Med Chem Lett*, **2018**, *28*, 2807–2815.
- [7] Zhavoronkov, A.; Ivanenkov, Y.A.; Aliper, A.; Veselov, M.S.; Aladinskiy, V.A.; Aladinskaya, A. v.; Terentiev, V.A.; Polykovskiy, D.A.; Kuznetsov, M.D.; Asadulaev, A.; Volkov, Y.; Zholus, A.; Shayakhmetov, R.R.; Zhebrak, A.; Minaeva, L.I.; Zagribelnyy, B.A.; Lee, L.H.; Soll, R.; Madge, D.; Xing, L.; Guo, T.; Aspuru-Guzik, A. Deep Learning Enables Rapid Identification of Potent DDR1 Kinase Inhibitors. *Nat Biotechnol*, **2019**, *37*, 1038–1040.
- [8] Yang, M.; Tao, B.; Chen, C.; Jia, W.; Sun, S.; Zhang, T.; Wang, X. Machine Learning Models Based on Molecular Fingerprints and an Extreme Gradient Boosting Method Lead to the Discovery of JAK2 Inhibitors. *J Chem Inf Model*, **2019**, *59*, 5002–5012.
- [9] Cai, C.; Guo, P.; Zhou, Y.; Zhou, J.; Wang, Q.; Zhang, F.; Fang, J.; Cheng, F. Deep Learning-Based Prediction of Drug-Induced Cardiotoxicity. *J Chem Inf Model*, **2019**, *59*, 1073–1084.
- [10] Sheridan, R.P.; Wang, W.M.; Liaw, A.; Ma, J.; Gifford, E.M. Extreme Gradient Boosting as a Method for Quantitative Structure–Activity Relationships. *J Chem Inf Model*, **2016**, *56*, 2353–2360.
- [11] Heikamp, K.; Bajorath, J. Support Vector Machines for Drug Discovery. *Expert Opin Drug Discov*, **2014**, *9*, 93–104.
- [12] Chen, H.; Engkvist, O.; Wang, Y.; Olivecrona, M.; Blaschke, T. The Rise of Deep Learning in Drug Discovery. *Drug Discov Today*, **2018**, *23*, 1241–1250.
- [13] Tin Kam Ho. *Random Decision Forests* In: Proceedings of the Proceedings of 3rd International Conference on Document Analysis and Recognition; IEEE Comput. Soc. Press; pp. 278–282.
- [14] Cunningham, P.; Delany, S.J. K-Nearest Neighbour Classifiers - A Tutorial. *ACM Comput Surv*, **2022**, *54*, 1–25.
- [15] Rogers, D.; Hahn, M. Extended-Connectivity Fingerprints. *J Chem Inf Model*, **2010**, *50*, 742–754.
- [16] Durant, J.L.; Leland, B.A.; Henry, D.R.; Nourse, J.G. Reoptimization of MDL Keys for Use in Drug Discovery. *J Chem Inf Comput Sci*, **2002**, *42*, 1273–1280.
- [17] Chuang, K. v.; Gunsalus, L.M.; Keiser, M.J. Learning Molecular Representations for Medicinal Chemistry. *J Med Chem*, **2020**, *63*, 8705–8722.
- [18] Schneider, G. Virtual Screening: An Endless Staircase? *Nat Rev Drug Discov*, **2010**, *9*, 273–276.



- [19] Scior, T.; Bender, A.; Tresadern, G.; Medina-Franco, J.L.; Martínez-Mayorga, K.; Langer, T.; Cuanalo-Contreras, K.; Agrafiotis, D.K. Recognizing Pitfalls in Virtual Screening: A Critical Review. *J Chem Inf Model*, **2012**, *52*, 867–881.
- [20] Kuntz, I.D.; Blaney, J.M.; Oatley, S.J.; Langridge, R.; Ferrin, T.E. A Geometric Approach to Macromolecule-Ligand Interactions. *J Mol Biol*, **1982**, *161*, 269–288.
- [21] Lyu, J.; Wang, S.; Balias, T.E.; Singh, I.; Levit, A.; Moroz, Y.S.; O’Meara, M.J.; Che, T.; Alga, E.; Tolmacheva, K.; Tolmachev, A.A.; Shoichet, B.K.; Roth, B.L.; Irwin, J.J. Ultra-Large Library Docking for Discovering New Chemotypes. *Nature*, **2019**, *566*, 224–229.
- [22] Sottriffer, C.A.; Gohlke, H.; Klebe, G. Docking into Knowledge-Based Potential Fields: A Comparative Evaluation of DrugScore. *J Med Chem*, **2002**, *45*, 1967–1970.
- [23] Liu, Z.; Su, M.; Han, L.; Liu, J.; Yang, Q.; Li, Y.; Wang, R. Forging the Basis for Developing Protein–Ligand Interaction Scoring Functions. *Acc Chem Res*, **2017**, *50*, 302–309.
- [24] Moitessier, N.; Englebienne, P.; Lee, D.; Lawandi, J.; Corbeil, C.R. Towards the Development of Universal, Fast and Highly Accurate Docking/Scoring Methods: A Long Way to Go. *Br J Pharmacol*, **2008**, *153*, S7–S26.
- [25] Gilson, M.K.; Zhou, H.-X. Calculation of Protein-Ligand Binding Affinities. *Annu Rev Biophys Biomol Struct*, **2007**, *36*, 21–42.
- [26] Ballester, P.J.; Mitchell, J.B.O. A Machine Learning Approach to Predicting Protein–Ligand Binding Affinity with Applications to Molecular Docking. *Bioinformatics*, **2010**, *26*, 1169–1175.
- [27] Brylinski, M. Nonlinear Scoring Functions for Similarity-Based Ligand Docking and Binding Affinity Prediction. *J Chem Inf Model*, **2013**, *53*, 3097–3112.
- [28] Ain, Q.U.; Aleksandrova, A.; Roessler, F.D.; Ballester, P.J. Machine-Learning Scoring Functions to Improve Structure-Based Binding Affinity Prediction and Virtual Screening. *Wiley Interdiscip Rev Comput Mol Sci*, **2015**, *5*, 405–424.
- [29] Li, G.-B.; Yang, L.-L.; Wang, W.-J.; Li, L.-L.; Yang, S.-Y. ID-Score: A New Empirical Scoring Function Based on a Comprehensive Set of Descriptors Related to Protein–Ligand Interactions. *J Chem Inf Model*, **2013**, *53*, 592–600.
- [30] Durrant, J.D.; McCammon, J.A. NNScore 2.0: A Neural-Network Receptor–Ligand Scoring Function. *J Chem Inf Model*, **2011**, *51*, 2897–2903.
- [31] Ballester, P.J.; Schreyer, A.; Blundell, T.L. Does a More Precise Chemical Description of Protein–Ligand Complexes Lead to More Accurate Prediction of Binding Affinity? *J Chem Inf Model*, **2014**, *54*, 944–955.
- [32] Li, H.; Leung, K.-S.; Wong, M.-H.; Ballester, P.J. Substituting Random Forest for Multiple Linear Regression Improves Binding Affinity Prediction of Scoring Functions: Cyscore as a Case Study. *BMC Bioinformatics*, **2014**, *15*, 291.
- [33] Kramer, C.; Gedeck, P. Leave-Cluster-Out Cross-Validation Is Appropriate for Scoring Functions Derived from Diverse Protein Data Sets. *J Chem Inf Model*, **2010**, *50*, 1961–1969.
- [34] Guedes, I.A.; Pereira, F.S.S.; Dardenne, L.E. Empirical Scoring Functions for Structure-Based Virtual Screening: Applications, Critical Aspects, and Challenges. *Front Pharmacol*, **2018**, *9*.
- [35] Gabel, J.; Desaphy, J.; Rognan, D. Beware of Machine Learning-Based Scoring Functions—On the Danger of Developing Black Boxes. *J Chem Inf Model*, **2014**, *54*, 2807–2815.
- [36] Li, H.; Leung, K.-S.; Wong, M.-H.; Ballester, P.J. Correcting the Impact of Docking Pose Generation Error on Binding Affinity Prediction. *BMC Bioinformatics*, **2016**, *17*, 308.
- [37] Wang, C.; Zhang, Y. Improving Scoring-Docking-Screening Powers of Protein-Ligand Scoring Functions Using Random Forest. *J Comput Chem*, **2017**, *38*, 169–177.

- [38] Li, H.; Leung, K.-S.; Wong, M.-H.; Ballester, P.J. Improving AutoDock Vina Using Random Forest: The Growing Accuracy of Binding Affinity Prediction by the Effective Exploitation of Larger Data Sets. *Mol Inform*, **2015**, *34*, 115–126.
- [39] Ragoza, M.; Hochuli, J.; Idrobo, E.; Sunseri, J.; Koes, D.R. Protein–Ligand Scoring with Convolutional Neural Networks. *J Chem Inf Model*, **2017**, *57*, 942–957.
- [40] Kinnings, S.L.; Liu, N.; Tonge, P.J.; Jackson, R.M.; Xie, L.; Bourne, P.E. A Machine Learning-Based Method To Improve Docking Scoring Functions and Its Application to Drug Repurposing. *J Chem Inf Model*, **2011**, *51*, 408–419.
- [41] Li, H.; Leung, K.-S.; Wong, M.-H.; Ballester, P. Low-Quality Structural and Interaction Data Improves Binding Affinity Prediction via Random Forest. *Molecules*, **2015**, *20*, 10947–10962.
- [42] Ballester, P.J.; Schreyer, A.; Blundell, T.L. Does a More Precise Chemical Description of Protein–Ligand Complexes Lead to More Accurate Prediction of Binding Affinity? *J Chem Inf Model*, **2014**, *54*, 944–955.
- [43] Yang, X.; Wang, Y.; Byrne, R.; Schneider, G.; Yang, S. Concepts of Artificial Intelligence for Computer-Assisted Drug Discovery. *Chem Rev*, **2019**, *119*, 10520–10594.
- [44] Johnson, M.A.; Maggiora, G.M.; American Chemical Society. Meeting (196th : 1988 : Los Angeles, Calif.). Concepts and Applications of Molecular Similarity. **1990**, 393.
- [45] Martin, E.J.; Polyakov, V.R.; Tian, L.; Perez, R.C. Profile-QSAR 2.0: Kinase Virtual Screening Accuracy Comparable to Four-Concentration IC<sub>50</sub>s for Realistically Novel Compounds. *J Chem Inf Model*, **2017**, *57*, 2077–2088.
- [46] Marchese Robinson, R.L.; Palczewska, A.; Palczewski, J.; Kidley, N. Comparison of the Predictive Performance and Interpretability of Random Forest and Linear Models on Benchmark Data Sets. *J Chem Inf Model*, **2017**, *57*, 1773–1792.
- [47] Tashchilova, A.; Podoplelova, N.; Sulimov, A.; Kutov, D.; Ilin, I.; Pantelev, M.; Shikhaliev, K.; Medvedeva, S.; Novichikhina, N.; Potapov, A.; Sulimov, V. New Blood Coagulation Factor XIIa Inhibitors: Molecular Modeling, Synthesis, and Experimental Confirmation. *Molecules*, **2022**, *27*, 1234.
- [48] Chen, J.J.F.; Visco, D.P. Developing an in Silico Pipeline for Faster Drug Candidate Discovery: Virtual High Throughput Screening with the Signature Molecular Descriptor Using Support Vector Machine Models. *undefined*, **2017**, *159*, 31–42.
- [49] Alvarsson, J.; Eklund, M.; Andersson, C.; Carlsson, L.; Spjuth, O.; Wikberg, J.E.S. Benchmarking Study of Parameter Variation When Using Signature Fingerprints Together with Support Vector Machines. *J Chem Inf Model*, **2014**, *54*, 3211–3217.
- [50] Rensi, S.E.; Altman, R.B. Shallow Representation Learning via Kernel PCA Improves QSAR Modelability. *J Chem Inf Model*, **2017**, *57*, 1859–1867.
- [51] Lenselink, E.B.; ten Dijke, N.; Bongers, B.; Papadatos, G.; van Vlijmen, H.W.T.; Kowalczyk, W.; Ijzerman, A.P.; van Westen, G.J.P. Beyond the Hype: Deep Neural Networks Outperform Established Methods Using a ChEMBL Bioactivity Benchmark Set. *J Cheminform*, **2017**, *9*.
- [52] Speck-Planche, A.; Cordeiro, M.N.D.S. Fragment-Based in Silico Modeling of Multi-Target Inhibitors against Breast Cancer-Related Proteins. *Mol Divers*, **2017**, *21*, 511–523.
- [53] Ma, J.; Sheridan, R.P.; Liaw, A.; Dahl, G.E.; Svetnik, V. Deep Neural Nets as a Method for Quantitative Structure–Activity Relationships. *J Chem Inf Model*, **2015**, *55*, 263–274.
- [54] Dahl, G.E.; Jaitly, N.; Salakhutdinov, R. Multi-Task Neural Networks for QSAR Predictions. **2014**.
- [55] Keiser, M.J.; Roth, B.L.; Armbruster, B.N.; Ernsberger, P.; Irwin, J.J.; Shoichet, B.K. Relating Protein Pharmacology by Ligand Chemistry. *Nat Biotechnol*, **2007**, *25*, 197–206.

- [56] Ramsundar, B.; Kearnes, S.; Riley, P.; Webster, D.; Konerding, D.; Pande, V. Massively Multitask Networks for Drug Discovery. **2015**.
- [57] Xu, Y.; Ma, J.; Liaw, A.; Sheridan, R.P.; Svetnik, V. Demystifying Multitask Deep Neural Networks for Quantitative Structure-Activity Relationships. *J Chem Inf Model*, **2017**, *57*, 2490–2504.
- [58] Mendenhall, J.; Meiler, J. Improving Quantitative Structure-Activity Relationship Models Using Artificial Neural Networks Trained with Dropout. *J Comput Aided Mol Des*, **2016**, *30*, 177–189.
- [59] Koutsoukas, A.; Monaghan, K.J.; Li, X.; Huan, J. Deep-Learning: Investigating Deep Neural Networks Hyper-Parameters and Comparison of Performance to Shallow Methods for Modeling Bioactivity Data. *J Cheminform*, **2017**, *9*, 1–13.
- [60] Winkler, D.A.; Le, T.C. Performance of Deep and Shallow Neural Networks, the Universal Approximation Theorem, Activity Cliffs, and QSAR. *Mol Inform*, **2017**, *36*.
- [61] Wu, Z.; Ramsundar, B.; Feinberg, E.N.; Gomes, J.; Geniesse, C.; Pappu, A.S.; Leswing, K.; Pande, V. MoleculeNet: A Benchmark for Molecular Machine Learning. *Chem Sci*, **2018**, *9*, 513–530.
- [62] Duvenaud, D.; Maclaurin, D.; Aguilera-Iparraguirre, J.; Gómez-Bombarelli, R.; Hirzel, T.; Aspuru-Guzik, A.; Adams, R.P. Convolutional Networks on Graphs for Learning Molecular Fingerprints. *Adv Neural Inf Process Syst*, **2015**, *2015-January*, 2224–2232.
- [63] High-throughput quantum chemistry and virtual screening for OLED material components. <https://www.lens.org/lens/scholar/article/082-561-627-619-240/main> (Accessed Nov 7, **2022**).
- [64] Halls, M.D.; Tasaki, K. High-Throughput Quantum Chemistry and Virtual Screening for Lithium Ion Battery Electrolyte Additives. *J Power Sources*, **2010**, *5*, 1472–1478.
- [65] Morgan, P.; van der Graaf, P.H.; Arrowsmith, J.; Feltner, D.E.; Drummond, K.S.; Wegner, C.D.; Street, S.D.A. Can the Flow of Medicines Be Improved? Fundamental Pharmacokinetic and Pharmacological Principles toward Improving Phase II Survival. *Drug Discov Today*, **2012**, *17*, 419–424.
- [66] Cumming, J.G.; Davis, A.M.; Muresan, S.; Haeberlein, M.; Chen, H. Chemical Predictive Modelling to Improve Compound Quality. *Nat Rev Drug Discov*, **2013**, *12*, 948–962.
- [67] In Vitro and In Vivo Assessment of ADME and PK Properties During Lead Selection and Lead Optimization – Guidelines, Benchmarks and Rules of Thumb - PubMed. <https://pubmed.ncbi.nlm.nih.gov/26561695/> (Accessed Nov 7, **2022**).
- [68] Durán-Iturbide, N.A.; Díaz-Eufracio, B.I.; Medina-Franco, J.L. In Silico ADME/Tox Profiling of Natural Products: A Focus on BIOFACQUIM. *ACS Omega*, **2020**, *5*, 16076–16084.
- [69] Harvey, A.L.; Edrada-Ebel, R.; Quinn, R.J. The Re-Emergence of Natural Products for Drug Discovery in the Genomics Era. *Nat Rev Drug Discov*, **2015**, *14*, 111–129.
- [70] van de Waterbeemd, H. Improving Compound Quality through in Vitro and in Silico Physicochemical Profiling. *Chem Biodivers*, **2009**, *6*, 1760–1766.
- [71] Morphy, R.; Rankovic, Z. The Physicochemical Challenges of Designing Multiple Ligands. *J Med Chem*, **2006**, *49*, 4961–4970.
- [72] Waring, M.J. Lipophilicity in Drug Discovery. *Expert Opin Drug Discov*, **2010**, *5*, 235–248.
- [73] Tetko, I. v.; Tanchuk, V.Yu. Application of Associative Neural Networks for Prediction of Lipophilicity in ALOGPS 2.1 Program. *J Chem Inf Comput Sci*, **2002**, *42*, 1136–1145.
- [74] Cheng, T.; Zhao, Y.; Li, X.; Lin, F.; Xu, Y.; Zhang, X.; Li, Y.; Wang, R.; Lai, L. Computation of Octanol–Water Partition Coefficients by Guiding an Additive Model with Knowledge. *J Chem Inf Model*, **2007**, *47*, 2140–2148.

- [75] Asanuma, D.; Sakabe, M.; Kamiya, M.; Yamamoto, K.; Hiratake, J.; Ogawa, M.; Kosaka, N.; Choyke, P.L.; Nagano, T.; Kobayashi, H.; Urano, Y. Sensitive  $\beta$ -Galactosidase-Targeting Fluorescence Probe for Visualizing Small Peritoneal Metastatic Tumours in Vivo. *Nat Commun*, **2015**, *6*, 6463.
- [76] Stegemann, S.; Leveiller, F.; Franchi, D.; de Jong, H.; Lindén, H. When Poor Solubility Becomes an Issue: From Early Stage to Proof of Concept. *European Journal of Pharmaceutical Sciences*, **2007**, *31*, 249–261.
- [77] Jorgensen, W.L.; Duffy, E.M. Prediction of Drug Solubility from Monte Carlo Simulations. *Bioorg Med Chem Lett*, **2000**, *10*, 1155–1158.
- [78] McDonagh, J.L.; Nath, N.; de Ferrari, L.; van Mourik, T.; Mitchell, J.B.O. Uniting Cheminformatics and Chemical Theory To Predict the Intrinsic Aqueous Solubility of Crystalline Druglike Molecules. *J Chem Inf Model*, **2014**, *54*, 844–856.
- [79] Wang, J.; Hou, T.; Xu, X. Aqueous Solubility Prediction Based on Weighted Atom Type Counts and Solvent Accessible Surface Areas. *J Chem Inf Model*, **2009**, *49*, 571–581.
- [80] Delaney, J.S. ESOL: Estimating Aqueous Solubility Directly from Molecular Structure. *J Chem Inf Comput Sci*, **2004**, *44*, 1000–1005.
- [81] Kühne, R.; Ebert, R.-U.; Kleint, F.; Schmidt, G.; Schüürmann, G. Group Contribution Methods to Estimate Water Solubility of Organic Chemicals. *Chemosphere*, **1995**, *30*, 2061–2077.
- [82] Roberto Todeschini, V.C.R.M.H.K.H.T. *Handbook of Molecular Descriptors*; Wiley: Weinheim, Germany, **2002**.
- [83] Kamlet, M.J.; Doherty, R.M.; Abboud, J.-L.M.; Abraham, M.H.; Taft, R.W. Linear Solvation Energy Relationships: 36. Molecular Properties Governing Solubilities of Organic Nonelectrolytes In Water. *J Pharm Sci*, **1986**, *75*, 338–349.
- [84] Unger, S.H. Molecular Connectivity in Structure–Activity Analysis. *J Pharm Sci*, **1987**, *76*, 269–270.
- [85] Zhou, D.; Alelyunas, Y.; Liu, R. Scores of Extended Connectivity Fingerprint as Descriptors in QSPR Study of Melting Point and Aqueous Solubility. *J Chem Inf Model*, **2008**, *48*, 981–987.
- [86] Hewitt, M.; Cronin, M.T.D.; Enoch, S.J.; Madden, J.C.; Roberts, D.W.; Dearden, J.C. *In Silico* Prediction of Aqueous Solubility: The Solubility Challenge. *J Chem Inf Model*, **2009**, *49*, 2572–2587.
- [87] Llinàs, A.; Glen, R.C.; Goodman, J.M. Solubility Challenge: Can You Predict Solubilities of 32 Molecules Using a Database of 100 Reliable Measurements? *J Chem Inf Model*, **2008**, *48*, 1289–1303.
- [88] Palmer, D.S.; Mitchell, J.B.O. Is Experimental Data Quality the Limiting Factor in Predicting the Aqueous Solubility of Druglike Molecules? *Mol Pharm*, **2014**, *11*, 2962–2972.
- [89] Zheng, T.; Hopfinger, A.J.; Esposito, E.X.; Liu, J.; Tseng, Y.J. Membrane-Interaction Quantitative Structure–Activity Relationship (MI-QSAR) Analyses of Skin Penetration Enhancers. *J Chem Inf Model*, **2008**, *48*, 1238–1256.
- [90] Tang, F.; Horie, K.; Borchardt, R.T. Are MDCK Cells Transfected with the Human MDR1 Gene a Good Model of the Human Intestinal Mucosa? *Pharm Res*, **2002**, *19*, 765–772.
- [91] Ballard, P.; Brassil, P.; Bui, K.H.; Dolgos, H.; Petersson, C.; Tunek, A.; Webborn, P.J.H. The Right Compound in the Right Assay at the Right Time: An Integrated Discovery DMPK Strategy. *Drug Metab Rev*, **2012**, *44*, 224–252.
- [92] Avdeef, A.; Artursson, P.; Neuhoff, S.; Lazorova, L.; Gråsjö, J.; Tavelin, S. Caco-2 Permeability of Weakly Basic Drugs Predicted with the Double-Sink PAMPA Method. *European Journal of Pharmaceutical Sciences*, **2005**, *24*, 333–349.

- [93] Chillistone, S.; Hardman, J.G. Factors Affecting Drug Absorption and Distribution. *Anaesthesia & Intensive Care Medicine*, **2017**, *18*, 335–339.
- [94] Shen, J.; Cheng, F.; Xu, Y.; Li, W.; Tang, Y. Estimation of ADME Properties with Substructure Pattern Recognition. *J Chem Inf Model*, **2010**, *50*, 1034–1041.
- [95] Billat, P.-A.; Roger, E.; Faure, S.; Lagarce, F. Models for Drug Absorption from the Small Intestine: Where Are We and Where Are We Going? *Drug Discov Today*, **2017**, *22*, 761–775.
- [96] Basant, N.; Gupta, S.; Singh, K.P. Predicting Human Intestinal Absorption of Diverse Chemicals Using Ensemble Learning Based QSAR Modeling Approaches. *Comput Biol Chem*, **2016**, *61*, 178–196.
- [97] Shin, M.; Jang, D.; Nam, H.; Lee, K.H.; Lee, D. Predicting the Absorption Potential of Chemical Compounds Through a Deep Learning Approach. *IEEE/ACM Trans Comput Biol Bioinform*, **2018**, *15*, 432–440.
- [98] Wang, N.-N.; Dong, J.; Deng, Y.-H.; Zhu, M.-F.; Wen, M.; Yao, Z.-J.; Lu, A.-P.; Wang, J.-B.; Cao, D.-S. ADME Properties Evaluation in Drug Discovery: Prediction of Caco-2 Cell Permeability Using a Combination of NSGA-II and Boosting. *J Chem Inf Model*, **2016**, *56*, 763–773.
- [99] McNamara, P.J.; Leggas, M. Drug Distribution. In: *Pharmacology*; Elsevier, **2009**; pp. 113–129.
- [100] Basant, N.; Gupta, S.; Singh, K.P. Predicting Binding Affinities of Diverse Pharmaceutical Chemicals to Human Serum Plasma Proteins Using QSPR Modelling Approaches. *SAR QSAR Environ Res*, **2016**, *27*, 67–85.
- [101] Sun, L.; Yang, H.; Li, J.; Wang, T.; Li, W.; Liu, G.; Tang, Y. In Silico Prediction of Compounds Binding to Human Plasma Proteins by QSAR Models. *ChemMedChem*, **2018**, *13*, 572–581.
- [102] Sheweita, S. Drug-Metabolizing Enzymes Mechanisms and Functions. *Curr Drug Metab*, **2000**, *1*, 107–132.
- [103] Zaretski, J.; Bergeron, C.; Huang, T. -w.; Rydberg, P.; Swamidass, S.J.; Breneman, C.M. RS-WebPredictor: A Server for Predicting CYP-Mediated Sites of Metabolism on Drug-like Molecules. *Bioinformatics*, **2013**, *29*, 497–498.
- [104] Zaretski, J.; Rydberg, P.; Bergeron, C.; Bennett, K.P.; Olsen, L.; Breneman, C.M. RS-Predictor Models Augmented with SMARTCyp Reactivities: Robust Metabolic Regioselectivity Predictions for Nine CYP Isozymes. *J Chem Inf Model*, **2012**, *52*, 1637–1659.
- [105] Zaretski, J.; Bergeron, C.; Rydberg, P.; Huang, T.; Bennett, K.P.; Breneman, C.M. RS-Predictor: A New Tool for Predicting Sites of Cytochrome P450-Mediated Metabolism Applied to CYP 3A4. *J Chem Inf Model*, **2011**, *51*, 1667–1689.
- [106] Taft, D.R. Drug Excretion. In: *Pharmacology*; Elsevier, **2009**; pp. 175–199.
- [107] Berellini, G.; Waters, N.J.; Lombardo, F. In Silico Prediction of Total Human Plasma Clearance. *J Chem Inf Model*, **2012**, *52*, 2069–2078.
- [108] Toshimoto, K.; Wakayama, N.; Kusama, M.; Maeda, K.; Sugiyama, Y.; Akiyama, Y. In Silico Prediction of Major Drug Clearance Pathways by Support Vector Machines with Feature-Selected Descriptors. *Drug Metabolism and Disposition*, **2014**, *42*, 1811–1819.
- [109] Guengerich, F.P. Mechanisms of Drug Toxicity and Relevance to Pharmaceutical Development. *Drug Metab Pharmacokinet*, **2011**, *26*, 3–14.
- [110] Xu, Y.; Pei, J.; Lai, L. Deep Learning Based Regression and Multiclass Models for Acute Oral Toxicity Prediction with Automatic Chemical Feature Extraction. *J Chem Inf Model*, **2017**, *57*, 2672–2685.

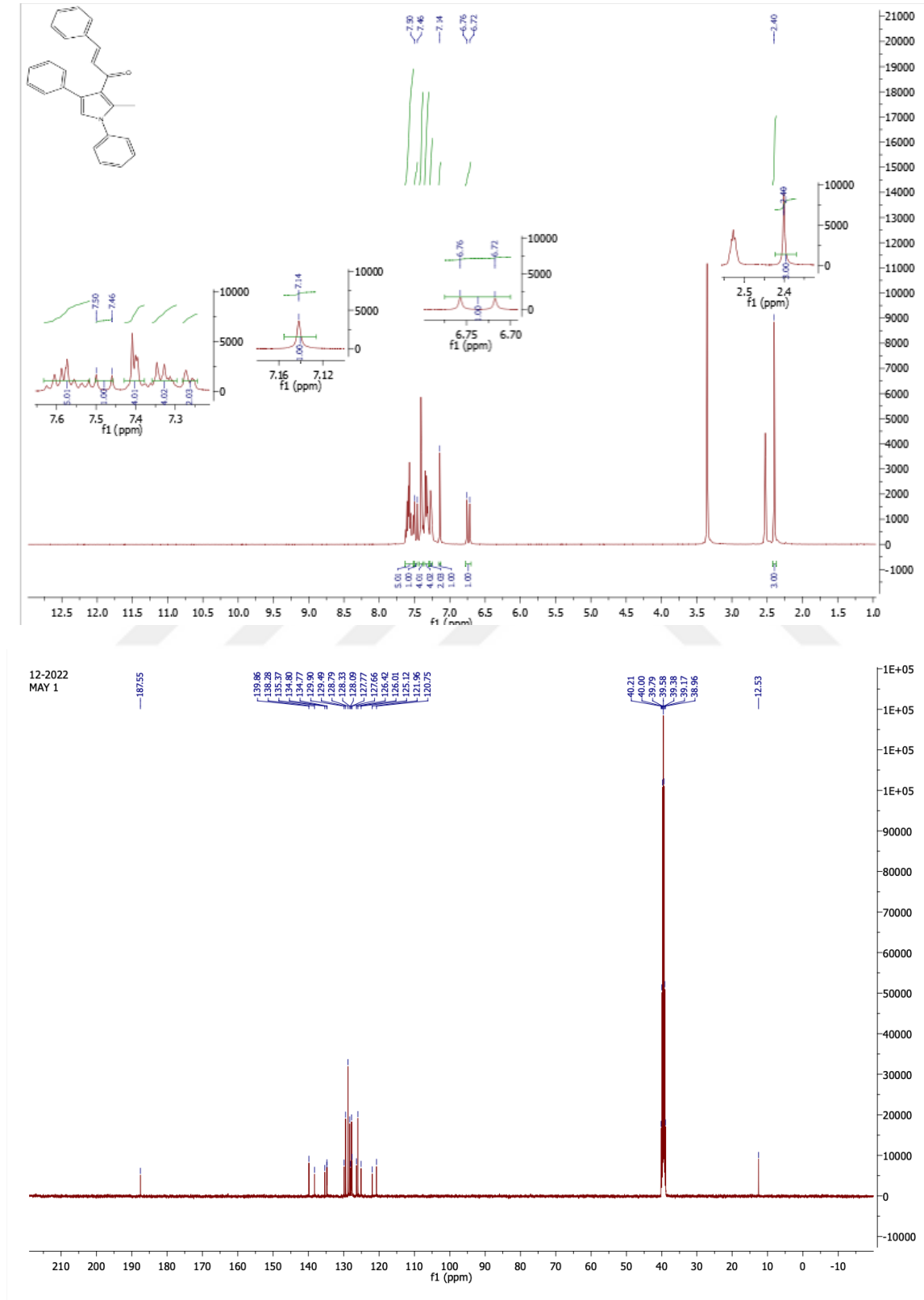
- [111] Patlewicz, G.; Jeliaskova, N.; Safford, R.J.; Worth, A.P.; Aleksiev, B. An Evaluation of the Implementation of the Cramer Classification Scheme in the Toxtree Software. *SAR QSAR Environ Res*, **2008**, *19*, 495–524.
- [112] Cheng, F.; Li, W.; Zhou, Y.; Shen, J.; Wu, Z.; Liu, G.; Lee, P.W.; Tang, Y. AdmetSAR: A Comprehensive Source and Free Tool for Assessment of Chemical ADMET Properties. *J Chem Inf Model*, **2012**, *52*, 3099–3105.
- [113] Pires, D.E. v.; Blundell, T.L.; Ascher, D.B. PkCSM: Predicting Small-Molecule Pharmacokinetic and Toxicity Properties Using Graph-Based Signatures. *J Med Chem*, **2015**, *58*, 4066–4072.
- [114] Cañada, A.; Capella-Gutierrez, S.; Rabal, O.; Oyarzabal, J.; Valencia, A.; Krallinger, M. LimTox: A Web Tool for Applied Text Mining of Adverse Event and Toxicity Associations of Compounds, Drugs and Genes. *Nucleic Acids Res*, **2017**, *45*, W484–W489.
- [115] Zhang, H.; Ren, J.-X.; Kang, Y.-L.; Bo, P.; Liang, J.-Y.; Ding, L.; Kong, W.-B.; Zhang, J. Development of Novel in Silico Model for Developmental Toxicity Assessment by Using Naïve Bayes Classifier Method. *Reproductive Toxicology*, **2017**, *71*, 8–15.
- [116] Uesawa, Y. Rigorous Selection of Random Forest Models for Identifying Compounds That Activate Toxicity-Related Pathways. *Front Environ Sci*, **2016**, *4*.
- [117] Zhang, C.; Cheng, F.; Li, W.; Liu, G.; Lee, P.W.; Tang, Y. In Silico Prediction of Drug Induced Liver Toxicity Using Substructure Pattern Recognition Method. *Mol Inform*, **2016**, *35*, 136–144.
- [118] Chavan, S.; Abdelaziz, A.; Wiklander, J.G.; Nicholls, I.A. A K-Nearest Neighbor Classification of HERG K<sup>+</sup> Channel Blockers. *J Comput Aided Mol Des*, **2016**, *30*, 229–236.
- [119] Mayr, A.; Klambauer, G.; Unterthiner, T.; Hochreiter, S. DeepTox: Toxicity Prediction Using Deep Learning. *Front Environ Sci*, **2016**, *3*.
- [120] Pushpakom, S.; Iorio, F.; Eyers, P.A.; Escott, K.J.; Hopper, S.; Wells, A.; Doig, A.; Williams, T.; Latimer, J.; McNamee, C.; Norris, A.; Sanseau, P.; Cavalla, D.; Pirmohamed, M. Drug Repurposing: Progress, Challenges and Recommendations. *Nat Rev Drug Discov*, **2019**, *18*, 41–58.
- [121] Zeng, X.; Zhu, S.; Lu, W.; Liu, Z.; Huang, J.; Zhou, Y.; Fang, J.; Huang, Y.; Guo, H.; Li, L.; Trapp, B.D.; Nussinov, R.; Eng, C.; Loscalzo, J.; Cheng, F. Target Identification among Known Drugs by Deep Learning from Heterogeneous Networks. *Chem Sci*, **2020**, *11*, 1775–1797.
- [122] Mater, A.C.; Coote, M.L. Deep Learning in Chemistry. *J Chem Inf Model*, **2019**, *59*, 2545–2559.
- [123] Stokes, J.M.; Yang, K.; Swanson, K.; Jin, W.; Cubillos-Ruiz, A.; Donghia, N.M.; MacNair, C.R.; French, S.; Carfrae, L.A.; Bloom-Ackermann, Z.; Tran, V.M.; Chiappino-Pepe, A.; Badran, A.H.; Andrews, I.W.; Chory, E.J.; Church, G.M.; Brown, E.D.; Jaakkola, T.S.; Barzilay, R.; Collins, J.J. A Deep Learning Approach to Antibiotic Discovery. *Cell*, **2020**, *180*, 688–702.e13.
- [124] Lane, T.; Russo, D.P.; Zorn, K.M.; Clark, A.M.; Korotcov, A.; Tkachenko, V.; Reynolds, R.C.; Perryman, A.L.; Freundlich, J.S.; Ekins, S. Comparing and Validating Machine Learning Models for *Mycobacterium Tuberculosis* Drug Discovery. *Mol Pharm*, **2018**, *15*, 4346–4360.
- [125] Ekins, S.; Godbole, A.A.; Kéri, G.; Orfi, L.; Pato, J.; Bhat, R.S.; Verma, R.; Bradley, E.K.; Nagaraja, V. Machine Learning and Docking Models for *Mycobacterium Tuberculosis* Topoisomerase I. *Tuberculosis*, **2017**, *103*, 52–60.
- [126] Lane, T.R.; Urbina, F.; Rank, L.; Gerlach, J.; Riabova, O.; Lepioshkin, A.; Kazakova, E.; Vocat, A.; Tkachenko, V.; Cole, S.; Makarov, V.; Ekins, S. Machine Learning Models for *Mycobacterium Tuberculosis* In Vitro Activity: Prediction and Target Visualization. *Mol Pharm*, **2022**, *19*, 674–689.

- [127] Ekins, S.; Madrid, P.B.; Sarker, M.; Li, S.-G.; Mittal, N.; Kumar, P.; Wang, X.; Stratton, T.P.; Zimmerman, M.; Talcott, C.; Bourbon, P.; Travers, M.; Yadav, M.; Freundlich, J.S. Combining Metabolite-Based Pharmacophores with Bayesian Machine Learning Models for Mycobacterium Tuberculosis Drug Discovery. *PLoS One*, **2015**, *10*, e0141076.
- [128] The pandas development team. Pandas. *Zenodo*.
- [129] Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; Kern, R.; Picus, M.; Hoyer, S.; van Kerkwijk, M.H.; Brett, M.; Haldane, A.; del Río, J.F.; Wiebe, M.; Peterson, P.; Gérard-Marchant, P.; Sheppard, K.; Reddy, T.; Weckesser, W.; Abbasi, H.; Gohlke, C.; Oliphant, T.E. Array Programming with NumPy. *Nature*, **2020**, *585*, 357–362.
- [130] Landrum, G. RDKit: Open-Source Cheminformatics. <https://www.rdkit.org>.
- [131] Batt, D.G.; Goodman, R.; Jones, D.G.; Kerr, J.S.; Mantegna, L.R.; McAllister, C.; Newton, R.C.; Nurnberg, S.; Welch, P.K.; Covington, M.B. 2'-Substituted Chalcone Derivatives as Inhibitors of Interleukin-1 Biosynthesis. *J Med Chem*, **1993**, *36*, 1434–1442.
- [132] *Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically Approved Standard (M7A5)*, National Committee for Clinical Laboratory Standards. 5th. Ed. *Methods for Dilution*; **2000**.
- [133] Clinical and Laboratory Standards Institute (CLSI). [www.clsi.org/wpcontent/uploads/sites/14/2013/07/CLSI-2015-Catalog.pdf](http://www.clsi.org/wpcontent/uploads/sites/14/2013/07/CLSI-2015-Catalog.pdf). (Accessed Mar 31, **2018**).

## EKLER

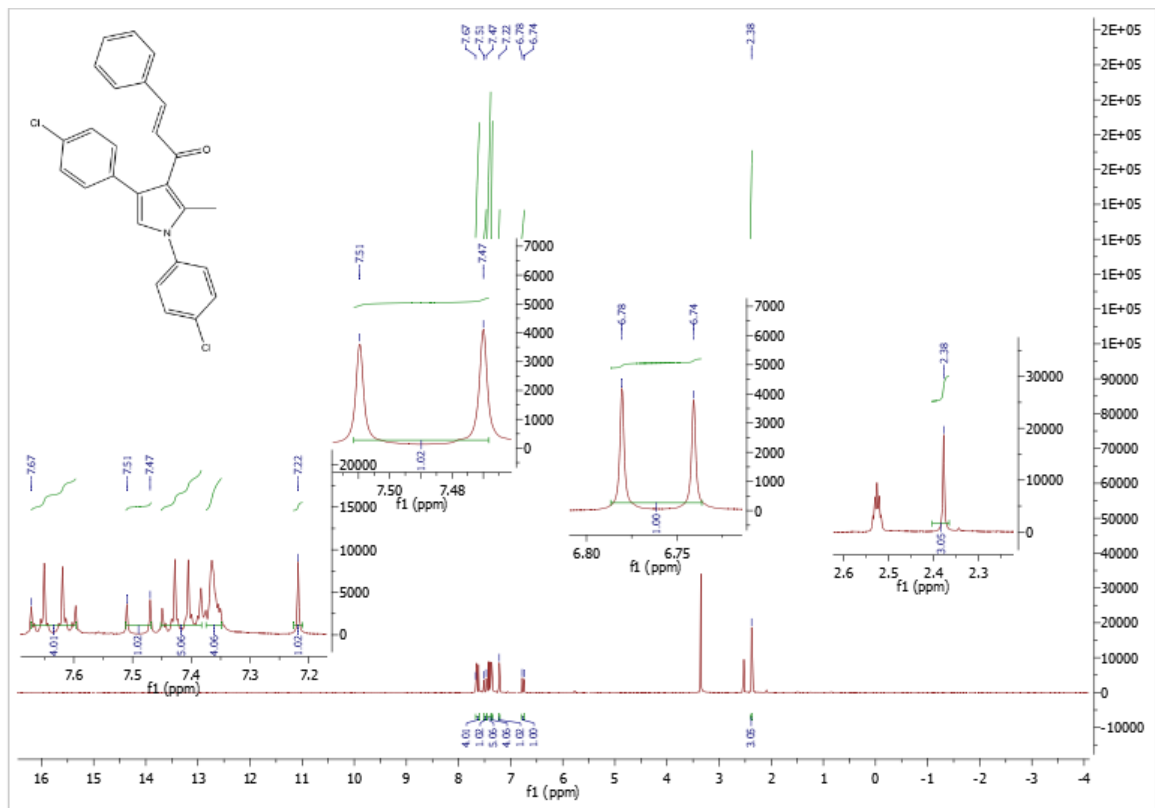
Ek 1: Sentezlenen moleküllerin  $^1\text{H}$  ve  $^{13}\text{C}$  NMR'larına ait spektrumlar

### MAY-1 molekülüne ait $^1\text{H}$ -NMR ve $^{13}\text{C}$ -NMR spektrumları

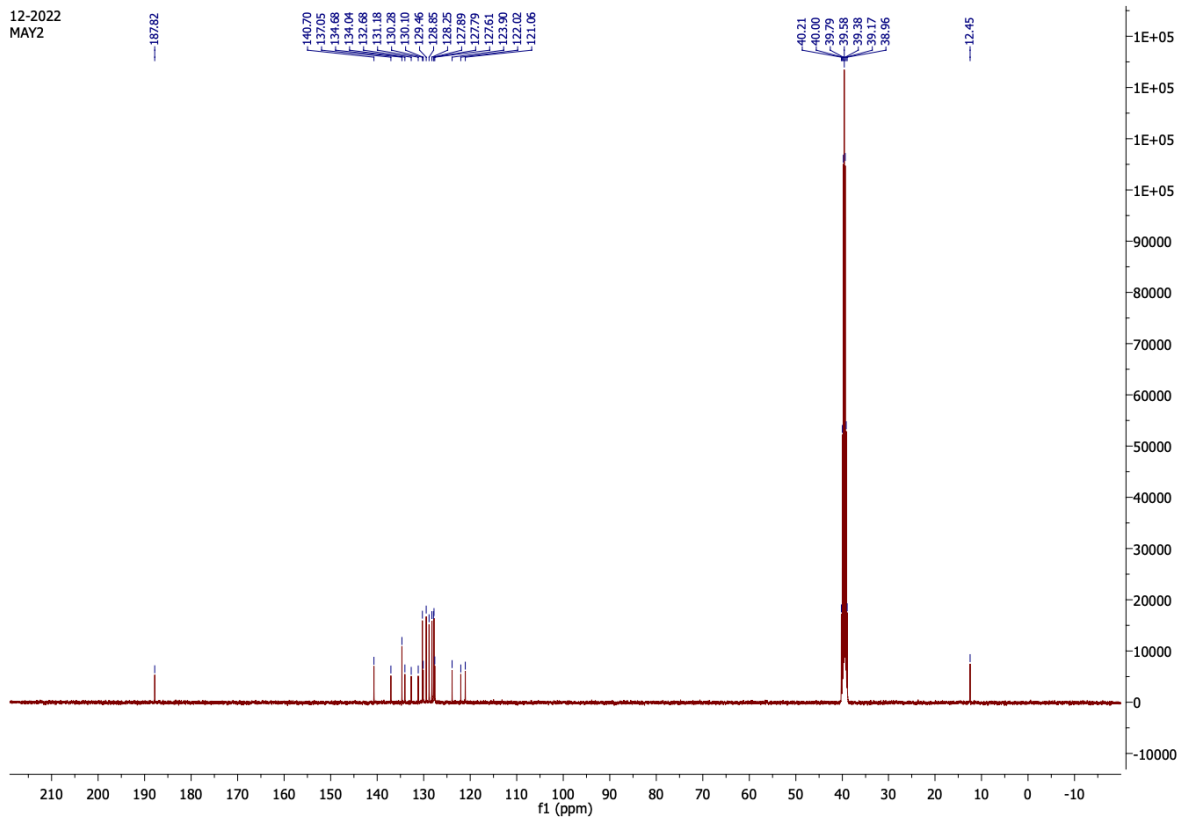




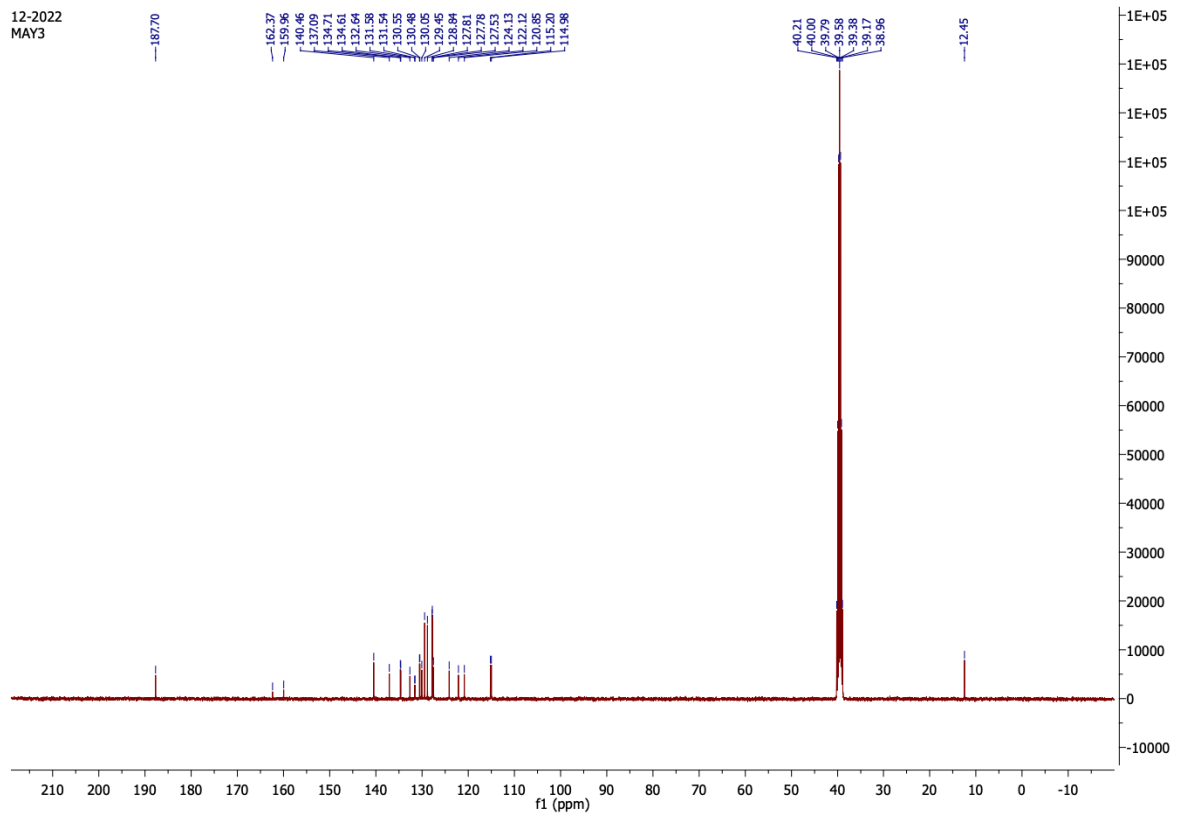
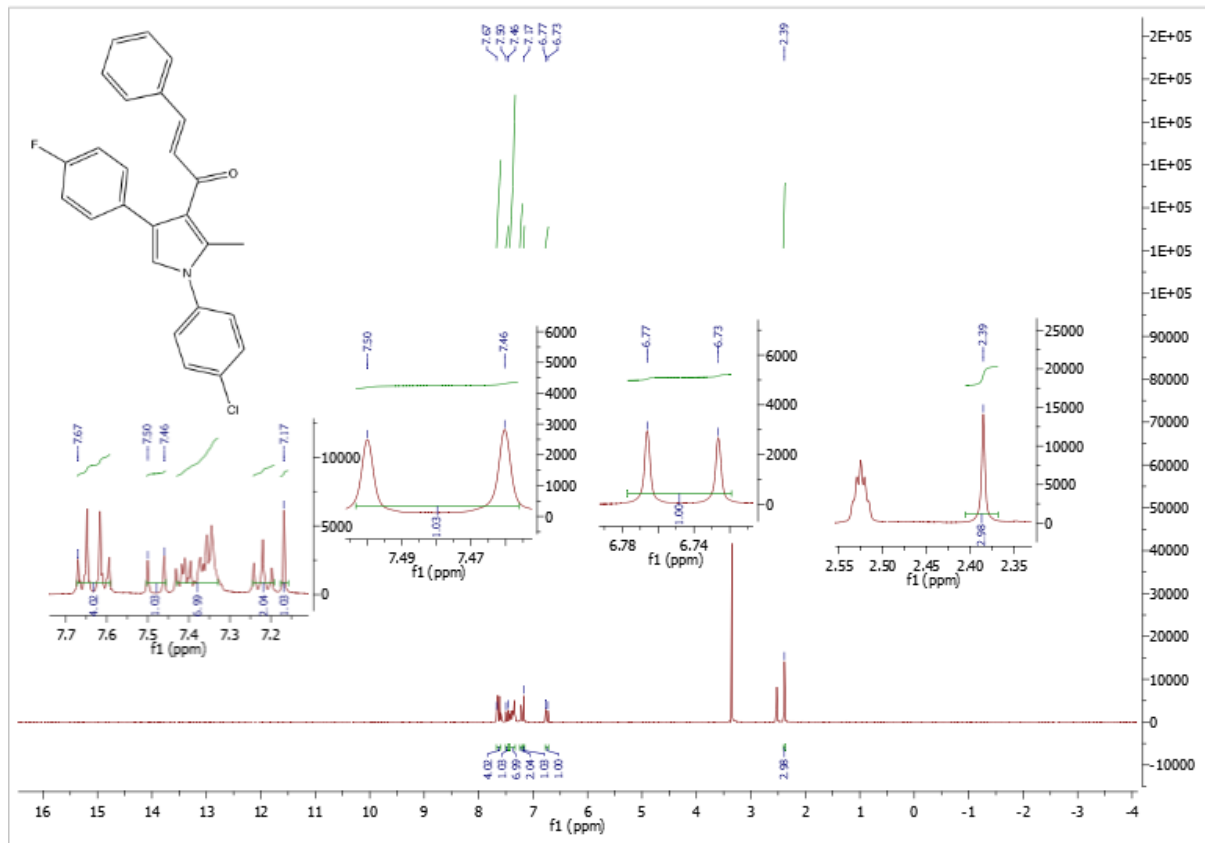
### MAY-2 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



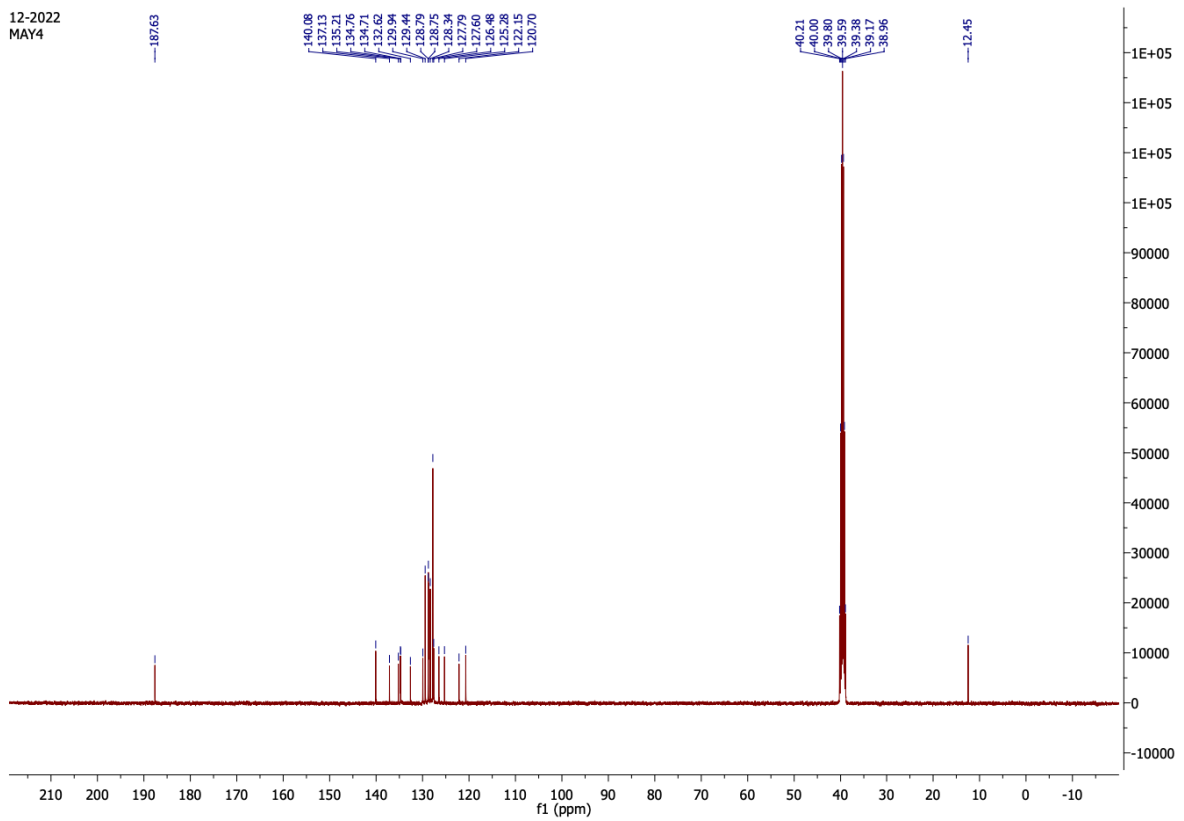
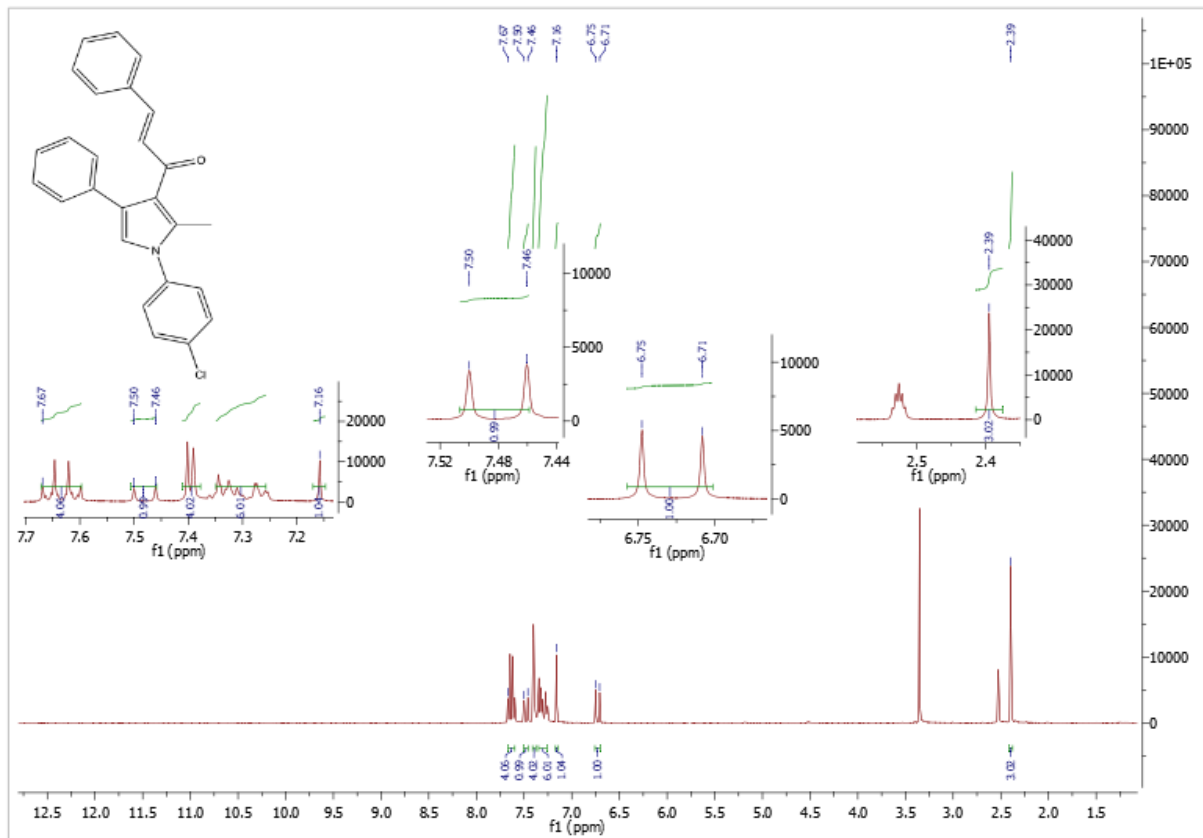
12-2022  
MAY2



### MAY-3 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları

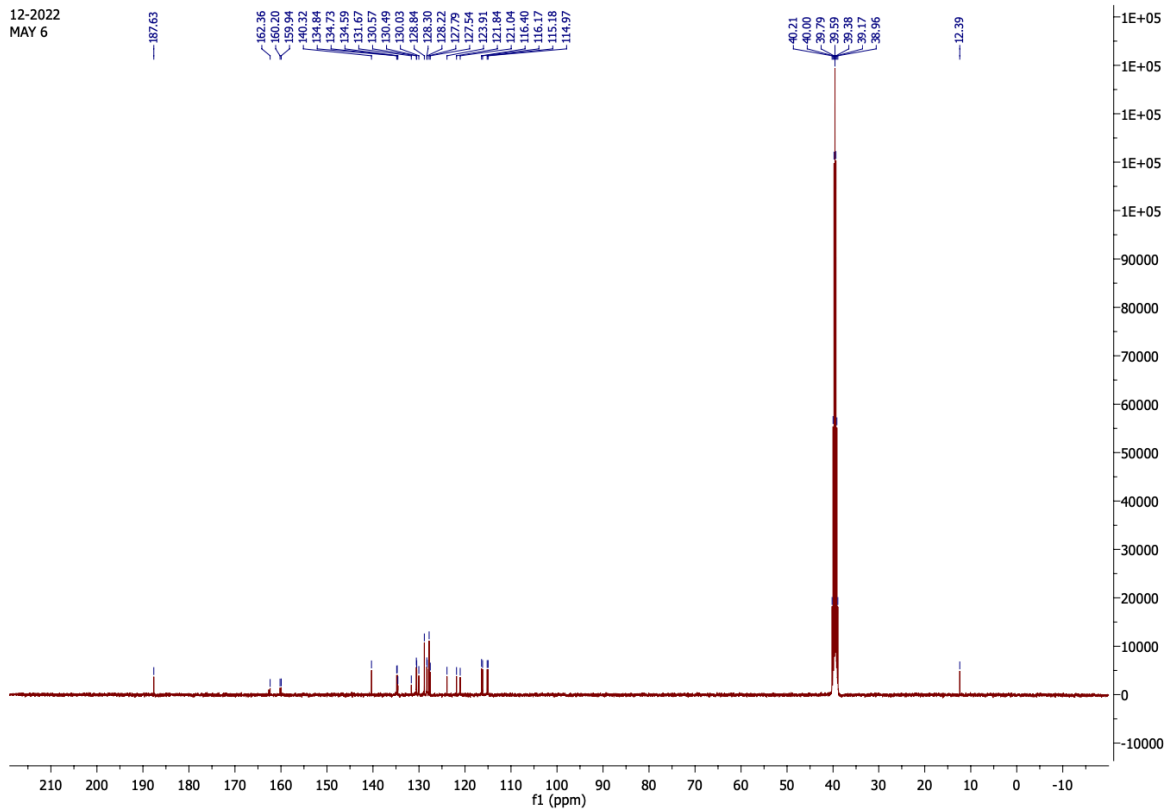
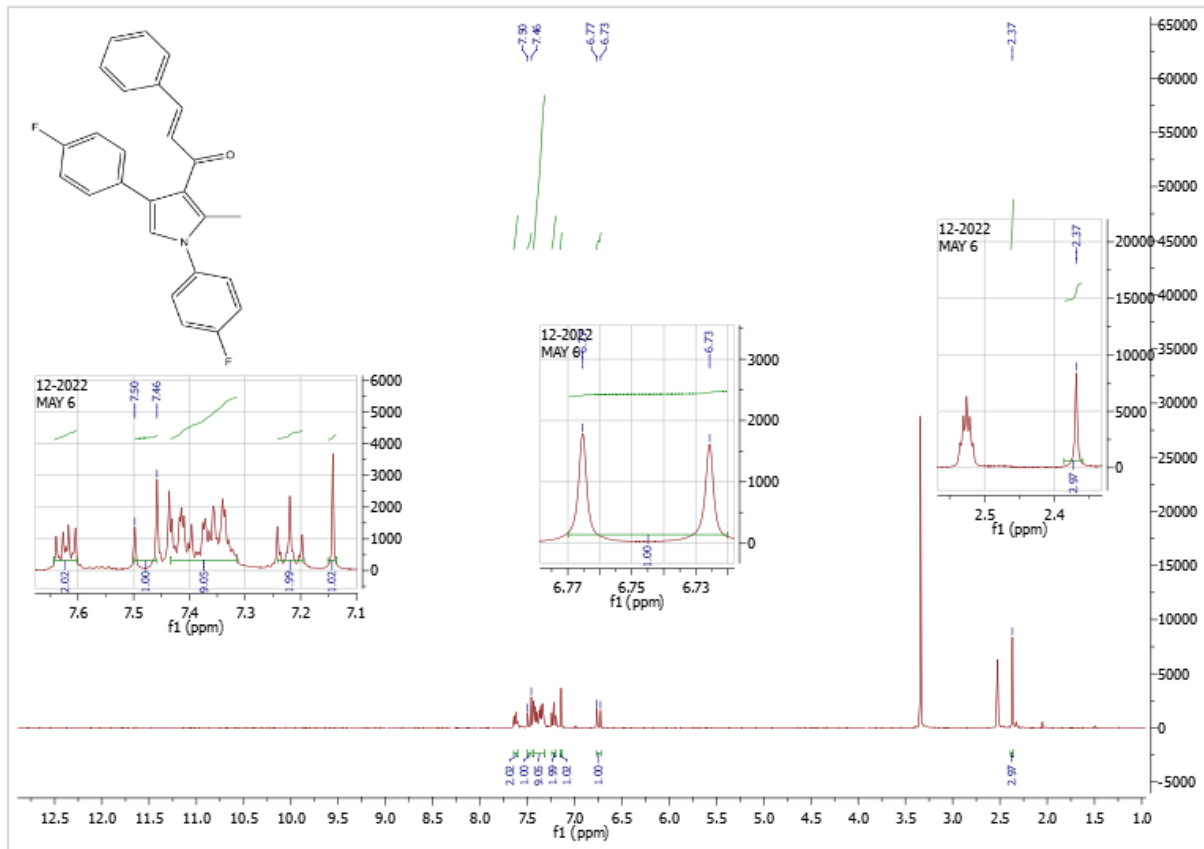


### MAY-4 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları

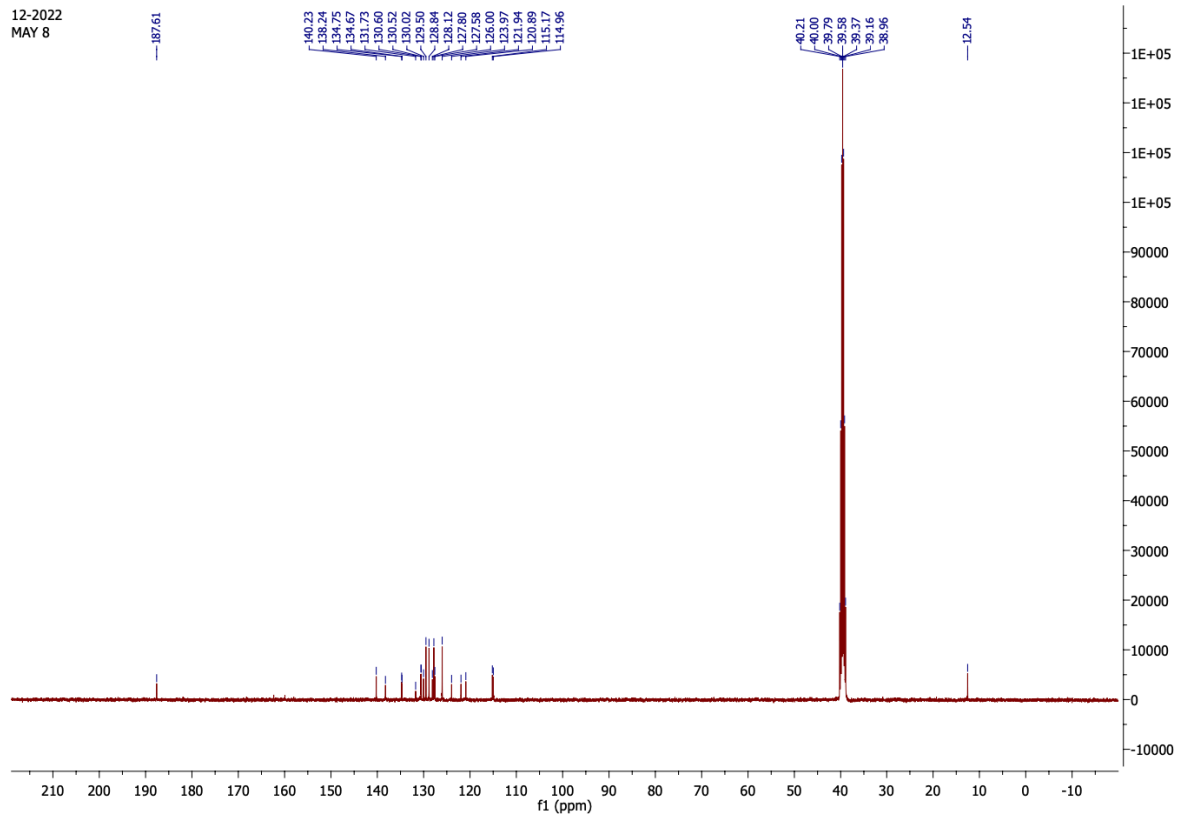
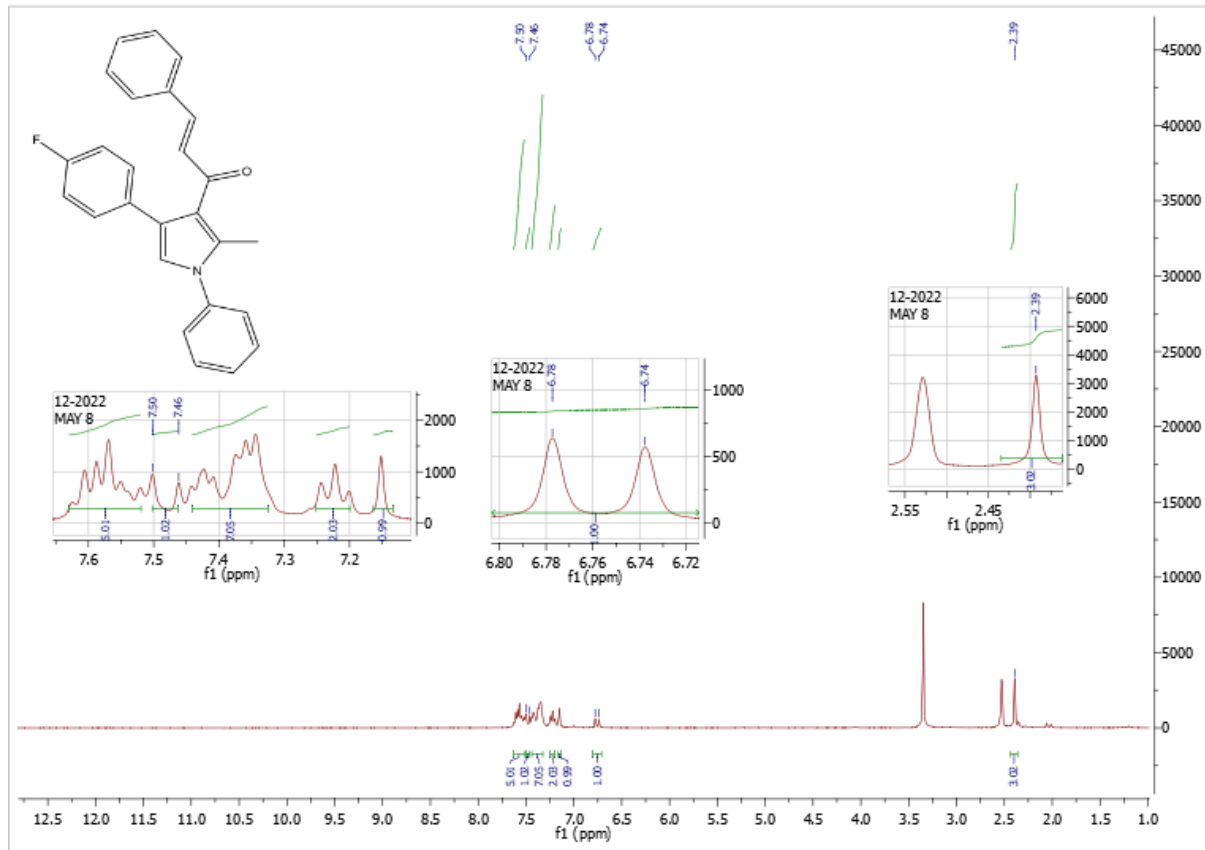




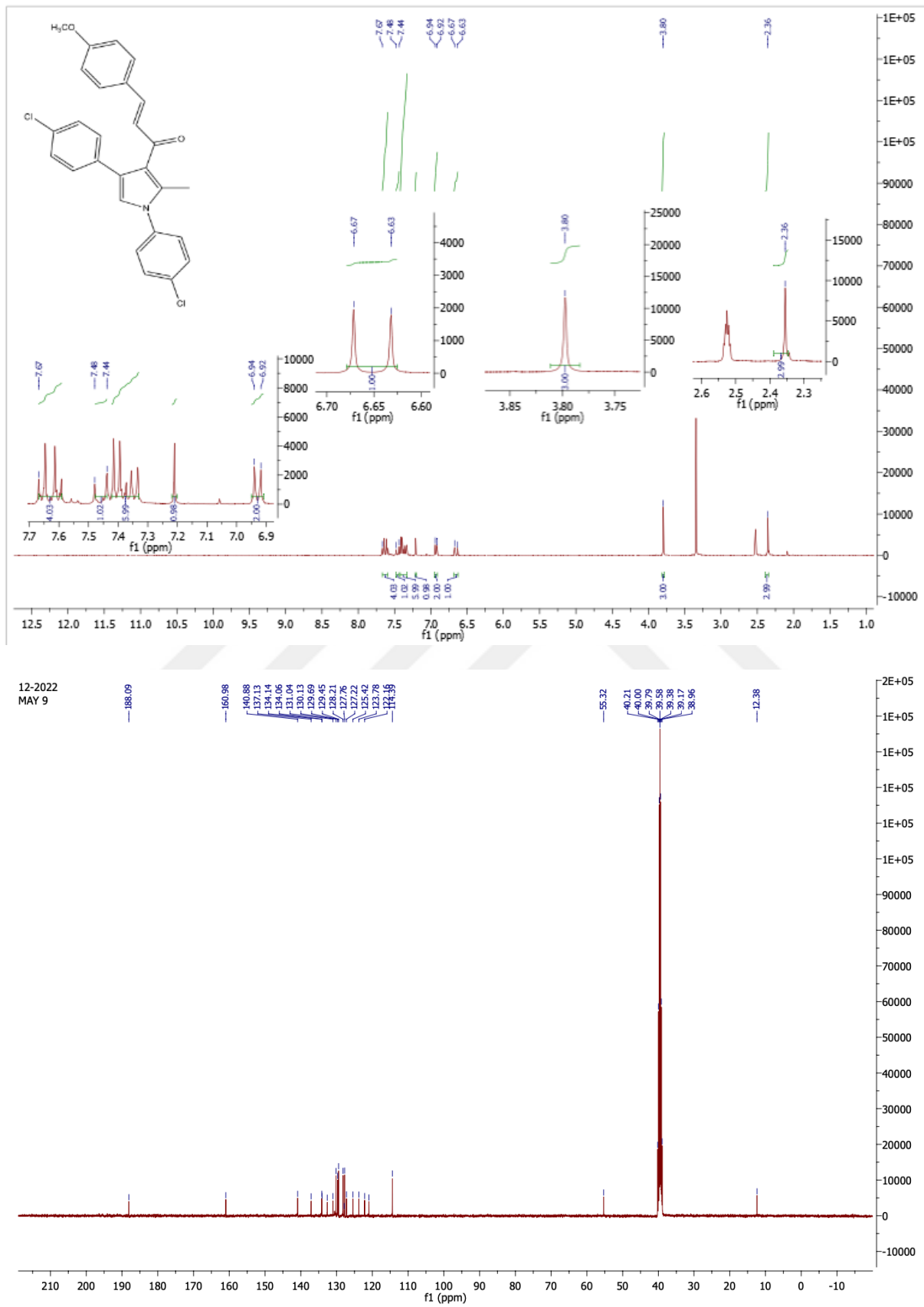
### MAY-6 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



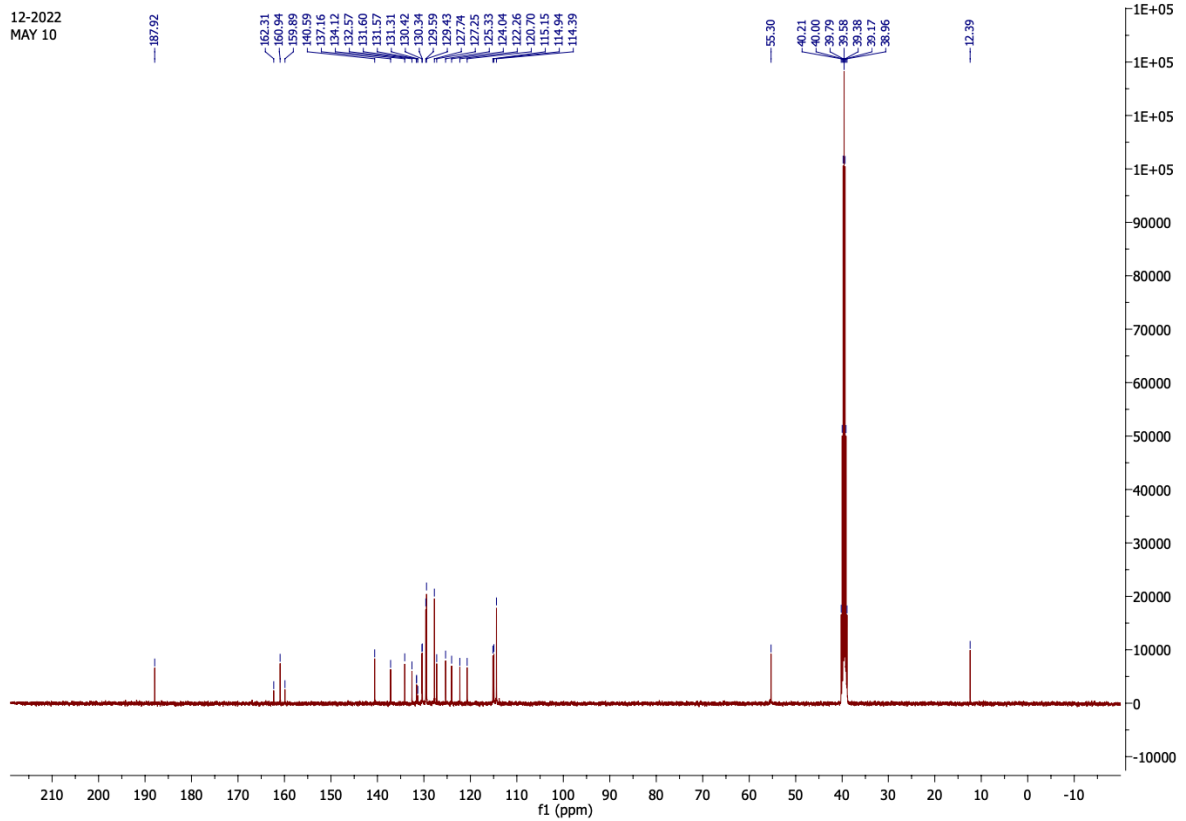
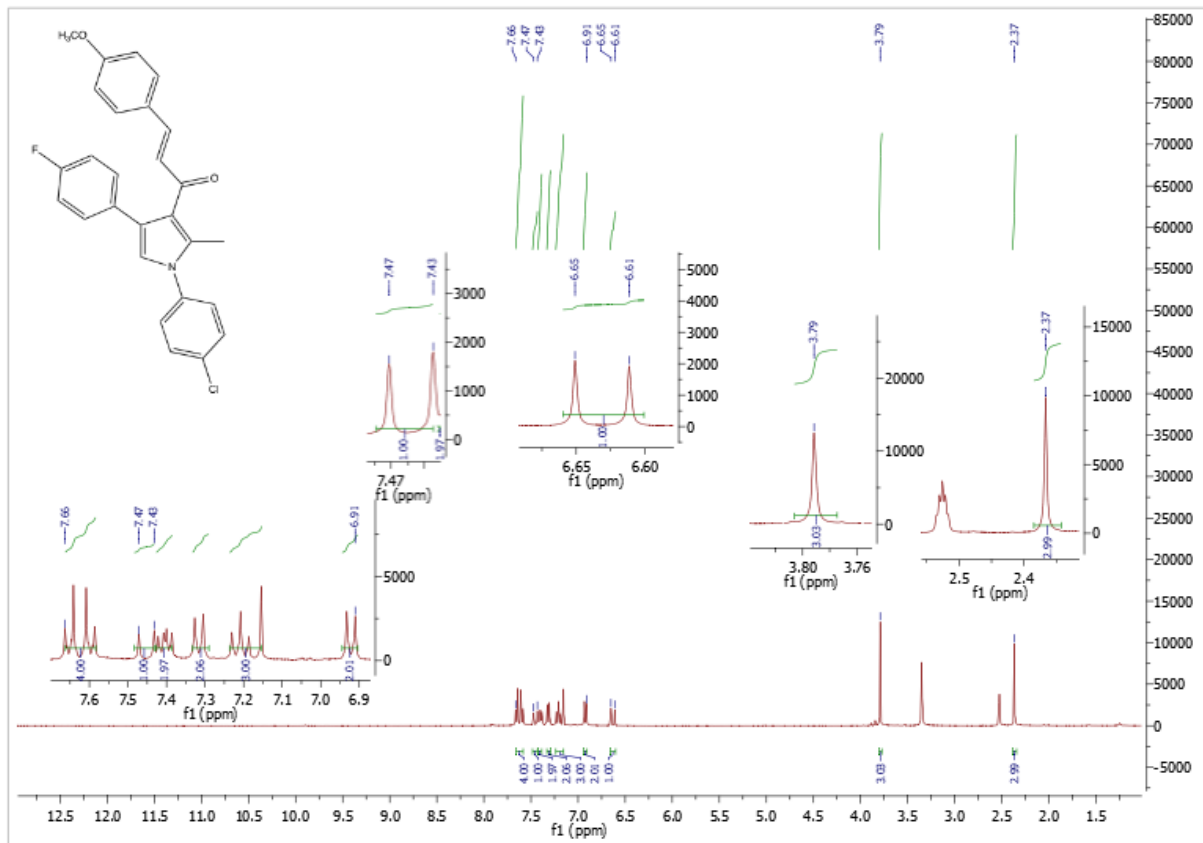
### MAY-8 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



### MAY-9 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları

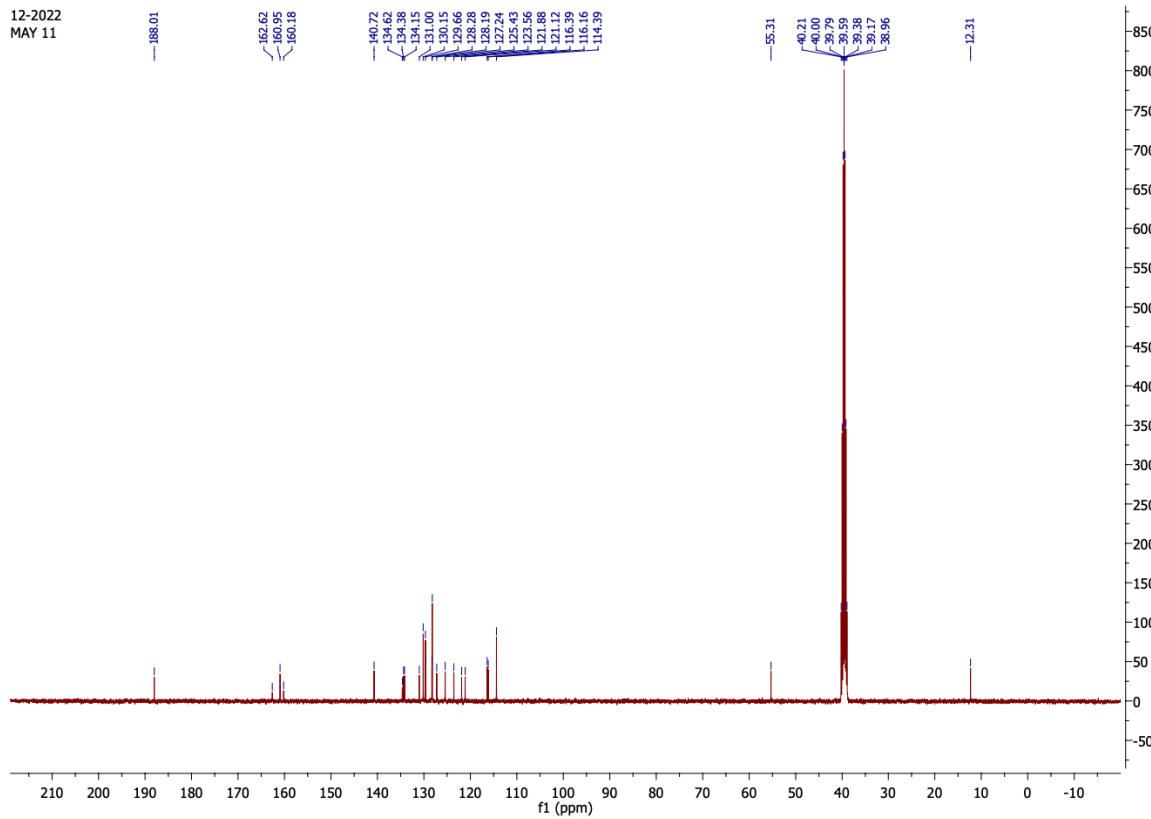
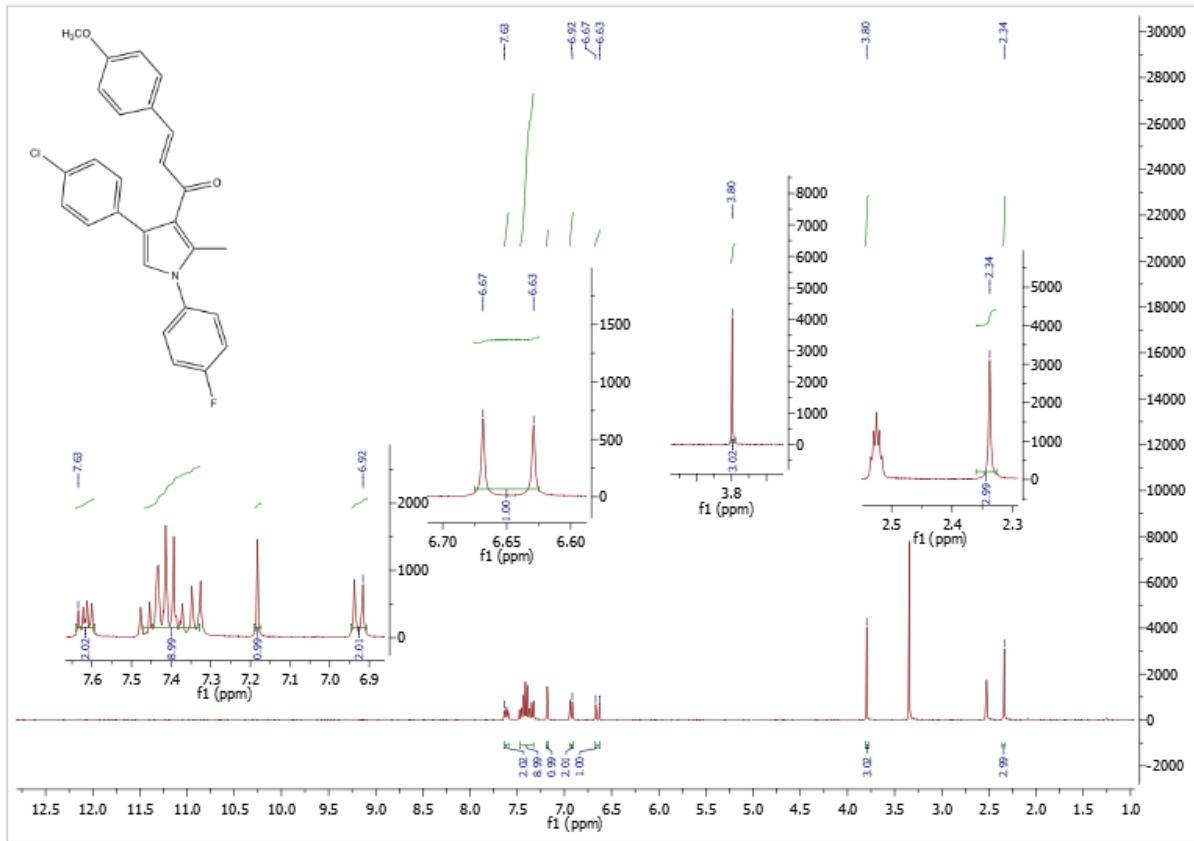


### MAY-10 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları

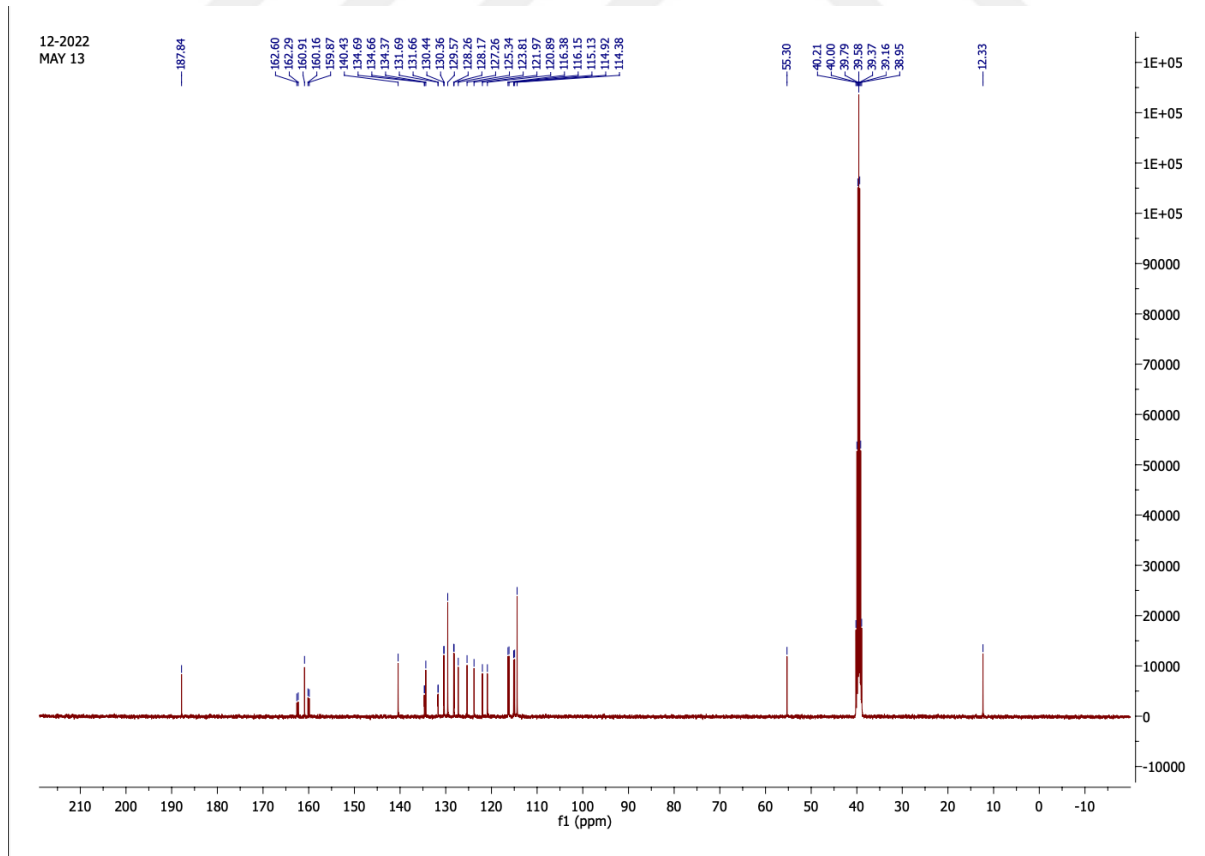
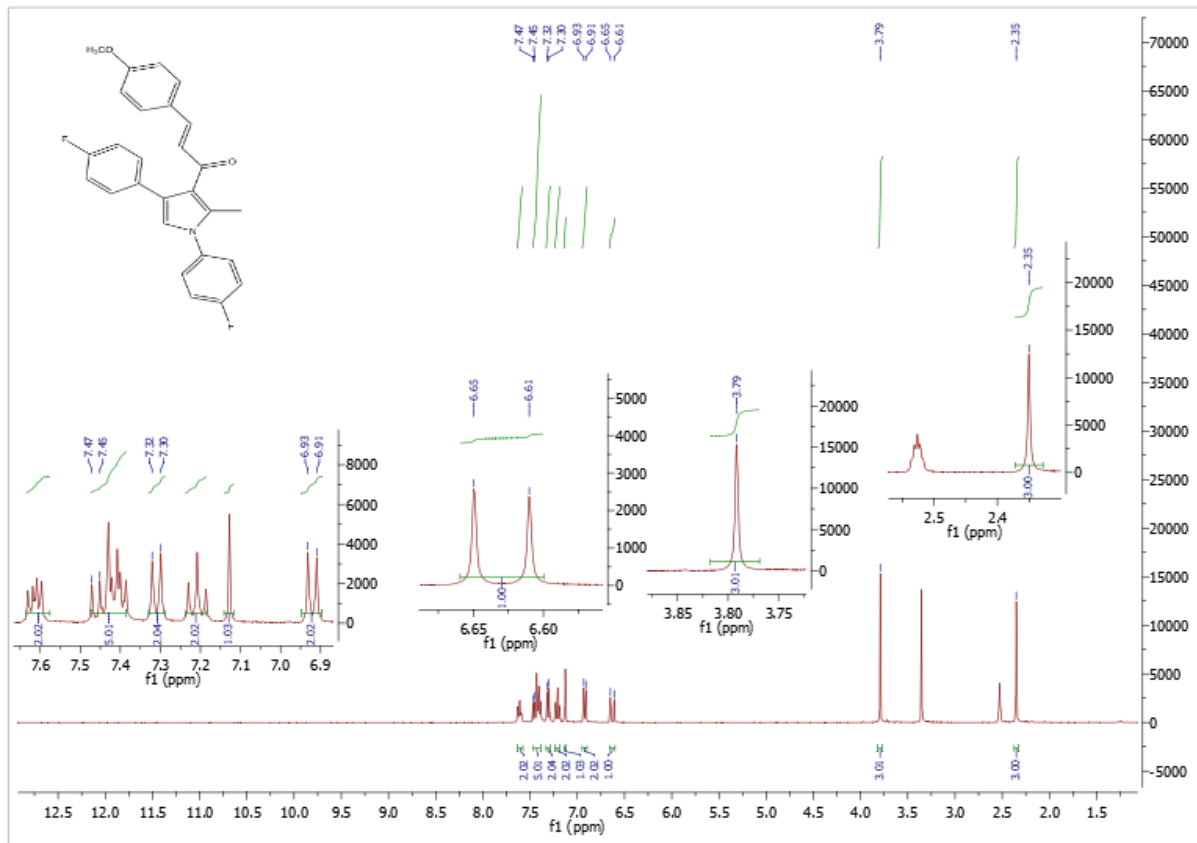




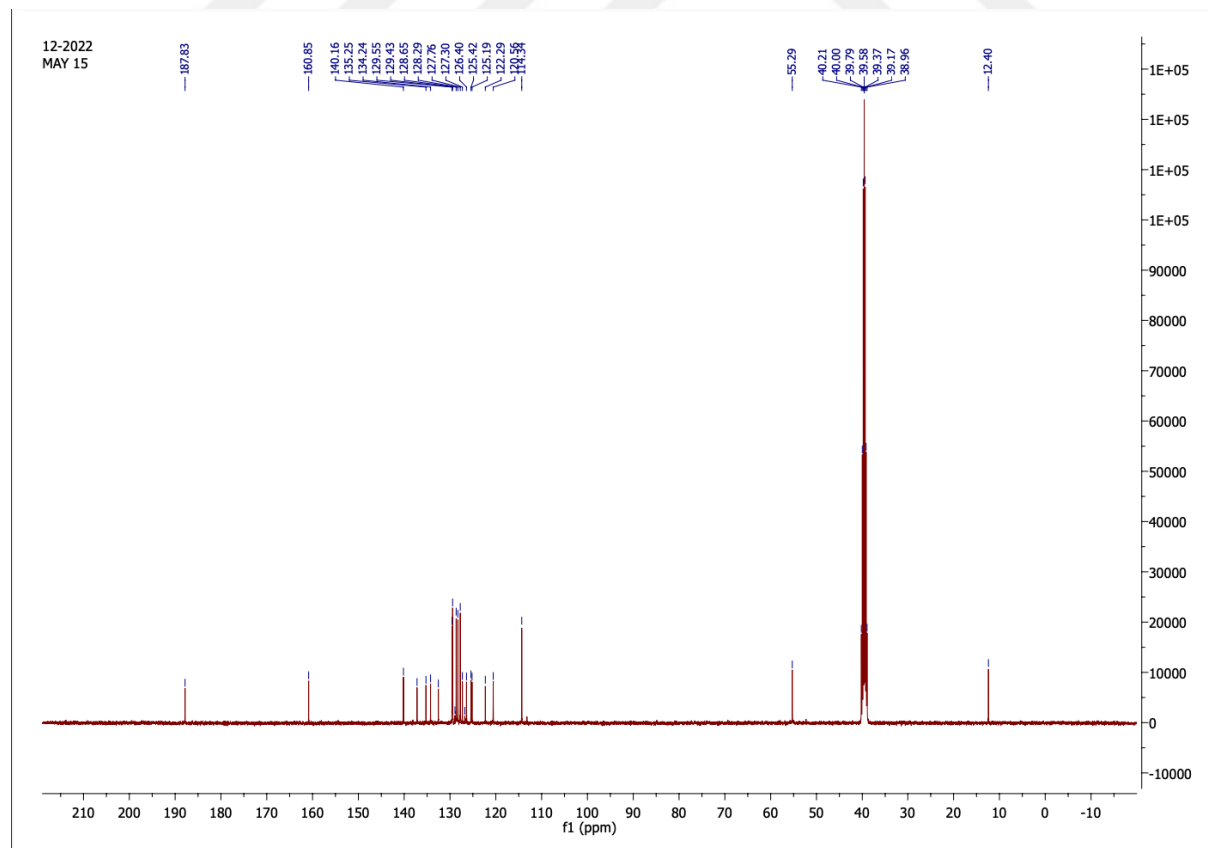
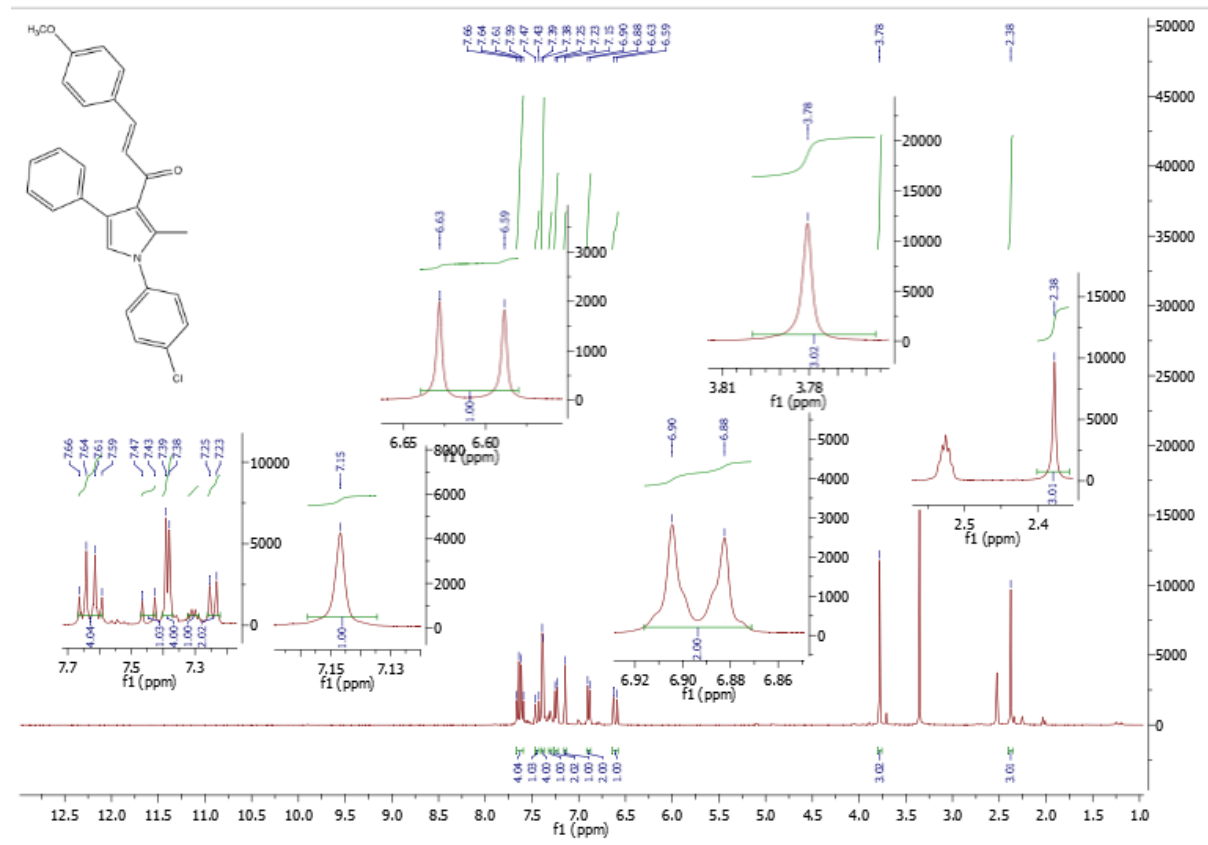
### MAY-11 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



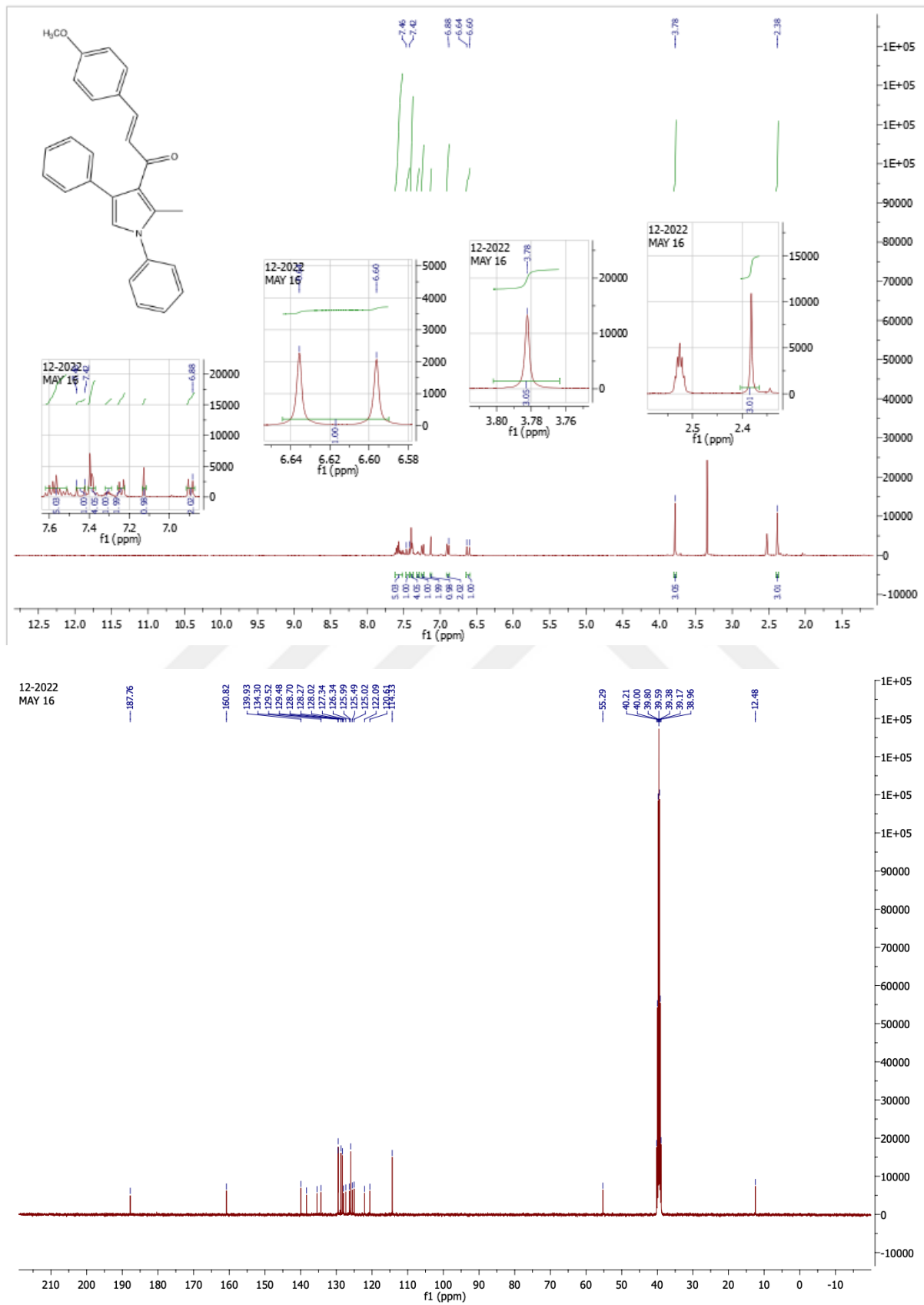
### MAY-13 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



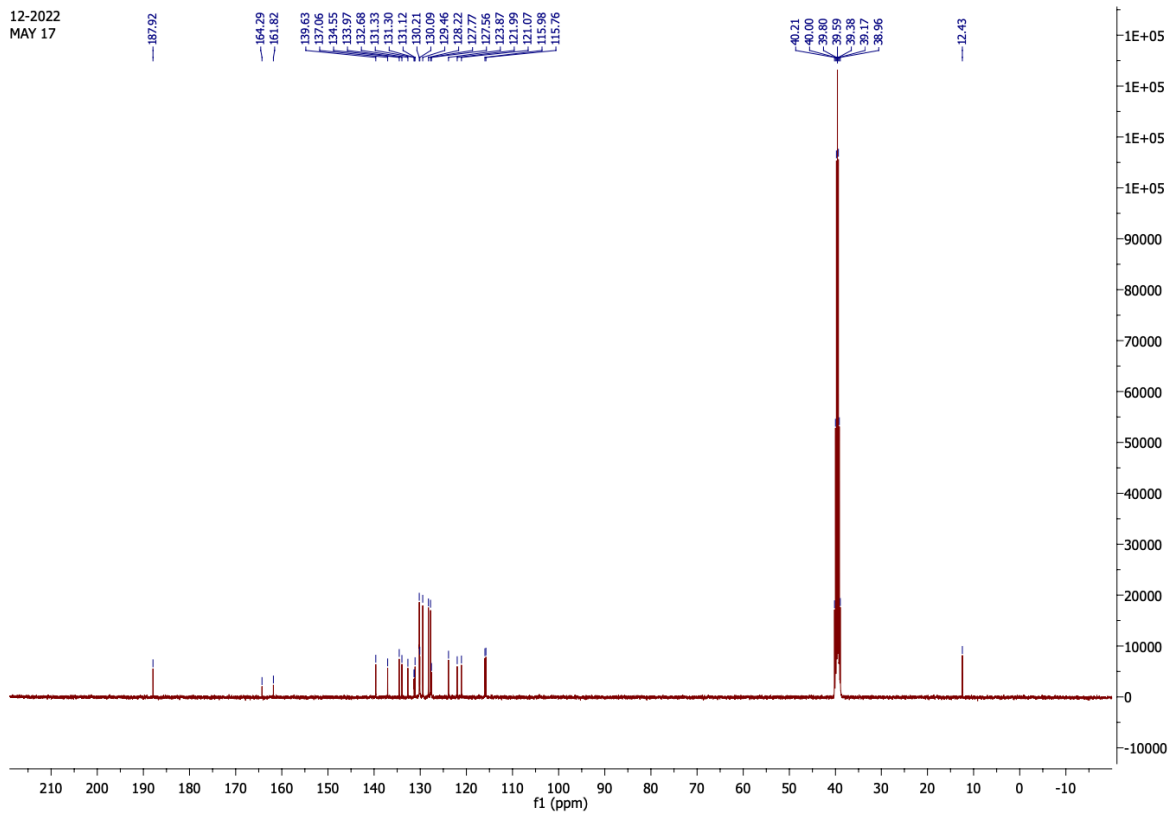
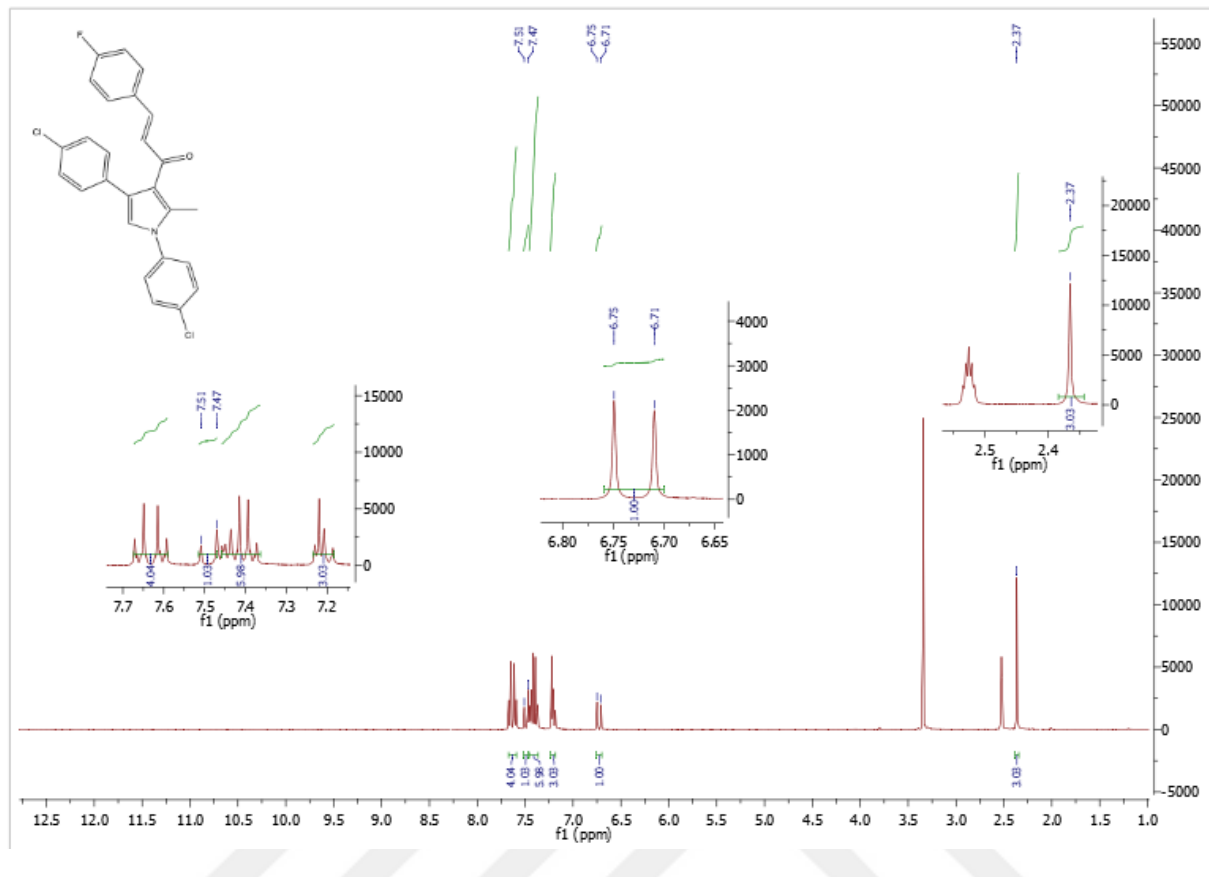
### MAY-15 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



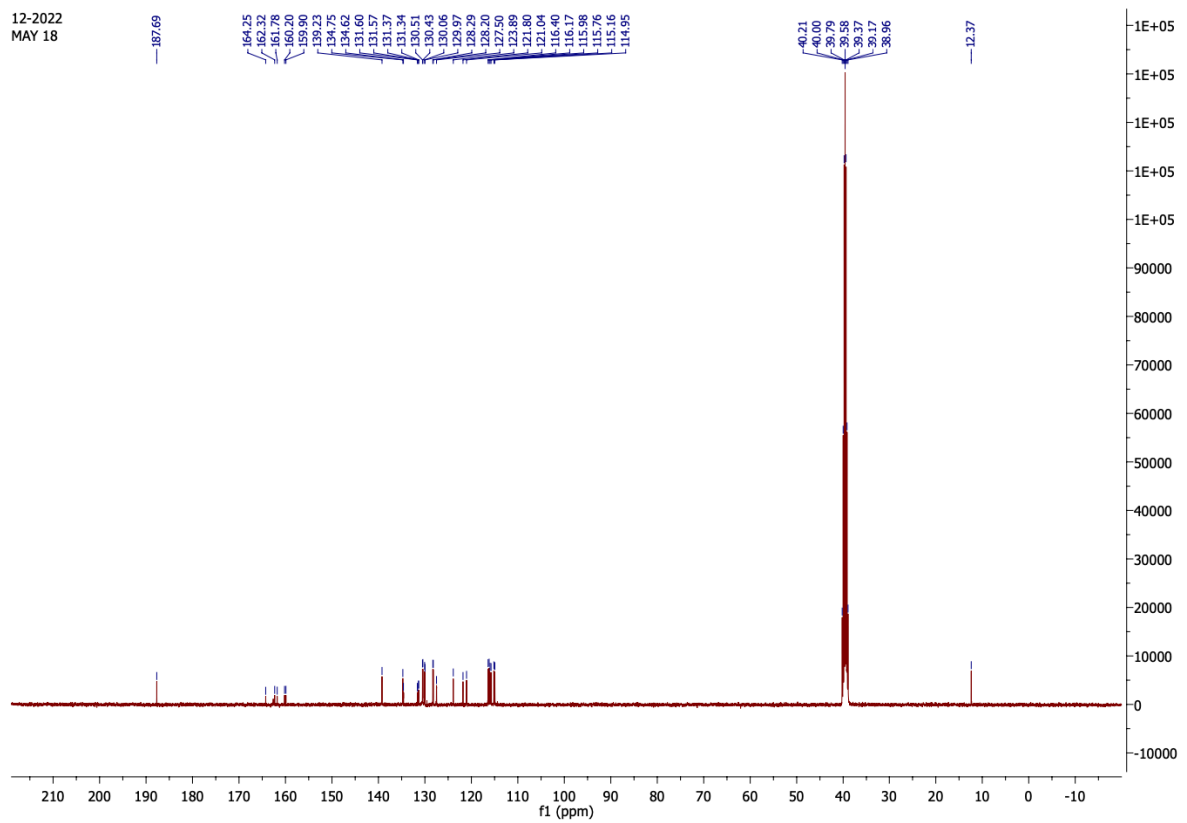
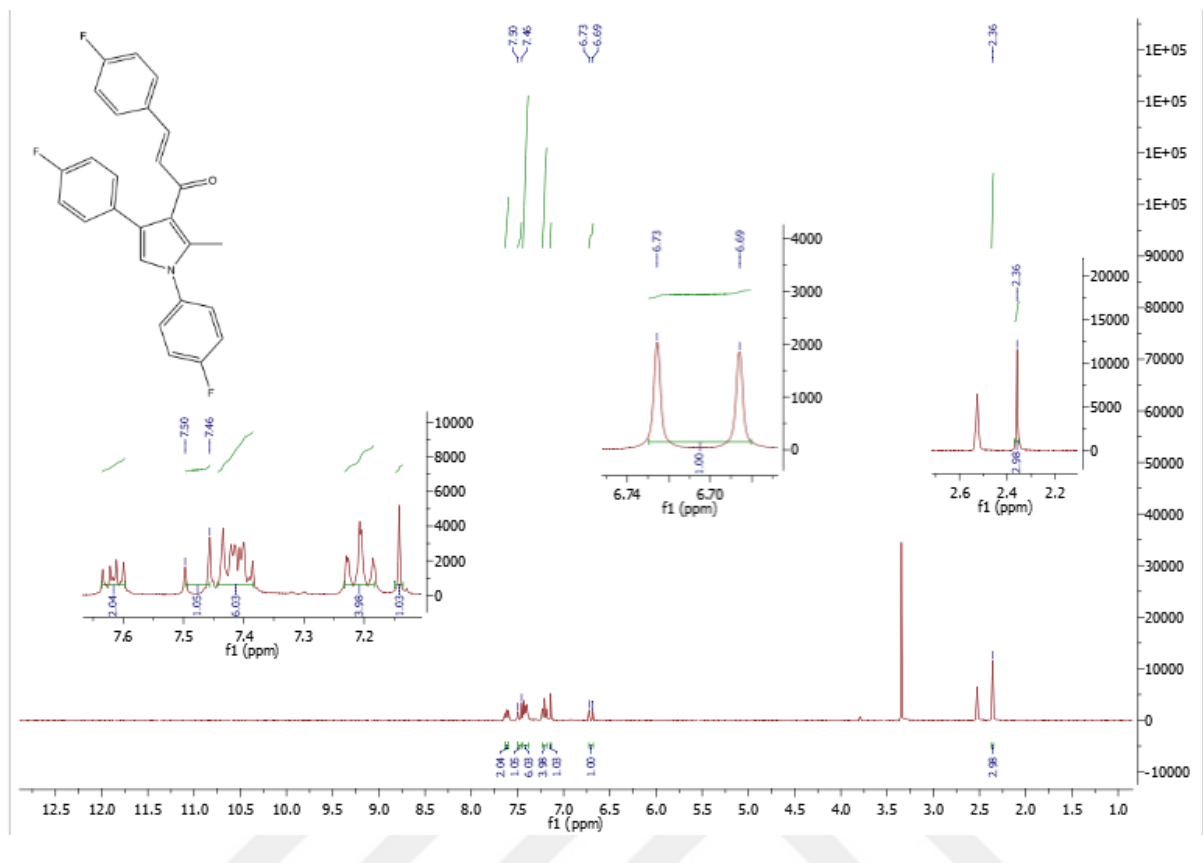
### MAY-16 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



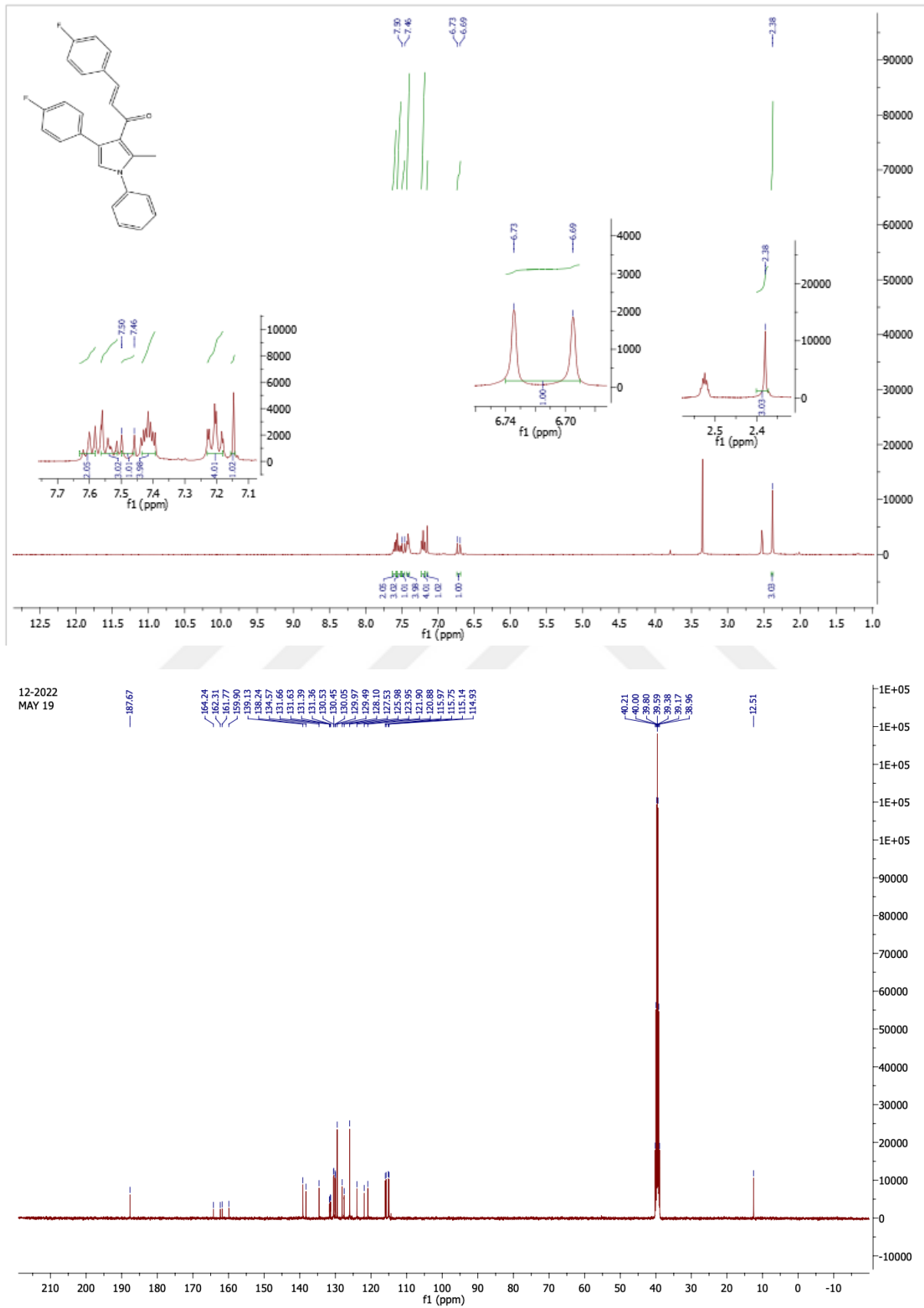
### MAY-17 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



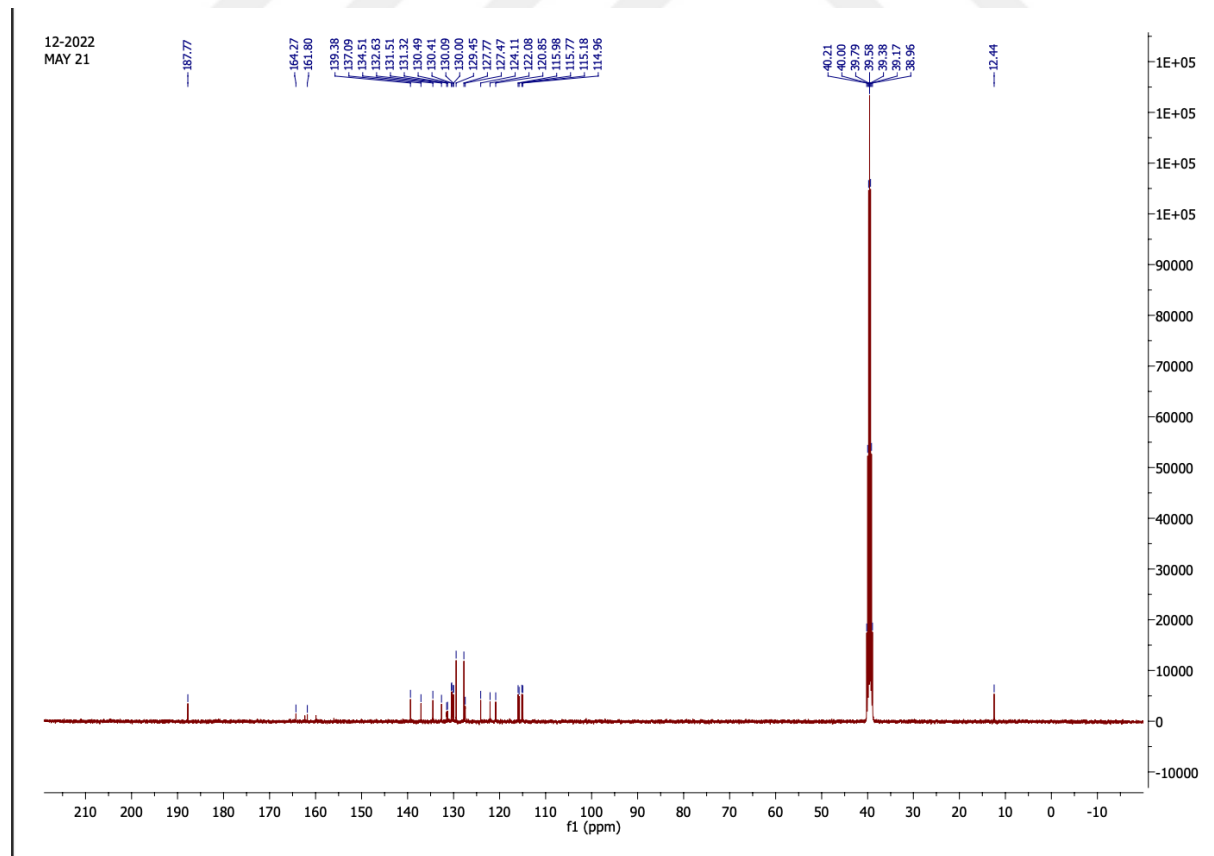
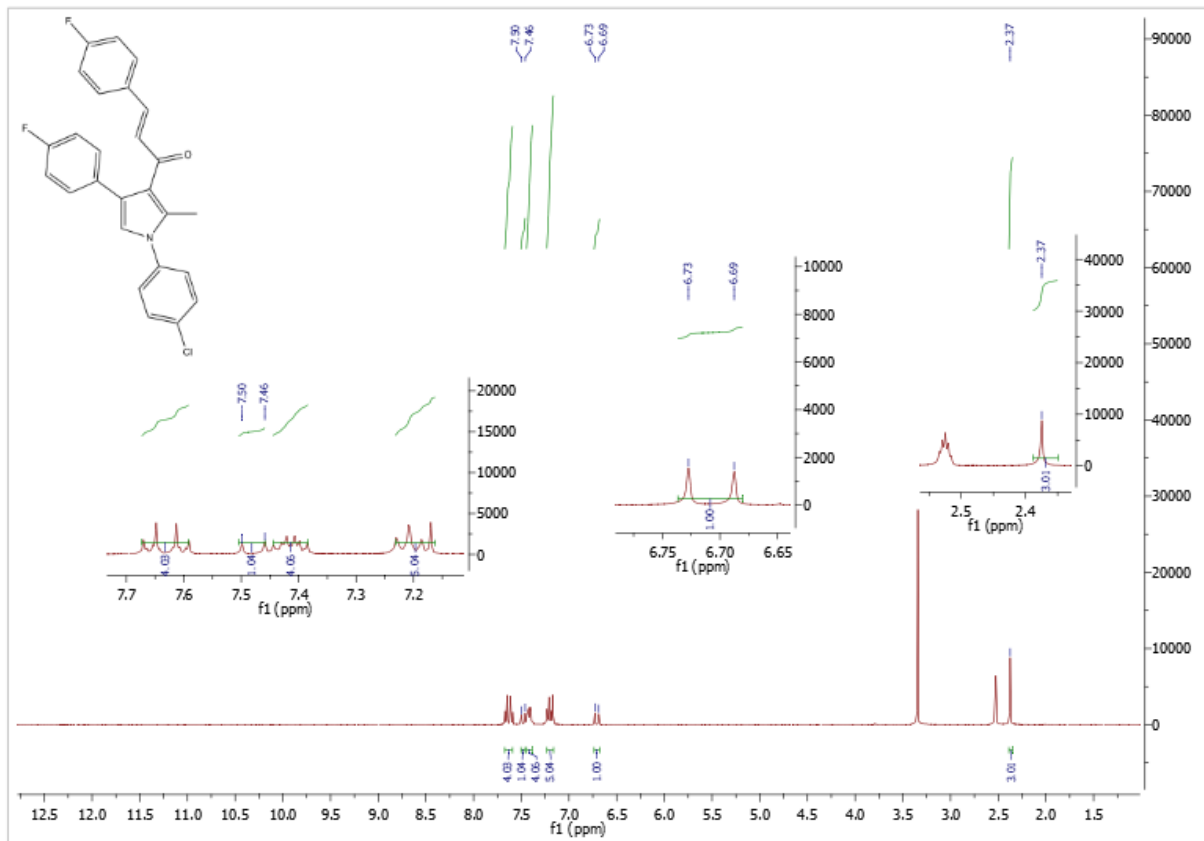
### MAY-18 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



### MAY-19 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları

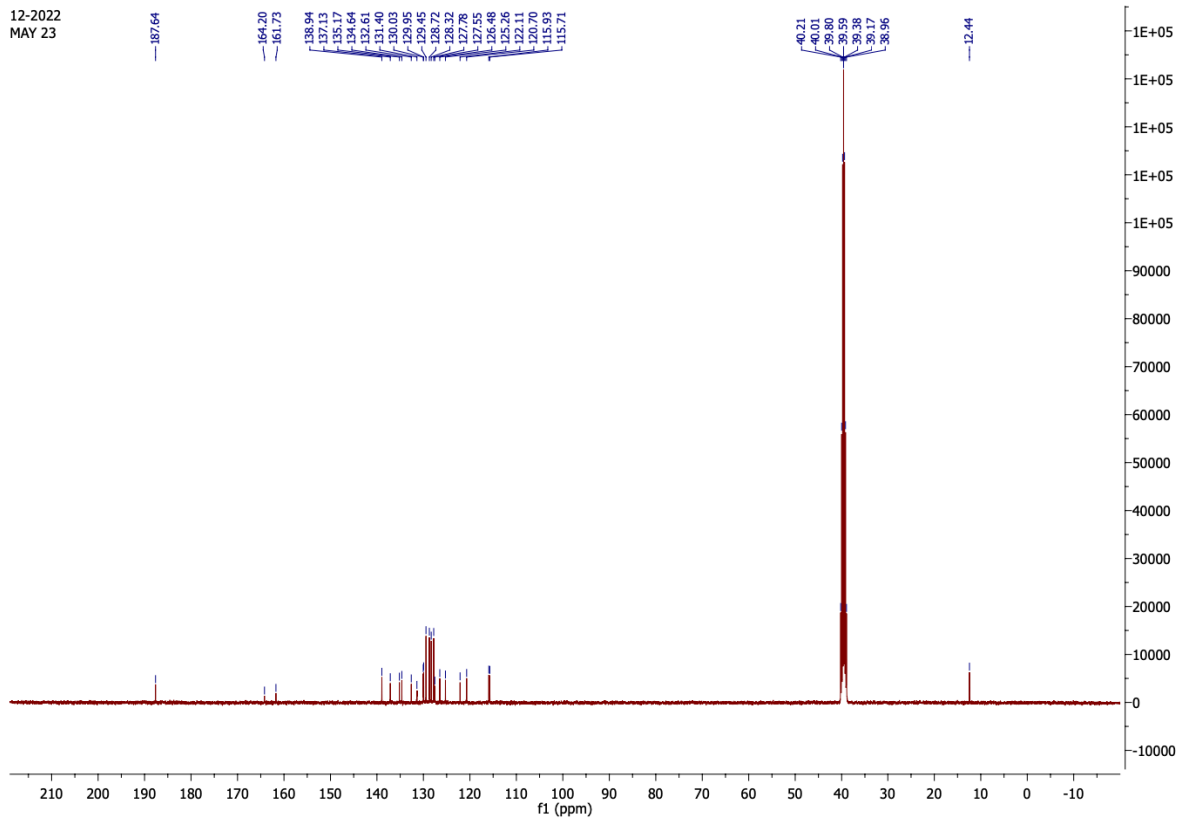
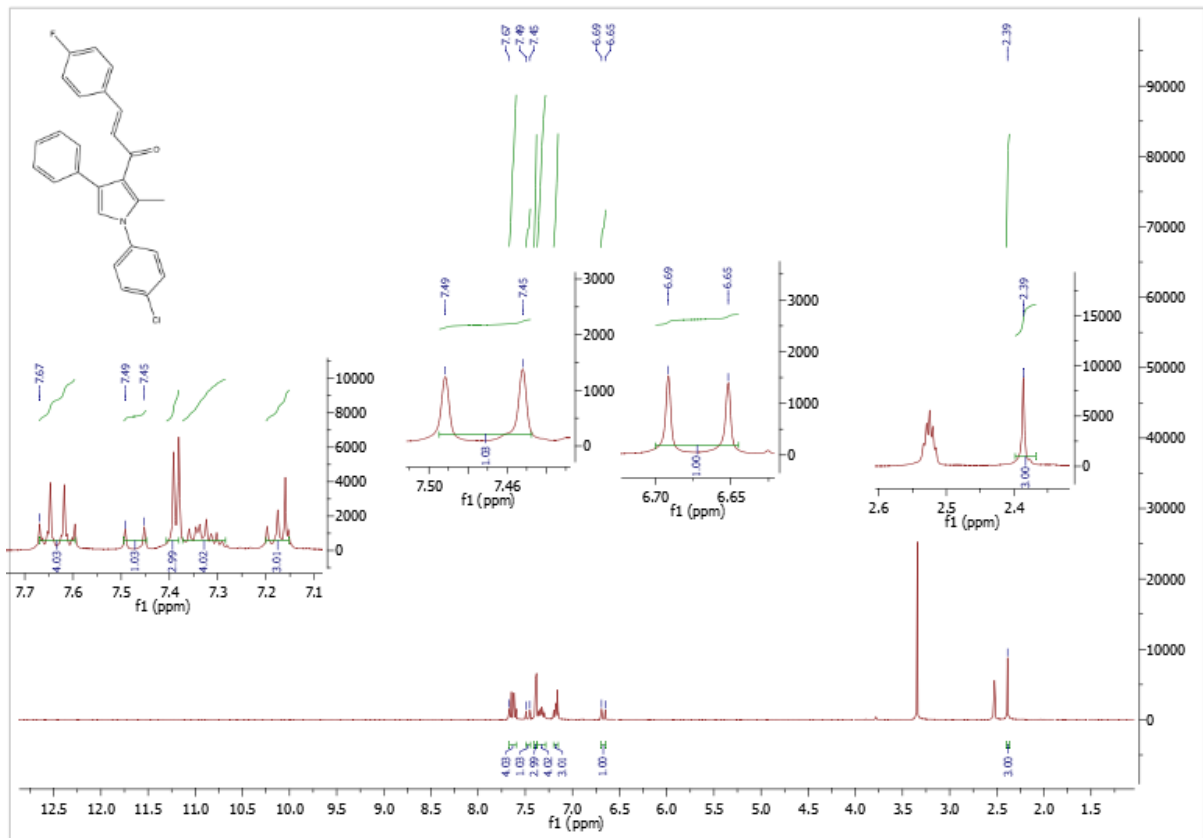


### MAY-21 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları

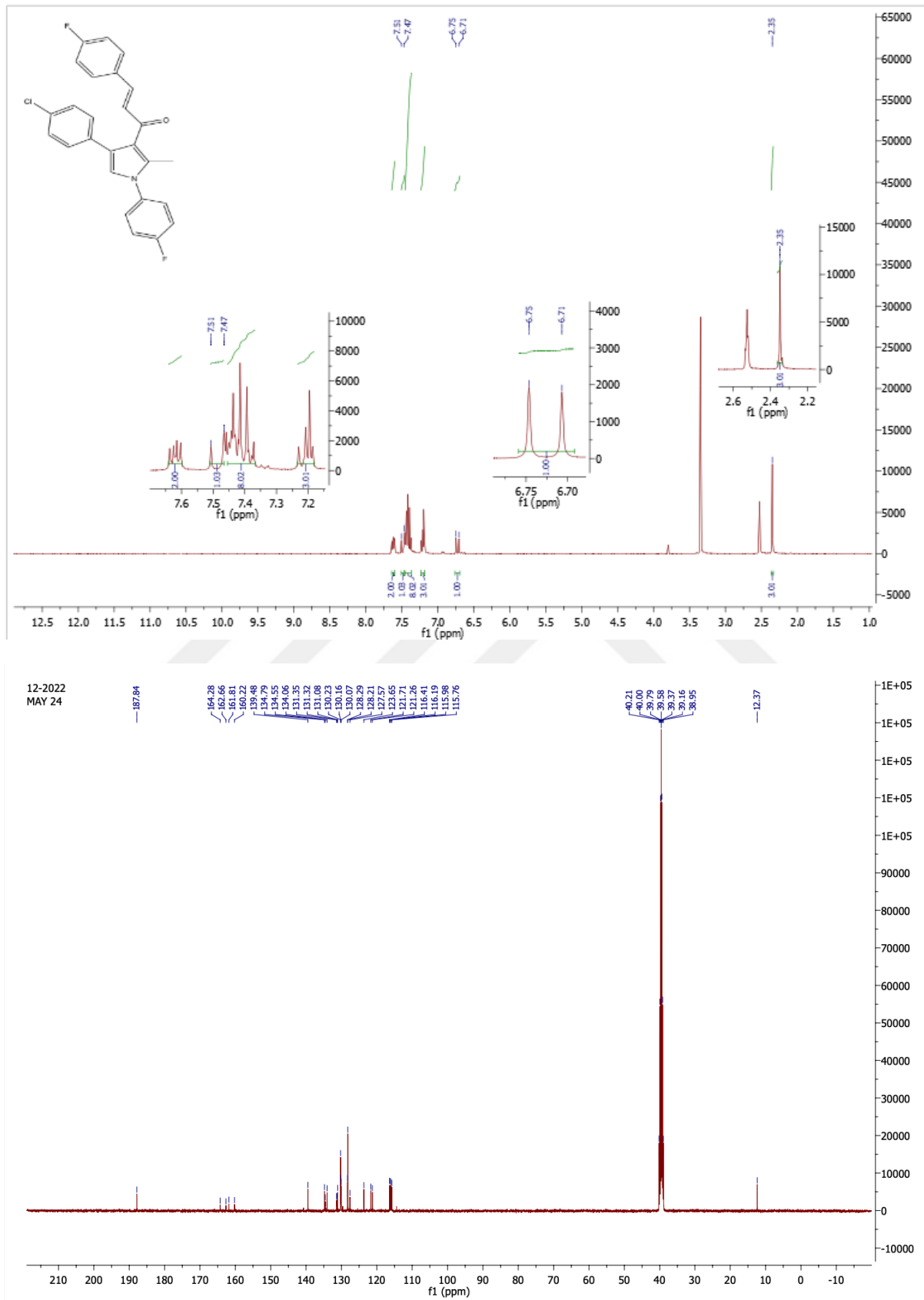




### MAY-23 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



### MAY-24 molekülüne ait $^1\text{H-NMR}$ ve $^{13}\text{C-NMR}$ spektrumları



## ÖZGEÇMİŞ

**Adı ve Soyadı :** Mehmet Ali YÜCEL

**Doğum Tarihi:** [REDACTED]

**E-mail :** [REDACTED]

### Öğrenim Durumu

Derece	Bölüm/Program	Üniversite	Yıl
Lisans	Eczacılık	Mersin Üniversitesi	2015-2020
Yüksek Lisans	Sağlık Bilimleri Enstitüsü/ Farmasötik Kimya Anabilim Dalı	Mersin Üniversitesi	2020- Devam etmektedir

### İş Deneyimi (Sondan geçmişe doğru sıralayınız)

Görevi	Kurum	Yıl
Araştırma Görevlisi	Erzincan Üniversitesi	2022- Devam etmektedir
Eczacı	Akgül Eczanesi	2020-2021

### YAYINLAR (Makaleler/Bildiriler/Ödüller)

#### Makaleler

1. Aktaş, Derya Anıl, et al. "Design, synthesis and biological evaluation of 3, 5-diaryl isoxazole derivatives as potential anticancer agents." *Bioorganic & Medicinal Chemistry Letters* 30.19 (2020): 127427.
2. Yucel, Mehmet Ali. "Comparing machine learning models for acetylcholine esterase inhibitors." *Bioorganic Medicinal Chemistry Reports* (2022).

#### Bildiriler

1. Beyza Şimşek, Dilek Akbaşlar, Tuğba Ertan Boelli, **Mehmet Ali Yücel**, Öztekin Algül., Synthesis and evaluation of in vitro antiproliferative activity some bisbenzazol derivatives as topoisomerase enzyme inhibitors., 8. International Drug Chemistry Conference, 27 Şubat -1 Mart 2020, Kemer, Antalya, Türkiye, pp.149.
2. **Mehmet Ali Yücel**, Öztekin Algül. Screening of  $\beta$ 2-Receptor Inhibitors with Machine Learning. 10. International Drug Chemistry Conference, 10-13 Mart 2022, Kemer, Antalya, Türkiye, pp.96-99.
3. **Mehmet Ali Yücel**, Öztekin Algül. 2.Computer-Aided Drug Design Symposium and Workshop. 11-12 Mayıs, İstanbul. Türkiye,pp. 64.

#### Ödüller

1. Mansiyon ödülü, Screening of  $\beta$ 2-Receptor Inhibitors with Machine Learning. 10. International Drug Chemistry Conference,