



**T.C.
AKSARAY ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

FİZİK ANABİLİM DALI

**YOĞUNLUK FONKSİYONEL TEORİ TEMELLİ KANTİTATİF
YAPI-ETKİ ANALİZLERİ (QSAR) : KUMARİN
MOLEKÜLLERİNİN KARBONİK ANHİDRAZ ENZİMİNE
KARŞI İNHİBİSYON ETKİSİNİN MODELLENMESİ**

YÜKSEK LİSANS TEZİ

Zekiye Eda BOYLUĞ

DANIŞMAN

Prof. Dr. Süleyman YILMAZ

AKSARAY, 2016



**T.C.
AKSARAY ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

FİZİK ANABİLİM DALI

**YOĞUNLUK FONKSİYONEL TEORİ TEMELLİ KANTİTATİF
YAPI-ETKİ ANALİZLERİ (QSAR) : KUMARİN
MOLEKÜLLERİNİN KARBONİK ANHİDRAZ ENZİMİNE
KARŞI İNHİBİSYON ETKİSİNİN MODELLENMESİ**

YÜKSEK LİSANS TEZİ

Zekiye Eda BOYLUĞ

**DANIŞMAN
Prof. Dr. Süleyman YILMAZ**

AKSARAY, 2016

DOĞRULUK BEYANI

Yüksek lisans tezi olarak sunduğum bu çalışmayı, bilimsel etik, ahlak ve geleneklere aykırı düşecek bir yol ve yardıma başvurmaksızın yazdığımı, yararlandığım eserlerin kaynakçada gösterilenlerden oluştuğunu ve bu eserleri her kullanışımında alıntı yaparak yararlandığımı belirtir; bunu şerefimle doğrularım.

Enstitü tarafından belli bir zamana bağlı olmaksızın, tezimle ilgili yaptığım bu beyana aykırı bir durumun saptanması durumunda, ortaya çıkacak tüm ahlaki ve hukuki sonuçlara katlanacağımı bildiririm.

Zekiye Eda BOYLUĞ

TEŞEKKÜR

Tez, emek ve sabır isteyen önemli çalışmalardan biridir. Bu çalışma süresince gerekli her türlü yardımı gösteren danışmanım Prof. Dr. Süleyman YILMAZ hocama; sorularımı, sorunlarımı dinleyen ve tez yazımı sürecine kadar benden desteğini esirgemeyen ikinci danışmanım Doç. Dr. Erol EROĞLU hocama ve bu çalışma süreci içerisinde birtakım yazılım ve programlardaki sorularımı sabırla cevaplayan Arş. Gör. Nuri YORULMAZ hocama teşekkür ederim.

Ayrıca bu tez çalışması kapsamında 114Z045 nolu “Karbonik Anhidraz İnhibitörü Kumarin Bileşiklerinin Etki Mekanizmalarının Moleküler Modelleme Teknikleri ile Araştırılması ve Biyolojik Etkisi Yüksek Yeni Kumarin Türevlerinin Tasarımı” projesi ile finansal desteğinden dolayı TÜBİTAK’a teşekkür ederim.

İÇİNDEKİLER

TEŞEKKÜR	i
İÇİNDEKİLER	ii
ÖZET.....	iii
ABSTRACT	iv
ŞEKİLLER DİZİNİ	v
ÇİZELGELER DİZİNİ.....	vii
KISALTMALAR DİZİNİ	ix
1.GİRİŞ	1
2.LİTERATÜR ÖZETİ	7
2.1. Giriş.....	7
2.2. Karbonik Anhidraz Enzimi İnhibisyonu QSAR Çalışmaları.....	8
3. GEREÇ VE YÖNTEM.....	14
3.1. Giriş.....	14
3.2. Karbonik Anhidraz.....	15
3.3. Karbonik Anhidraz İnhibitörü Kumarinler	20
3.4. QSAR (Kantitatif Yapı-Etki İlişkisi) Analizleri	25
3.5 Çoklu Doğrusal Regresyon (ÇDR) Analizi	26
3.6.Moleküler Tanımlayıcı Hesaplama Yoğunluk Fonksiyonel Teorisi	30
3.7. Moleküler Tanımlayıcılar	42
3.8. QSAR Modellerinin Uygulanabilirlik Sınırlarının Tanımlanması (AD)	50
3.9 Kullanılan Donanım ve Yazılımlar	52
4. BULGULAR VE TARTIŞMA.....	54
4.1 Giriş.....	54
4.2 Karbonik Anhidraz II İzoenzimi QSAR Modelleri.....	55
4.3 Karbonik Anhidraz IX İzoenzimi QSAR Modelleri	81
4.4. QSAR Modelleri İle Seçiciliği Yüksek Sanal Kumarin Türevleri.....	104
5. SONUÇ VE ÖNERİLER.....	106
6. KAYNAKLAR	107
ÖZGEÇMİŞ.....	116

ÖZET

YOĞUNLUK FONKSİYONEL TEORİ TEMELLİ KANTİTATİF YAPI-ETKİ ANALİZLERİ (QSAR) : KUMARİN MOLEKÜLLERİNİN KARBONİK ANHİDRAZ ENZİMİNE KARŞI İNHİBİSYON ETKİSİNİN MODELLENMESİ

Bu tez çalışmasında, 50 den fazla kumarin, kükürt içeren kumarin ve kükürt içermeyen kumarin türevlerinin Karbonik anhidraz izoformlarından IX (CAIX) ve II (CA II) ye karşı deneysel inhibisyon katsayısı (K_i) ve hesaplanmış moleküler tanımlayıcıları arasındaki korelasyon ile DFT (Yoğunluk Fonksiyonel Teorisi) bazlı QSAR(Kantitatif Yapı-Etki Analizi İlişkileri) modelleri elde edildi. Kumarinlerin 3 boyutlu yapıları Gaussian 09 yazılımı ile hesaplandı. Oluşturulan bu yapılar Codessa 3 programına yüklendi ve pek çok sayıda moleküler tanımlayıcı hesaplanarak deneysel olarak ölçülmüş biyolojik aktivite değerleri ile arasındaki korelasyon Çoklu Doğrusal Regresyon Analizi ile sağlandı.

CAIX için kükürt içeren kumarin modeli, fiziksel olarak yorumlanabilir moleküler tanımlayıcılar içeren istatistiksel olarak önemli bir küresel model olan 3 parametrelili lineer denklemin istatistiksel değerleri $N=17$, $R^2=0.943$, $Q^2=0.872$, $F=44.586$ ve $R_{ext}^2=0.993$ 'dir.

Anahtar Kelimeler: Kumarin, Karbonik Anhidraz, DFT, QSAR, Çoklu Doğrusal Regresyon Analizi.

ABSTRACT

DENSITY FUNCTIONAL THEORY BASED QUANTITATIVE STRUCTURE ACTIVITY RELATIONSHIP (QSAR): MODELING OF THE INHIBITORY EFFECT AGAINST THE ENZYME CARBONIC ANYDRASE MOLECULES OF COUMARIN

In this study present DFT (Density Functional Theory) based QSAR (Quantitative Structure Activity Relationship) models that demonstrate the correlations between calculated molecular descriptors and the experimental inhibition constant (K_i) data of more than fifty derivatives of coumarin, sulfo coumarin and nonsulfo coumarin measured against Carbonic anhydrase isoforms IX (CA IX) and II (CA II). 3D structures of coumarins have been calculated by Gaussian 09. These structures, was loaded into the Codessa 3 programme, and many molecular descriptor values calculated and experimentally measured the correlation between biological activity with Multi Linear Regression Analysis was performed.

The sulfo coumarin model for CA IX, statistically the most significant one which comprises physically interpretable molecular descriptors is a five-parameter linear equation with the statistical values $N=17$, $R^2=0.943$, $Q^2=0.872$, $F=44.586$ ve $R_{ext}^2=0.99$.

Keywords: Coumarin, Carbonic Anyhrase, DFT, QSAR, Multi Linear Regression Analysis.

ŞEKİLLER DİZİNİ

Sayfa

Şekil 3.1: CA II- ((2Z) -3- {2-hydroxy-5- [(1S) -1-hydroxy-3-methylbutyl]-4-methoxyphenyl} prop-2-enoic asit kompleksi (PDB kod: 3F8E).	16
Şekil 3.2: CA II aktif sitesi içerisindeki aminoasitlerin hidrolize kumarin türevi ile etkileşimleri.	17
Şekil 3.3: PDB’de bulunan üç kumarin bileşiğinin üst üste görüntüsü.	17
Şekil 3.4: CA IX 5- (1-naphthalen-1-yl-1,2,3-triazol-4-yl) thiophene-2 sulfonamide kompleksi (PDB kod:5FL4).	18
Şekil 3.5: CA IX dimer modeli.....	19
Şekil 3.6: CA izoenzimlerinin CO ₂ hidrasyonu katalitik mekanizması.	20
Şekil 3.7: 82 adet Karbonik anhidraz inhibitörü kumarin türevinin yapısı.	21
Şekil 3.8: Şekil 4.7’de verilen 1 no’luk kumarin türevinin karboksilik asit formu.	25
Şekil 4.1: CA II kapalı form tüm kumarinler QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	60
Şekil 4.2: CA II kapalı form tüm kumarinler QSAR modeli sapan değer tayini.	61
Şekil 4.3: CA II kapalı form kükürt içeren kumarinler QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	65
Şekil 4.4: CA II kapalı form kükürt içeren kumarinler QSAR modeli sapan değer tayini.	65
Şekil 4.5: CA II kapalı form kükürt içermeyen kumarinler QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	68
Şekil 4.6: CA II kapalı form kükürt içermeyen kumarinler QSAR modeli uygulanabilirlik sınırları ile sapan değer tayini.	69
Şekil 4.7: CA II karboksilik asit (açık form) tüm kumarinler QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	73
Şekil 4.8: CA II karboksilik asit (açık form) tüm kumarinler QSAR modeli sapan değer tayini.	73
Şekil 4.9: CA II karboksilik asit (açık form) kükürt içeren moleküller QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	77
Şekil 4.10: CA II karboksilik asit (açık form) kükürt içeren kumarinler QSAR modeli sapan değer tayini.	77
Şekil 4.11: CA II karboksilik asit (açık form) kükürt içermeyen moleküller QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	81
Şekil 4.12: CA II karboksilik asit (açık form) kükürt içermeyen moleküller QSAR modeli sapan değer tayini.	81
Şekil 4.13: CA IX kapalı form tüm kumarinler QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği.	85
Şekil 4.14: CA IX kapalı form tüm kumarinler QSAR modeli sapan değer tayini	86
Şekil 4.15: CA IX kapalı form kükürt içeren kumarinler QSAR modeli deneysel K _i ’nin hesaplanan K _i değerine karşı grafiği	89

Şekil 4.16:. CA IX kapalı form kükürt içeren kumarinler QSAR modeli sapan değer tayini.....	89
Şekil 4.17: CAIX kapalı form kükürt içermeyen kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği.	93
Şekil 4.18: CA IX kapalı form kükürt içermeyen kumarinler QSAR modeli sapan değer tayini.	93
Şekil 4.19: CAIX karboksilik asit formu tüm kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği.	97
Şekil 4.20: CAIX karboksilik asit formu tüm kumarinler QSAR modeli sapan değer tayini.	97
Şekil 4.21: CAIX karboksilik asit formu kükürt içeren kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği.	100
Şekil 4.22: CAIX karboksilik asit formu kükürt içeren kumarinler QSAR modeli sapan değer tayini.	101
Şekil 4.23: CAIX karboksilik asit formu kükürt içermeyen kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği.	101
Şekil 4.24: CAIX karboksilik asit formu kükürt içermeyen kumarinler QSAR modeli sapan değer tayini.	104

ÇİZELGELER DİZİNİ

Sayfa

Çizelge 3.1: Yapıları şekil 4. 7’de verilen kumarin türevlerinin deneysel ölçülmüş inhibisyon katsayısı verileri.	23
Çizelge 4.1: CA II İzoenzimi kapalı form tüm kumarinler QSAR modeli.	57
Çizelge 4.2.: CA II İzoenzimi kapalı form tüm kumarinler QSAR modeli tahmin çizelgesi.	59
Çizelge 4.3: CA II İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli.....	63
Çizelge 4.4: CA II İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli tahmin çizelgesi.	64
Çizelge 4.5: CA II İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli.	67
Çizelge 4.6: CA II İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi.	68
Çizelge 4.7: CA II İzoenzimi açık (karboksilik asit) formu tüm kumarinler QSAR modeli.	70
Çizelge 4.8: CA II İzoenzimi açık form (karboksilik asit) tüm kumarinler QSAR modeli tahmin çizelgesi.	72
Çizelge 4.9: CA II İzoenzimi karboksilik asit (açık form) kükürt içeren kumarinler QSAR modeli.	75
Çizelge 4.10: CA II İzoenzimi karboksilik asit (açık form) kükürt içeren kumarinler QSAR modeli tahmin çizelgesi.	76
Çizelge 4.11: CA II İzoenzimi karboksilik asit (açık form) kükürt içermeyen kumarinler QSAR modeli.	79
Çizelge 4.12: CA II İzoenzimi açık form (karboksilik asit) kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi.	80
Çizelge 4.13: CA IX İzoenzimi kapalı form tüm kumarinler global QSAR modeli.	83
Çizelge 4.14: CA IX İzoenzimi kapalı form tüm kumarinler QSAR modeli tahmin çizelgesi.	84
Çizelge 4.15: CA IX İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli.	87
Çizelge 4.16: CA IX İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli tahmin çizelgesi.	88
Çizelge 4.17: CA IX İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli.	91
Çizelge 4.18: CA IX İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi.	92
Çizelge 4.19: CA IX İzoenzimi karboksilik asit tüm kumarinler QSAR modeli. ..	95
Çizelge 4.20: CA IX İzoenzimi karboksilik asit tüm kumarinler QSAR modeli tahmin çizelgesi.	96

	<u>Sayfa</u>
Çizelge 4.21: CA IX İzoenzimi karboksilik asit kükürt içeren kumarinler QSAR modeli.	99
Çizelge 4.22: CA IX İzoenzimi karboksilik asit kükürt içeren kumarinler QSAR modeli tahmin çizelgesi.	100
Çizelge 4.23: CA IX İzoenzimi karboksilik asit kükürt içermeyen kumarinler QSAR modeli.	102
Çizelge 4.24: CA IX İzoenzimi karboksilik asit kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi.	103
Çizelge 4.25: QSAR modelleri ile tasarlanan sanal kumarin türevleri.	106



KISALTMALAR DİZİNİ

ACE	Conditional Expectations Technique
AD	Application Domain
BMLR	En İyi Çoklu Doğrusal Regresyon
CA	Karbonik Anhidraz
CA II	Karbonik Anhidraz II İzoenzimi
CA IX	Karbonik Anhidraz IX İzoenzimi
CNDO/2	Complete Neglect of Differential Overlap/2
CoMFA	Comparative Molecular Field Analysis
CoMSIA	Comparative Molecular Similarity Index Analysis
ÇDR	Çoklu Doğrusal Regresyon
DFT	Yoğunluk Fonksiyonel Teorisi
ESS	Açıklanan Kareler Toplamı
FOPA	Frontier Orbitals Phase Angles
HOMO	En Yüksek Dolu Moleküler Orbital
KEKKR	Kısmi En Küçük Kareler Regresyonu
K_i	İnhibisyon Sabiti
LUMO	En Düşük Dolu Olmayan Moleküler Orbital
PDB	Protein Data Bank
PLS	Kısmi Doğrusal Kareler Metodu
PRESS	Tahmin Edilen Residüel Kareler Toplamı
RSS	Rezidüel Kareler Toplamı
SCF	Öz Tutarlılık Kapsamı
TSS	Toplam Kareler Toplam

1. GİRİŞ

İlaçların çoğunun hedefi hücre proteinleridir. İlaç hücre içerisine ulaştığında ilgili protein ile seçici şekilde etkileşimi neticesinde tedavi veya teşhis etkisi ortaya çıkar. 2002 yılında tamamlanan insan genoması çalışmaları sonuçlarına göre, insan geni tarafından üretilen 6000 ile 8000 arasında farmakolojik açıdan ilaç hedefi olabilecek proteinin varlığı tahmin edilmiştir (Landry ve Gies, 2008). 2003 yılında yapılan bir çalışma insan genoması çalışmalarında ortaya konulan proteinler arasından sadece 273 proteinin o günkü mevcut ilaçların hedefi olduğunu göstermiştir (Golden, 2003). Günümüzde 330 civarında ilaç hedefi protein vardır, bunlardan 270 kadarı insan genoması tarafından üretilen proteinler iken geriye kalanları ise patojen organizmalar (bakteri, virüs, mantar vs.) tarafından üretilenlerdir (Supuran ve Winum, 2009). 2006 yılında Amerika Birleşik Devletlerinde 21000 civarında ilaç ürünü olduğu bildirilmiştir. Bunlardan 1357 tanesinin özgün ilaç olup, bu ilaçlardan 1204 tanesinin küçük molekül tipi ilaç, 192 tanesinin prodrug (ilaç ana maddesi; vücutta parçalanıp bir parçanın ilaç etkisi göstermesi) olduğu ve 166 tanesinin de biyolojik ilaç olduğu bildirilmiştir (Overington vd., 2006). İlaç hedefi olan proteinler sınıflandırıldığında bunlardan 66 tanesi insan enzimleri, 20 tanesi patojen (bakteri, virüs, mantar vs.) enzimleri, 20 tanesi G-proteine kenetli aile reseptörleri, 12 tanesi nükleer reseptörler, 7 tanesi sitokin reseptörleri, 10 tanesi iyon kanalı reseptörleri ve 10 tanesi de hücre membranı reseptörleri olduğu bildirilmiştir (Imming vd., 2006).

Karbonik anhidraz (CA, EC 4.2.1.1) en basitinden en gelişmiş olana kadar doğadaki tüm canlılarda var olan bir metalo-enzimdir. Bu enzimin doğada pek çok formu olmakla beraber, bulundukları canlıya bağlı olarak beş gruba (α -CA, β -CA, δ -CA, γ -CA ve ϵ -CA) ayrılmıştır. Bu gruplardan α -Karbonik anhidraz memelilerde bulunur ve en temel fonksiyonu kanda ve diğer dokularda



reaksiyonunu katalize etmektedir. Bu basit reaksiyonun vücuttaki birçok fizyolojik süreçte önemli fonksiyonu vardır. Bu reaksiyon; metabolize edici dokular ile akciğer arasında karbondioksit/bikarbonat transferi, dokulardaki asit-baz dengesinin ayarlanması, birçok organda elektrolit salgısının düzenlenmesi, kemik erimesi, kireçlenme, biyosentetik reaksiyonlar (örneğin; glukoneogenez, lipogenez ve ureagenez) ve tümör oluşumu gibi fizyolojik süreçler ile doğrudan ilişkilidir (Supuran, 2009). α -Karbonik anhidraz sınıfındaki enzimlerin katalitik fonksiyonu bu enzimlerin aktif sitesi içerisinde yerleşmiş kofaktör Çinko (Zn^{+2}) iyonu vasıtasıyla gerçekleşir. Çinko iyonu, yaklaşık 15 Å derinliğinde huni şeklinde, bir tarafı hidrofobik diğer tarafı hidrofilik aminoasitler içeren enzim aktif sitesinin tabanında bulunur. Çinko iyonu aktif site içerisinde tetrahedral olarak yerleşmiştir. Bu yapı bağlarının üçü Histidinlerin (His 94, His 96, His 119) yan zincirlerindeki İmidazol halkalarının Azot atomları üzerinden olurken dördüncü bağ su veya hidroksil iyonu ile oluşmaktadır (Supuran vd., 2004). Çinko iyonuna yerleşmiş olan su molekülü enzim aktif sitesinde bulunan Treonin'in (199) hidroksil grubuyla ve Glütamikasitin (106) karboksilat grubuyla Hidrojen bağı yapmaktadır. Bu etkileşim su molekülünü karbondioksit tarafından yapılacak nükleofilik atak için uygun bir hedef haline getirmektedir (Supuran, 2009). α -Karbonik anhidraz sınıfındaki enzimlerin katalitik reaksiyonu hakkında daha detaylı bilgi bölüm 3'te verilecektir.

α -Karbonik anhidraz sınıfındaki enzimler yapısal olarak 15 farklı alt izo-enzim olarak ayrılmıştır. Bu izo-enzimler yapısal açıdan çok az farklılık göstermelerine karşın katalitik aktivitelerinin kinetik hızları, hücre içerisindeki lokalizasyonları ve organ/doku içerisindeki dağılımları açısından oldukça farklıdır. 15 adet izo-enzimden 12 tanesi (CA I, CA II, CA III, CA IV, CA VA, CA VB, CA VI, CA VII, CA IX, CA XII, CA XIII ve CA XIV) enzimik aktiviteye sahiptir. 3 tanesi (CA VIII, CA X ve CAXI) ise karbonik anhidraz ilişkili protein olarak adlandırılır ve bunların katalitik aktiviteleri gözlenmemektedir. Bu 12 katalitik izo-enzimin aktif sitesinin tamamında Çinko atomu mevcut olup 3 Histidin aminoasidi ve su/hidroksil iyonu ile bağ haldedir. Aktif site içerisinde bulunan dördüncü bir Histidin rezüdü (His 64) sürekli konformasyon değişimi yaparak katalitik bölgeye Hidrojen taşıyan bir pompa gibi davranır (Pastorekova vd., 2009). Bu izo-enzimlerin hücre içerisindeki lokalizasyonları farklılık gösterir.

Hücre içerisinde CA I, CA II, CA III, CA VII, CA VIII, CA X, CA XI ve CA XII sitoplazmada; CA IV, CA IX, CA XII ve CA XIV hücre zarında; CA VA ve CA VB mitokondride ve CA VI ise tükürük bezi ve meme bezi salgısında bulunmaktadır (Alterio vd., 2012). CAI ve CAII izoenzimleri insan eritrosit hücreleri ile ilişkili olup aynı zamanda en önemli işlevleri solunumda yer almalarıdır. İnsan dokularında en fazla miktarda bulunan CA II izoenzimi aynı zamanda göz lensi, kornea ve epitelyumda bol miktarda, böbrek korteksinde ise membrana yapışık halde mevcuttur. CA IX ve CA XII izoenzimleri tümörlerle ilişkili olmanın yanı sıra kanser hücrelerinde de baskındır.

Karbonik anhidraz enzimin inhibisyonu temeline dayanan ilk ilaçlar 1950-1960 yıllarda piyasaya çıkmıştır. Bunlar hali hazırda kullanılmakta olan acetazolamide, methazolamide, ethoxzolamide, sulthiame ve dichlorophenamide'dir. Bunlar diüretik (idrar söktürücü) ve antiepileptik (epilepsi tedavisi ilacı) olarak kullanılırken karbonik anhidraz enzimi inhibisyonunun Glukoma (göz tansiyonu) tedavisinde de kullanılabileceğinin farkına varılmıştır (Supuran, 2008). Bu bahsedilen ilaçlar halen göz tansiyonu tedavisinde kullanılmaktadır. Bunlara ilaveten 1990'lı yıllarda damla formunda kullanılabilen iki göz tansiyonu ilaç daha piyasaya çıkmıştır. Bunlar dorzolamide ve brinzolamide olup öncekilere göre çok daha az istenmeyen yan etkiye sahiptir.

2000'li yıllardan itibaren karbonik anhidraz enzimin inhibisyonu temeline dayanan kanser tümörü tedavisi ilacı geliştirme çalışmaları popüler bir hal almaya başladı. 2006 ve 2007 yılında yapılan çalışmalarda CA IX izo-enzimin kanser tümöründeki hücrelerde asit-baz dengesinin bozulmasında ve bu kanser hücrelerin fizyolojisinde önemli bir fonksiyonunun olduğunun tespit edilmesi (Brahimi-Horn ve Pouyssegur, 2007) ve CA XII izo-enziminin sağlıklı hücrelere oranla bazı tip kanser tümöründeki hücrelerde daha fazla oranda üretiliyor olması bu iki izo-enzimi seçici olarak inhibe etme temeline dayalı kanser ilacı geliştirme çalışmalarına temel oluşturmuştur (Pastorekova vd., 2006). Günümüzde Karbonik anhidraz enzimin inhibisyonu temeline dayanan ilaç geliştirme çabaları göz tansiyonu ve kanser ilacı geliştirme konusunda yoğunlaşmakla beraber ödem giderici, epilepsi, obezite, kısırlık, irtifa rahatsızlığı gibi birçok hastalığın tedavisinde kullanılmak üzere ilaç geliştirme çalışmalarında hedef enzim olarak kullanılmaktadır.

Farklı hastalıklar için farklı izo-enzimlerin seçici şekilde inhibisyonunun yapılması ilaç geliştirme çalışmalarında hayati önem taşımaktadır. Hedef dışındaki farklı izo-enzimlerin katalitik aktivitesi inhibe edildiğinde birçok durumda tedavide istenmeyen yan etki olarak ortaya çıkması mümkündür. Bu nedenle izo-enzim seçici inhibitörlerin tasarımı bu alanda çok önemlidir. CA I izo-enziminin seçici inhibisyonu ile retina ve beyin ödemeine karşı etkili ilaç geliştirme çalışması yapılmaktadır (Gao vd., 2007). Göz tansiyonu ilacı geliştirilmesinde CA II, CA IV ve CA XII hedef iken CA IX ve CA XII kanser tümörü tedavisi ilaç için hedef enzimdir. Kanser tedavisi amaçlı ilaç molekülünün CA I ve CA II izo-enzimlerini inhibe etmemesi istenir. Çünkü CA I ve CA II vücutta en yaygın ve yoğun olan izo-enzimler olduğundan bunların inhibisyonu ciddi yan etkiler olarak ortaya çıkar. CA VA ve CA VB izo-enzimlerinin seçici inhibisyonu anti-obezite ilaç geliştirme konusunda hedef enzim olarak belirlenmiştir (Supuran vd., 2008). CA II, CA VII, CA XII ve CA XIV seçici inhibisyonu epilepsi tedavisi için ilaç geliştirme konusunda hedef enzim olarak belirlenmiştir (Hen vd., 2011; De Simone vd., 2009).

Karbonik anhidraz inhibitörü olarak kumarin türevlerinin bazıları da izoenzim seçici inhibisyon sergilemesi sebebiyle keşfedilmiştir. Bu durum sülfonamidlerin CA enziminin aktif bölgesindeki Zn^{+2} iyonu ile doğrudan bağ yaparak etki gösterirken; kumarinlerin Zn^{+2} ile bağ yapmadan enzimin aktif bölgesinin ağzına “kapak” şeklinde bağlanarak reaksiyonu inhibe etmesi olarak açıklanmıştır. Böylelikle üzerinde yoğun olarak çalışılan sülfonamid tabanlı bileşiklerin seçicilik etkisine karşın kumarin türevlerinin daha seçici olması çalışmaların bu alanda yapılmasına bir etken olabilir. Kumarin türevi bir molekülün karbonik anhidraz enzimini inhibe ettiği ilk defa 2009 yılında rapor edilmiştir (Maresca vd., 2009). Kumarin türevlerinin karbonik anhidraz enzimini inhibisyonunda ki en ilginç durum inhibisyonun hemen gerçekleşmeyip 6 saat kadar bir süre enzim ve inhibitörün inkübasyonuna gereksinim duyulmasıdır. X-ışınları kırstalografisi sonuçları kumarin türevlerinin karbonik anhidraz enzimini inhibisyonunda ikinci ilginç ve şaşırtıcı bir sonuç daha ortaya çıkarmıştır. Kristalografi sonuçları inhibitör-enzim kompleksinde kumarin türevlerinin halka yapısının aktif site içerisinde hidrolize olduğunu ve enzime cis-2-hydroxy-sinamik asit formunda kompleks yaptığını göstermiştir.

Benzer şekilde, kristalografi sonuçları Sülfokumarin yapısının da karbonik anhidraz aktif sitesi içerisinde hidrolize olmuş formuyla kompleks oluşturduğunu göstermektedir (Tars vd., 2013).

Son olarak ise günümüzde yapılmakta olan QSAR (Quantitative Structure-Activity Relationship / Kantitatif Yapı-Etki Analizleri) çalışmaları ilaç geliştirmede kullanılan vazgeçilmez ve yaygın bir yöntemdir. Bu çalışmanın temeli mevcut olan bir grup molekül ile bu moleküllerin deneysel olarak ölçülmüş olan aktivite değerleri (biyolojik etkiyi arttırıcı ya da azaltıcı yöndeki değerleri) arasındaki ilişkinin istatistiksel yöntemlerle tayin edilmesi esasına dayanır. Ayrıca bu çalışma ile elde edilen modeldeki yeni moleküllerin aktivite değerlerini tahmin edebilme öngörüsü de sağlanır.

Bu çalışma, göz tansiyonu ve ya kanser tedavisinde ilaç olma potansiyeline sahip sanal moleküllerin tasarımını amaçlamaktadır. Yukarıdaki bilgiler göz önüne alındığında göz tansiyonu ilacı adayı moleküllerin CA II, CA IV ve CA XII için yüksek inhibisyon aktivitesi göstermesi beklenirken diğer izo-enzimler için düşük inhibisyon aktivitesi beklenir. CA IV ve CA XII'nin vücutta CA II'ye oranla çok daha az miktarda olduğu göz önüne alındığında CA II için yüksek inhibisyon aktivitesi göstermesi bu amaç için yeterli sayılabilir. Kanser tedavisinde ilaç olma potansiyeline sahip sanal moleküllerin ise CA IX ve CA XII'ye karşı yüksek inhibisyon aktivitesi göstermesi beklenirken diğer izo-enzimler için düşük inhibisyon aktivitesi beklenir. Özellikle CA I ve CA II'ye karşı düşük inhibisyon aktivitesi olması gerekir. Çünkü CA I ve CA II vücuttaki doku ve organlarda yaygın ve diğer izo-enzimlere oranla çok daha yoğun olarak bulunduklarından bu izo-enzimlerin inhibisyonu istenmeyen yan etki olarak ortaya çıkabilir.

Karbonik anhidraz enzimi, inhibitörleri ve bu enzimin izo-enzimlerinin seçici inhibisyonunun ilaç geliştirme çalışmaları açısından yeri ortaya konulduktan sonra bu çalışmanın amacı şu şekilde ifade edilebilir. Glukom (göz tansiyonu) ve kanser tümörü tedavisi için ilaç geliştirme çalışmalarında hedef enzim olarak yaygın biçimde kullanılan karbonik anhidraza karşı, inhibitör özelliği deneysel yöntemle belirlenmiş literatürde bulunan kumarin bileşiklerini kullanarak, QSAR modelleri oluşturmak amaçlanmaktadır. Bu modellerdeki moleküler tanımlayıcıların yorumlanması yoluyla moleküler yapı ile biyolojik etki arasındaki ilişkinin rasyonel bir açıklaması yapılmaya çalışılacaktır.

Glukom tedavisi için karbonik anhidrazın ikinci izo-enziminin seçici şekilde inhibe edilmesi hedeflenirken, kanser tümörlerinin tedavisinde karbonik anhidrazın dokuzuncu (CAIX) izo-enzimlerinin seçici şekilde inhibe edilmesi hedeflenir. Bu projede, literatürde deneysel K_i verileri olan CAII ve CAIX inhibitörü kumarin bileşikleri için kuantum mekanik yöntemle hesaplayacağımız moleküler tanımlayıcıları, deneysel K_i değerleri ile Çoklu Doğrusal Regresyon (ÇDR) işlemine tabi tutarak QSAR modelleri oluşturulacaktır. Elde edilecek QSAR modelleri tasarlanan yeni sanal moleküllerin inhibisyon sabitlerinin hesaplanmasında kullanılacaktır. Ayrıca QSAR modellerinin analizi yoluyla, kumarin ve sülfokumarin türevlerinin CAII ve CAIX izo-enzimlerini inhibisyonunun mekanizmaları mekaniksel olarak açıklanmaya çalışılacaktır.

2. LİTERATÜR ÖZETİ

2.1 Giriş

Karbonik anhidraz enzimi 1933 yılında Meldrum ve Roughton tarafından keşfedilmesinden bugüne üzerinde yapılan araştırmalar hız kesmeden yoğun şekilde devam etmektedir. Bu alandaki çalışmalar 1970’li yıllara kadar enzimin izolasyonu, karakterizasyonu, inhibitor sentezi ve kinetik ölçümü gibi alanlarda yoğunlaşırken Liljas vd, (1972) tarafından ilk olarak karbonik anhidrazın x-ışınları kristalografisi ile üç boyutlu yapısının ortaya çıkartılması ile yeni bir aşamaya gelmiştir. Bu enzimin üç boyutlu yapısının ortaya koyulması vücuttaki fonksiyonlarının anlaşılmasında bir mihenk taşı oluşturmaktadır. Karbonik anhidraz ve bu enzime inhibitör etkisi gösterdiği bilinen en eski molekül grubu sülfonamidlerin QSAR çalışması literatürde ilk defa Testa vd, (1978) yaptığı çalışma ile başlamaktadır. Ayrıca 2000’li yıllarda bu enzimin aktif sitesi içerisinde bulunan çinkoya inhibitörlerin bağlanmasını modelleyen çok sayıda kuantum kimyasal hesaplama çalışmaları da konu olmuştur. Bu bölümde karbonik anhidrazın inhibisyonunu konu alan moleküler modelleme çalışmalarının bir özeti verilecektir. Bu alanda yüzlerce çalışma olduğundan, en fazla katkı yaptığı düşünülen çalışmalardan bir seçki yapılarak birbirinin tekrarı niteliğindeki çalışmalara değinilmeyecektir.

2.2 Karbonik Anhidraz Enzimi İnhibisyonu QSAR Çalışmaları

Karbonik anhidraz inhibitörleri ile yapılan QSAR çalışmaları 1970’li yıllarda başlamıştır. Bizim araştırmalarımıza göre bu çalışmaların tamamında sülfonamit tabanlı inhibitörlerin karbonik anhidraza karşı ölçülmüş aktivite verileri kullanılmıştır. Literatürdeki çalışmaların tamamına yakını karbonik anhidraza II (CA II) izoenziminin biyolojik aktivite verilerinin kullanılmasına yoğunlaşmıştır. Bu alandaki çalışmalar başlıca aşağıda verilmektedir.

Testa vd, (1978) yaptığı ilk QSAR çalışmasında sülfonamitlere bağlı aromatik halka üzerindeki sapsütütentlerin değişimi ile ortaya çıkan sterik etkideki değişimin ligandın inhibisyon gücünü belirleyen önemli faktörlerden biri olduğu ortaya koyulmuştur. Debenedetti vd, (1985 ve 1987) yıllarında yaptıkları QSAR modelinde bir seri sülfonamit türevi moleküle CNDO/2 teori seviyesinde hesaplama yaparak kuantum kimyasal hesaplama yöntemiyle elde ettikleri elektronik indeksler ile bu moleküllerin inhibisyon aktiviteleri arasında çoklu regresyon analizi kullanarak yaptıkları çalışmada aktiviteye etki eden elektronik faktörleri incelemişlerdir. Scholz vd, (1993) yılında bir seri sülfonli metan sülfonamit molekülü ile yaptıkları QSAR çalışmasında, sülfonamit grubuna komşu karbon atomuna bağlanan sapsütütentlerin toplam indüktif etkisinin aktiviteyi tayin etmede en dominant faktör olduğunu göstermişlerdir.

1994 yılından başlayarak Supuran ve Clare’in karbonik anhidraz inhibitörleri ile yapılan pek çok sayıda QSAR çalışması yayınlanmıştır. Yazarlar, hesaplanan moleküler tanımlayıcılar ile çoğu durumda CA II izoenzimine karşı ölçülen deneysel inhibisyon sabiti arasında korelasyon denklemini çoğu kez Çoklu Doğrusal Regresyon (ÇDR) ve Kısmi En Küçük Kareler Regresyonu (KEKKR) tekniklerini bazen de doğrusal olmayan istatistiki bir yöntem olan Koşullu Beklentiler Metodunu (ACE: conditional expectations technique) kullanarak elde etmişlerdir (Breiman ve Friedman, 1985). Clare ve Supuran (1994) yılında sülfonamitler ile yaptıkları ilk QSAR çalışmasında CNDO/2 kuantum kimyasal hesaplama yöntemi ile elde ettikleri moleküler tanımlayıcılar ile CA II izoenzimine karşı ölçülen deneysel inhibisyon sabiti arasında korelasyon denklemlerini elde etmişler.

Yazarlar bu çalışmada ligantlardaki oksijen ve azot atomları üzerindeki kısmi yük miktarı ile molekülün boyutlarının aktiviteyi belirlemede en etkin moleküler tanımlayıcılar olduğunu göstermiştir. Supuran ve Clare (1995) yılında içinde iyonik sülfonamitlerin de bulunduğu QSAR çalışmasında PM3 kuantum kimyasal hesaplama yöntemi ile elde ettikleri moleküler tanımlayıcıları kullanmışlardır. Yazarlar, SO_2NH_2 grubuna bağlı pridin halkasında pozitif yük miktarındaki artışın ve hem x hem de z yönlerinde polarizibilite tensörleri bileşenlerinin değerlerindeki artışın inhibisyon aktivitesini arttırdığını, buna karşın ligantların HOMO enerjilerinde ve z yönündeki polarizibilite tensörü bileşenindeki artışın ise inhibisyon aktivitesini düşürdüğünü göstermiştir.

Supuran ve Clare (1990)'larda birçok benzer QSAR çalışmalarına devam etmişlerdir. Supuran ve Clare (1997) yılında yaptıkları QSAR çalışmasında aktiviteyi domine eden moleküler tanımlayıcıların SO_2NH_2 grubundaki atomların Mulliken kısmi yükleri, ligandların boyutları ve lipofilisiti özelliğini temsil eden LogP (octanol-su partision katsayısı) olduğunu bildirmişlerdir. Supuran ve Clare (1998) yılında yaptıkları çalışmada ise SO_2NH_2 grubundaki atomların elektrostatik potansiyel tabanlı hesaplanmış kısmi yüklerini, ligandların dipol momentlerini AM1 metodu ile çözücü içerisinde hesaplayarak korelasyon denklemini elde etmişlerdir. Yazarlar CA I ve CA II izoenzimlerinin sülfonamitler ile etkileşimlerinde seçiciliği sağlayan birtakım faktörlerin ne olabileceği konusunda önermelerde bulunmuşlardır. Ayrıca yazarların 1998 ve 1999 yıllarında yaptıkları benzer QSAR çalışmalarında da SO_2NH_2 grubundaki atomların ESP tabanlı kısmi yükleri, Homo, Lumo, dipol moment, polarizibilite tensörleri ve ligandların COSMO yaklaşımıyla hesaplanmış çözünme enerjilerini moleküler tanımlayıcı olarak başarı ile uygulamışlardır. Bu moleküler tanımlayıcıların mekanistik yorumlamaya uygun olmaları yeni sentezlenecek ligandların tasarımına katkı yaptığına şüphe yoktur.

Supuran ve Clare, (2000)'li yılların başından itibaren yaptıkları çalışmalarda benzene-sülfonamit grubuna bağlı sabsüte benzene ve pridin halkalarındaki pi tipi bağların düğüm noktalarının yönelim açılarının karbonik anhidraz enzimi inhibisyonunun QSAR modellemesinde etkili bir moleküler tanımlayıcı olduğunu ortaya koymuşlardır. Supuran ve Clare, (2001) yılında bu seride yaptıkları ilk çalışmalarını 27 adet Schiff bazlı benzene-sülfonamit türevi ile yürütmüşlerdir.

Yazarlar bu çalışmada moleküllerdeki benzene ve imidazol gibi halka sistemlerindeki (pi-tipi orbital) öncü orbitallerin faz açılarını (FOPA; Frontier Orbitals Phase Angles) moleküler tanımlayıcı olarak kullanırken bu tanımlayıcıların aktiviteye nasıl etki yaptığı hususunda bazı yorumlar ortaya koymuşlardır. Bu yorumlara göre FOPA'nın aktiviteye etkisi; ligandlar üzerindeki pi-orbitali benzer orbitallerdeki yükler ile enzimin pi-orbitali benzeri orbitaller arasında yük transferi etkileşimi oluşmaktadır. FOPA moleküler tanımlayıcılarının CA II inhibisyonu QSAR modellemesinde etkin olurken CA I modellemesinde etkin olmayışı buna kanıt olarak sunulmaktadır. Çalışılan seride bulunan ligandların CA I ve CA II'ye karşı farklı efinite sergilemesinin sebebi şu şekilde açıklanmıştır. CA II aktivitesinde His64 rezidüsünün imidazol halkası ile ligandın sabsüte benzene halkası arasında yük transfer etkileşiminin olmasına karşın CA I katalitik aktivitesinde His64'ün rolünün olmadığından dolayı yük transferi etkileşimi de söz konusu olmadığı için bu moleküler tanımlayıcıların CA I QSAR modelinde yer almamış olması sonucuna götürmüştür. Supuran ve Clare, (2005) yaptığı iki çalışmada daha geleneksel olarak kullandıkları, sülfonamit grubunda bulunan N, O, S atomları üzerindeki Mulliken kısmi yükleri, molekülün dipol momenti, polarizibilite tensörleri, öncü orbitallerin (Homo, Lumo vs.), çözünme enerjileri yanında pi-tipi orbitallerin faz açılarını moleküler tanımlayıcı olarak kullanmışlardır.

2000'li yıllardan sonra topolojik indeksleri kullanarak sülfonamitlerin karbonik anhidrazın başta CA II olmak üzere birçok farklı izoenzimine karşı ölçülmüş inhibisyon aktivitesi verisini modelleyen QSAR çalışması yapılmıştır. Topolojik indekslerin moleküler tanımlayıcı olarak kullanılması ile elde edilen QSAR modelleri her ne kadar istatistiki açıdan tatminkâr olsa da bu moleküler tanımlayıcıların doğaları gereği fiziksel yorumlamaya pek uygun değildir.

Khadikar vd,(2005) benzene sülfonamitlerin CA II ile bağlanma sabitlerini modelleyen QSAR çalışmasında Wiener, Szeget, Randic, Kier Hall ve Padmakar-Ivan topolojik indekslerini üç gösterge parametresi ile beraber kullanmışlardır. Aynı çalışmada kullanılan topolojik indekslere ilaveten moleküllerin hidrofobisinin derecesini temsil eden logP modellere eklendiğinde modellerin istatistik değerlerinde bariz bir artışın olduğu rapor edilmiştir.

Agrawal vd, (2002) 39 adet sülfonamitin CA II izoenzimine karşı ölçülmüş inhibisyon sabitlerini modelleyen QSAR çalışmasında Wiener, Szeget, Balaban, dallanma ve bağlanma topolojik indekslerini iki gösterge parametresi ile beraber kullanmışlardır. Thakur vd, (2004) 29 adet sapsüte benzene sülfonamitin CA II ye karşı ölçülmüş inhibisyon sabitleri ile yaptıkları QSAR modellemesinde tek bir moleküler tanımlayıcı (Balaban indeksi) içeren QSAR modelinin istatistiği oldukça iyi sonuç verdiği rapor edilmiştir (n=29, r=-0.95, F=274.19 ve SE=0.42).

Jaiswal vd, (2004) topolojik indeksleri kullanarak bir QSAR çalışması yapmışlardır. Bu çalışmanın en özgün tarafı CA IX izoenzimine karşı ölçülmüş inhibisyon sabitlerinin ilk defa bir QSAR çalışmasında kullanılmış olmasıdır. Ancak bu çalışmada sunulan QSAR modelleri istatistiki açıdan sorunludur. Dolayısıyla kullanılan veri seti QSAR modellemesi yapmaya uygun değildir. Literatürde yukarıda bahsedilenler dışında topolojik indekslerin kullanıldığı, CAII yanında CA I, CA IV ve CA VII'ye karşı ölçülmüş inhibisyon aktivitelerini kullanan çok sayıda QSAR çalışması mevcuttur.

2000'li yıllarda karbonik anhidraz inhibisyonunu konu alan sınırlı sayıda 3D-QSAR çalışması da yayınlanmıştır. Bunlar başlıca şöyledir. Weber vd, (2006) yaptıkları kapsamlı bir 3D-QSAR çalışmasını 87 adet aromatik ve heterosilik sülfonamit türevinin CA I, CA II ve CA IV'e karşı ölçülmüş inhibisyon sabitlerini kullanarak CoMFA ve CoMSIA tekniklerini uygulayarak yürütmüşlerdir. CA I, CA II ve CA IV için ayrı ayrı elde edilen, 60 adet eğitim molekülünden oluşturulan CoMFA ve CoMSIA modelleri 27 test molekülünün aktivitelerini başarılı şekilde belirlemişlerdir (tüm modellerde $R^2 > 0.85$). Elde edilen modellerin sterik, hidrojen donor ve hidrojen akseptör alanlarının grafik olarak görselleştirilmesinin yorumlanması ligandların izoenzim seçiciliği konusunda bilgiler vermiştir. Aynı araştırma grubu, Hiilebrecht vd., (2006) aynı yıl yayınladıkları ikinci makalede, CoMFA ve CoMSIA tekniklerine ilave olarak AFMoC (Adaptation of Fields for Molecular Comparison) (Gohkle ve Klebe, 2002) yaklaşımını da CA I, CA II ve CA IV inhibisyon aktivitelerinin modellenmesinde daha geniş bir sülfonamit seti için kullanmışlardır.

Tuccinardi vd, (2007) yılında yayınladıkları bir makalede moleküler kenetlenme ve reseptör tabanlı 3D-QSAR tekniklerini birlikte kullandıkları kapsamlı bir çalışmanın sonuçlarını sunmuşlardır. Protein Data Bank (PDB)'tan alınan 40 adet CA II sülfonamit kompleksi ile GOLD yazılımını kullanarak moleküler kenetlenme çalışması yürütmüşlerdir. 294 sülfonamit molekülü ile de reseptör tabanlı 3D QSAR çalışması yürütülmüştür. 220 molekül eğitim seti olarak ayrılmıştır. Diğer kalan moleküller üç farklı test sete ayrılmıştır. PLS yaklaşımı kullanılarak elde edilen 3D QSAR modellerinin istatistik düzeyi iyi seviyede olduğu görülmüştür ($R^2=0.88$ ve $Q^2=0.72$). Tuccinardi vd, (2007) yılında yayınladıkları ikinci makalede yukarıda anlatılan reseptör tabanlı 3D-QSAR tekniğini CA IX için uygulamışlardır. Elde edilen 3D QSAR modellerinin istatistik seviyesi bir önceki çalışmaya sunulan verilere benzerlik göstermiştir.

Huang vd, (2007) yılında CoMFA ve CoMSIA teknikleri ile yaptıkları 3D QSAR çalışmasında sülfonamit türevlerinin CA II izoenzimine karşı inhibisyon aktivitelerini modellemişlerdir. Yazarlar CoMFA ve CoMSIA teknikleri hem ligand tabanlı (ligandların minimum enerji konformasyonları karşılaştırılarak elde edilen alanlar) alanları hem de reseptör tabanlı (ligandlar CA II aktif sitesi içine kenetlenerek elde edilen bio-aktif konformasyonları karşılaştırılarak elde edilen alanlar) alanları kullanarak yürütmüşlerdir. Sonuçlar karşılaştırıldığında reseptör tabanlı CoMFA modelinin en iyi istatistiki veriye sahip olduğu görülmüştür ($n_{\text{training}}=51$, $R^2=0.88$, $Q^2=0.72$, $n_{\text{test}}=9$ ve $R_{\text{ext}}^2=0.74$).

Literatürde bunlar haricinde de çok sayıda CA inhibisyonu konusunda QSAR çalışması mevcuttur. Ancak burada verilenler ve literatürde olanların tamamı sülfonamitler ile yapılan çalışmalardır. Supuran, (2015) yılında CA izoenzimlerinin inhibisyonun da şimdiye kadar bilinen beş farklı mekanizmanın olduğunu söylemiştir. Bu mekanizmalardan en yaygın olanı yukarıda bahsedilen çalışmalara konu olan sülfonamitlerin de bağlanma mekanizması olan SO_2NH^- grubunun aktif site içerisindeki Zn^{+2} iyonu ile yarı kovalent bağlanma şeklidir. Literatürdeki tüm QSAR çalışmaları bu mekanizma ile bağlanan ligandlar ile yürütülmüştür. Bu çalışmanın esasını oluşturan kumarinlerin, Maresca vd, (2009) yılında yaptığı deneysel x-ışınları kırınımı çalışmasında gösterdiği üzere inhibisyon mekanizması oldukça farklı olup, enzim aktif sitesi ağzına kovalent bağ yapmaksızın bir takım elektrostatik etkileşim ve hidrojen bağları ile tutundukları bilinmektedir. Literatüre

bakıldığında kumarinler ile yapılmış bir QSAR çalışmasına rastlanmamaktadır. Bu çalışma bu alanda bir ilk olacaktır.



3. GEREÇ VE YÖNTEM

3.1 Giriş

Bu bölümde üzerinde moleküler modelleme çalışması yürüttüğümüz enzim olan Karbonik anhidrazın inhibisyon mekanizması ve inhibitörleri arasında yer alan kumarinler hakkında genel bilgiler verildikten sonra modelleme yöntemi olarak kullandığımız QSAR metodu hakkında teorik arka plan verilecektir. Bir veri seti için QSAR modellerini oluşturmada çok farklı moleküler tanımlayıcılar kullanmak mümkündür. Ancak elde edilen modellerin kullanışlı olabilmesi için modeli oluşturan moleküler tanımlayıcıların yorumlanabilir olması oldukça önemlidir. Fiziksel yorumlamaya en uygun moleküler tanımlayıcılar arasında kuantum mekaniksel hesaplama ile elde edilen tanımlayıcılar en önde yer alır. Bu çalışmada moleküler tanımlayıcıları hesaplamada kuantum mekaniksel hesaplama yöntemi olarak DFT (Density Functional Theory/Yoğunluk Fonksiyonel Teorisi) kullanılmıştır. Dolayısıyla moleküler kuantum mekanik hesaplamaları hakkında genel ve DFT için özel teorik arka planda bu bölümde verilecektir. Ayrıca dördüncü bölümde vereceğimiz QSAR modelleri içerisinde yer alan moleküler tanımlayıcıların tanımları da bu bölümde verilecektir.

Bir veri seti için QSAR modelleri birçok farklı algoritma ile elde edilebilir. Bunlardan en yaygın olanlar arasında ÇDR Analizi de yer almaktadır. Bu araştırmada kullanılan bu yönteme dair teorik arka plan bu bölümde verilecektir. Elde edilen QSAR modellerinin uygulanabilirlik sınırlarının (AD: Application Domain) belirlenmesi önemlidir. Hiçbir QSAR modelinin tüm kimyasal evren için geçerli olması mümkün değildir. Onun için bir QSAR modelinin uygulanabilirlik sınırlarını belirleyen algoritma geliştirilmiştir. Bu çalışmada popüler uygulanabilirlik sınırı hesaplama yöntemleri arasında yer alan Leverage (kaldıraç gücü) yöntemi kullanılmıştır. Bu yönteme dair arka plan da bu bölümde verilecektir. Son olarak tüm bu bahsedilen algoritmaların yer aldığı bu araştırmada kullanılan yazılım ve donanım hakkında genel bilgiler sunulacaktır.

3.2 Karbonik Anhidraz

Meldrum ve Roughton (1933) yılında bulunan karbonik anhidraz (CA, EC 4.2.1.1) en basitinden en gelişmiş olana kadar doğadaki tüm canlılarda var olan bir metaloenzimdir. Bu enzimin doğada pek çok formu olmakla beraber, bulundukları canlıya bağlı olarak beş sınıfa (α -CA, β -CA, δ -CA, γ CA ve ε -CA) ayrılmıştır. Bu sınıflardan α -Karbonik anhidraz memelilerde bulunur ve en temel fonksiyonu kanda ve diğer dokularda



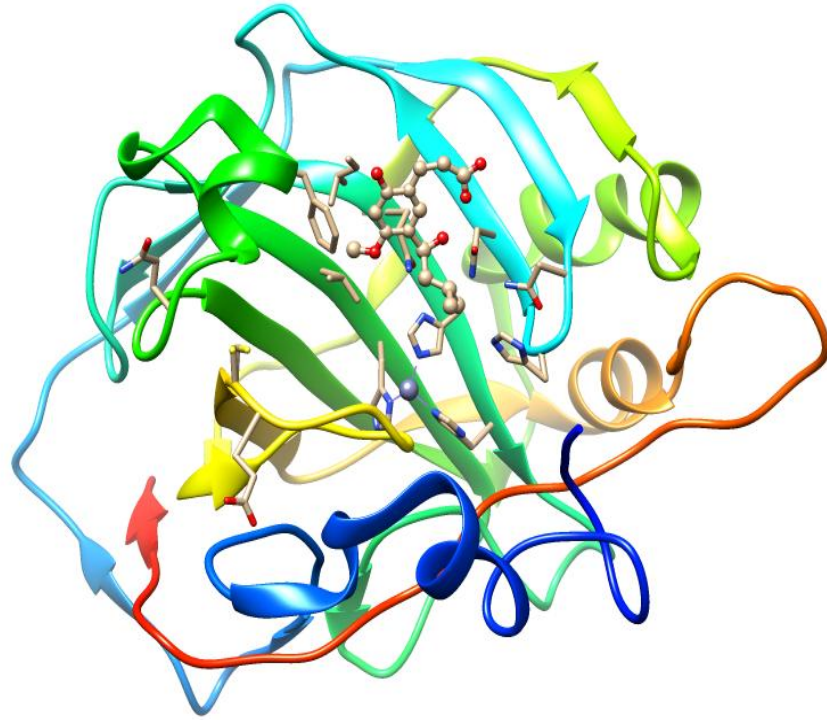
CO₂'nin hidratasyonu ve HCO₃⁻ iyonunun dehidratasyonu reaksiyonunu katalize etmek vasıtası ile birçok fonksiyonu yerine getirir. Karbonik anhidraz fonksiyonları arasında asit baz dengesini ayarlamak, ve dokulardan karbondioksit transferini sağlamaktır.

Aynı canlı türünde aynı reaksiyonu katalizleyen ancak farklı kimyasal ve fiziksel özellikleri olan enzimlere *izoenzim* veya *izozim* denir. İzoenzimlerin substratlarına, kofaktörlerine ve inhibitörlerine karşı ilgileri farklıdır. İzoenzimlerin başlıca özellikleri arasında aminoasit sayı ve sırasının farklı olması, izoelektrik pH değerinin farklı olması, her bir alt ünitenin ayrı geninin olması ve elektroforetik hareketliliklerinin farklı olması sayılabilir. İzoenzimler farklı dokularda lokalize olabildiği gibi, bir hücrenin subsellüler fraksiyonlarında da yerleşebilirler (Keha ve Küfrevioğlu, 2004).

α -Karbonik anhidraz sınıfındaki enzimler yapısal olarak 15 farklı alt izoenzim olarak ayrılmıştır. Bu izoenzimler yapısal açıdan çok az farklılık göstermelerine karşın katalitik aktivitelerinin kinetik hızları, hücre içerisindeki lokalizasyonları ve organ/doku içerisindeki dağılımları açısından oldukça farklıdır. Ayrıca bu izoenzimlerin 12 tanesi (CAI, CAII, CAIII, CAIV, CAVA, CAVB, CAVI, CAVII, CAIX, CAXII, CAXIII ve CAXIV) enzimik aktiviteye sahipken, 3 tanesi (CAVIII, CAX ve CAXI) karbonik anhidraz ilişkili protein olarak adlandırılır ve bunların katalitik aktiviteleri gözlenmemektedir. Bu izoenzimlerin hücre içerisindeki lokalizasyonları farklılık gösterir. Hücre içerisinde CAI, CAII, CAIII, CAVII, CAVIII, CAX, CAXI ve CAXII sitoplazmada bulunur. CAIV, CAIX, CAXII ve

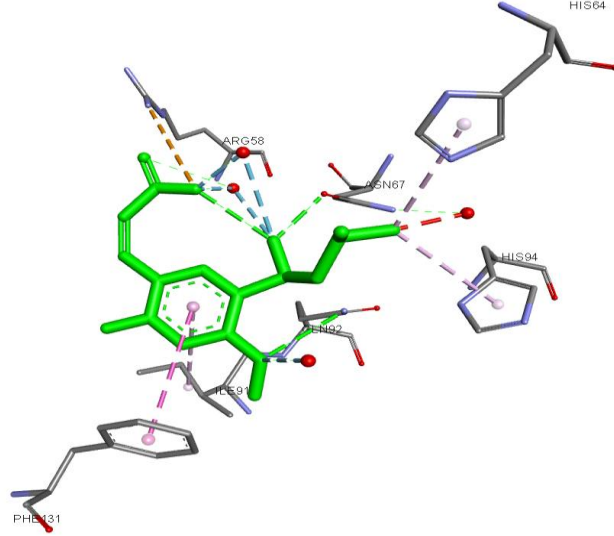
CAXIV hücre zarında, CAVA ve CAVB mitokondride, CAVI ise tükürük bezi ve meme bezi salgısında bulunmaktadır (Alterio vd., 2012).

CA'nın farklı izoenzimlerinin seçici inhibisyonu farklı hastalıkların tedavisinde kullanılacak ilaçların geliştirilmesinde hedef enzim olarak kullanılmaktadır. Bu enzimin aktif bölgesinde Çinko (Zn^{+2}) atomu vardır. Zn atomu enzimin inhibisyonu esnasında genellikle inhibitör ile yarı kovalent karakterli tetrahedral bir bağ yapısı gerçekleştirir. Ancak bu çalışmaya konu olan kumarin bileşikleri farklı bir mekanizma ile CA'ya bağlanmaktadır. Şekil 3,1'de Protein Data Bank'da 3F8E kodu ile bulunan CAII'nin ((2Z)-3-{2-hydroxy-5-[(1S)-1-hydroxy-3-methylbutyl]-4-methoxyphenyl}prop-2-enoic asit ile kompleksinin X-ışınları kristalografisi ile elde edilen üç boyutlu yapısı gösterilmektedir. Şekilde görüldüğü gibi hidrolize olmuş kumarin türevinin Zn ile etkileşimi söz konusu değildir. Bağlanma aktif site içerisinde birçok elektrostatik etkileşimler ve hidrojen bağları ile gerçekleşmektedir.



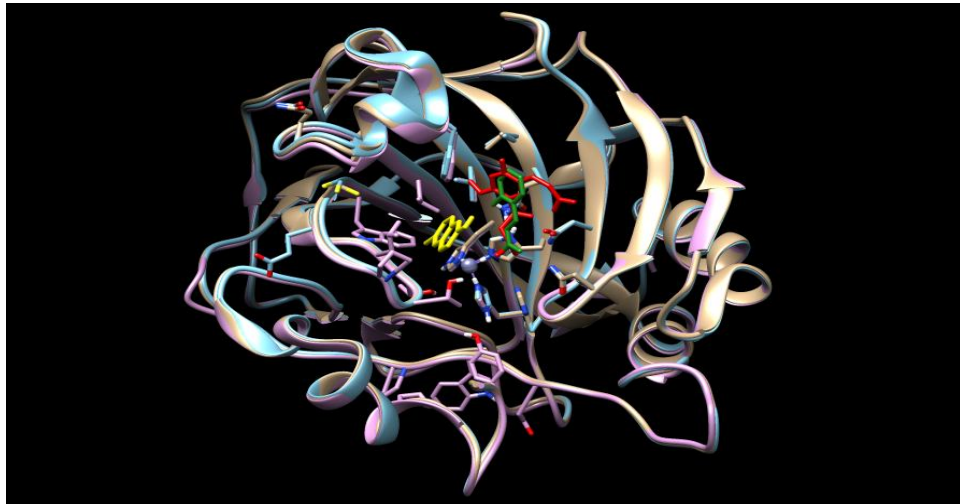
Şekil 3.1: CAII- ((2Z)-3-{2-hydroxy-5-[(1S)-1-hydroxy-3-methylbutyl]-4-methoxyphenyl}prop-2-enoic asit kompleksi(PDB kod: 3F8E).

Bu bağlanmada CA II aktif sitesi içerisinde yer alan birçok aminoasit rol oynamaktadır. Şekil 3.2'de ise CAII'nin ligand ile yaptığı etkileşimleri göstermektedir. Bu şekil Discovery Studio (Biovia, 2016) yazılımı ile PDB'deki 3F8E yapısından üretilmiştir.



Şekil 3.2: CAII aktif sitesi içerisindeki aminoasitlerin hidrolize kumarin türevi ile etkileşimleri.

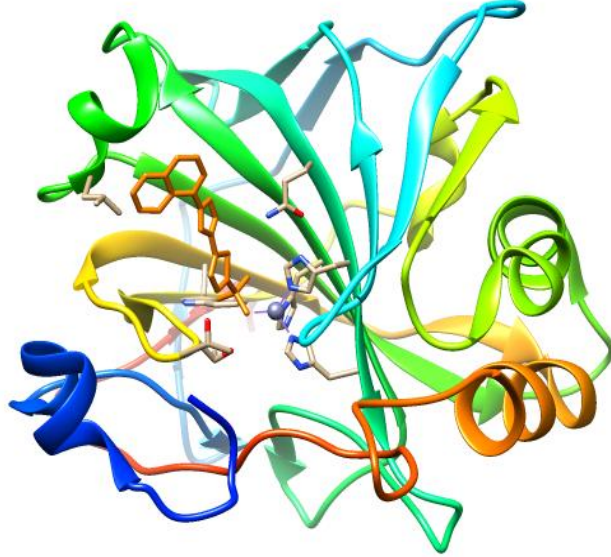
Bu ligandın sergilediği etkileşimler arasında öne çıkanların, PHE131'in fenil halkası ile ligandın fenil halkasının pi-stok (edge-to-face) ve Asn 67'nin karbonil oksijeni ile ligandın S-OH bölgesi arasındaki dipol-dipol etkileşimleri olduğu söylenebilir. Hali hazırda PDB'de üç adet (3F8E, 5BNL ve 4WL4) kumarin türevinin CA II ile oluşturduğu enzim ligand kompleksi mevcuttur. Şekil 3.3'de bu üç yapının üst üste (superimpose) getirilmiş görüntüsü verilmektedir. Üst üste getirme işlemi UCSF Chimera yazılımı (Pettersson vd., 2013) ile gerçekleştirildi.



Şekil 3.3: PDB'de bulunan üç kumarin bileşiğinin üst üste (superimpose) görüntüsü.

Yukarıdaki Şekil 3.3'den de görüldüğü gibi kumarin türevlerinin aktif sitesi içerisindeki lokasyonları farklılık göstermektedir. Sarı renkli tio-kumarin (6-hydroxy-2H-chromene-2-thione) hidrolize olmaksızın aktif sitenin dibine yakın bir yerde yer alırken, turuncu renkli ((2Z)-3-{2-hydroxy-5-[(1S)-1-hydroxy-3-methylbutyl]-4-methoxyphenyl}prop-2-enoic acid) ve yeşil renkli aktif sitenin ağız kısmında yer almaktadır. Yeşil ve turuncu kumarinlerdeki açılan lakton halkalarının farklı yönelimlerde olduğu dikkat çekmektedir.

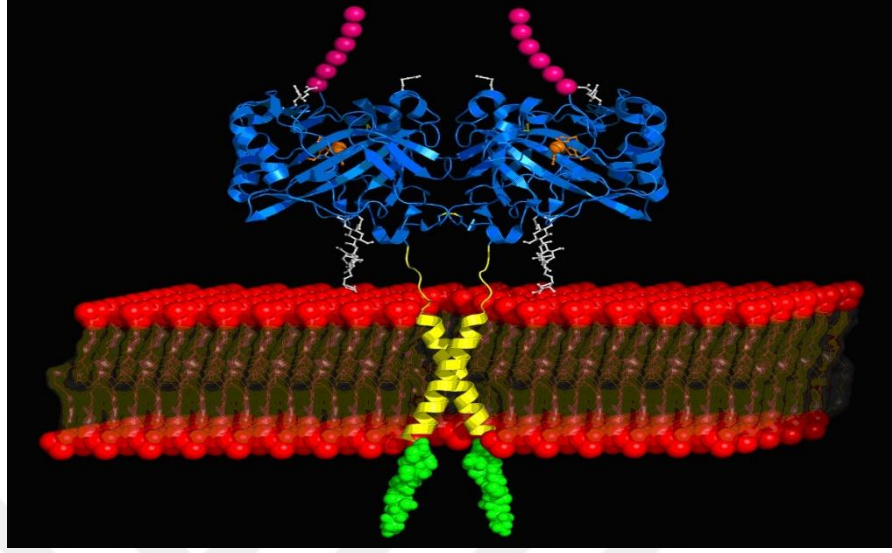
Bu çalışmaya konu olan diğer izoenzim olan CAIX'un herhangi bir kumarin türevi ile PDB'de kompleksinin yapısı bulunmamaktadır. CAIX'un CAII'ye benzerliğinin sergilenmesi açısından CAIX'un bir sülfonamid türevi ile olan kompleksinin üç boyutlu yapısı Şekil 3.4'de gösterilmektedir. Şekil 3.4 PDB'de 5FL4 kodla bulunan CAIX sülfonamid kompleksinin tetramer yapısından UCSF Cimer yazılımı ile elde edilmiştir (Pettersen vd., 2013). CAII sitoplazmada bulunan bir izoenzim iken CAIX trans-membran bir izoenzim olup makro molekülün büyük bölümü sitoplazmada değil de hücre zarı dışında bulunur. Şekil 3.4'de gösterilen bölüm hücre zarı dışında bulunan kısmı içermektedir ve CAIX'un katalitik domaini olarak adlandırılır.



Şekil 3.4: CAIX5-(1-naphthalen-1-yl-1,2,3-triazol-4-yl)thiophene-2-sulfonamide kompleksi (PDB kod:5FL4).

CA farklı izoenzimlerinin aminoasit sekansları büyük oranda benzerlik gösterir. Monomerik CAIX 422 aminoasitten oluşur. CAIX yapı olarak, bir sinyal peptit, proteoglikan bölge, CA katalitik aktif bölgesi (Şekil 3.4'de görülen), transmembran

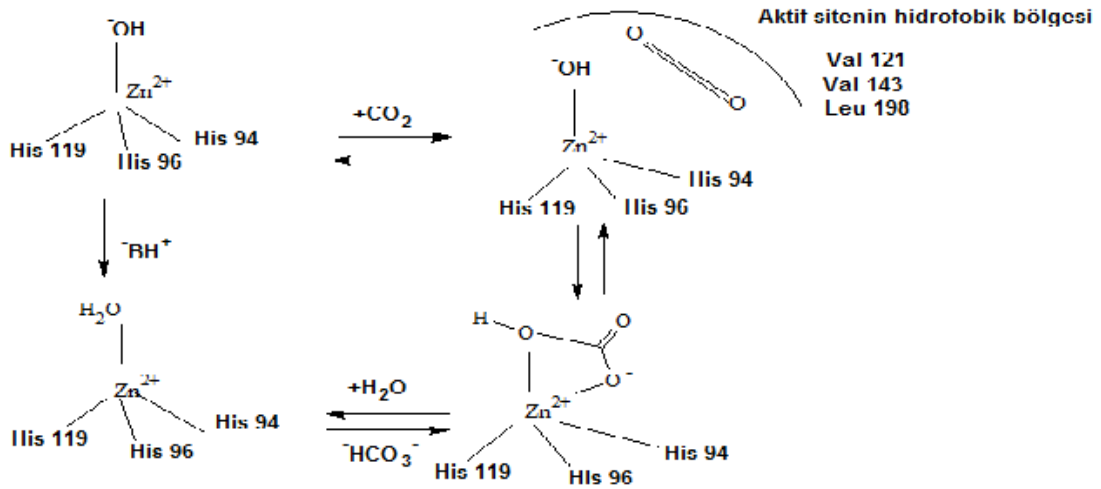
bölgesi ve hücre içi intrastoplazmik kuyruk bölümlerinden oluşur. CAIX izoenziminin bahsedilen tüm bu bölümlerini gösteren bir model Şekil 3.5’de gösterilmektedir (Alterio vd., 2009).



Şekil 3.5: CA IX dimer modeli.

Karbonik Anhidraz Enziminin Katalitik Mekanizması

CA’nın tüm izoenzimlerinin aktif bölgesinde Zn^{+2} iyonu ve ona bağlı bir su/hidroksil grubu mevcuttur. Aktif bölgenin civarındaki aminoasitler, proton verici ve proton gradienti oluşturacak şekilde düzenlenmiştir. Zn^{+2} ’ye bağlı H_2O , Glu-106 rezidüsünün (CAII’nin sekansına göre numaralandırma) karboksilat grubuna sırayla köprü oluşturan Thr-109 rezidüsünün hidroksil grubuyla, hidrojen bağı etkileşimleri sonucu tutunmaktadır. Bu etkileşimler, Zn^{+2} ’e bağlı H_2O molekülünün nükleofilitesini arttırmakta ve molekül nükleofilik atak için uygun bir yerdeki CO_2 molekülüne doğru hareket etmektedir (Supuran, 2009). Oluşan zincirleme reaksiyon Şekil 3.6’da gösterilmektedir.



Şekil 3.6: CA izoenzimlerinin CO₂ hidrasyonu katalitik mekanizması.

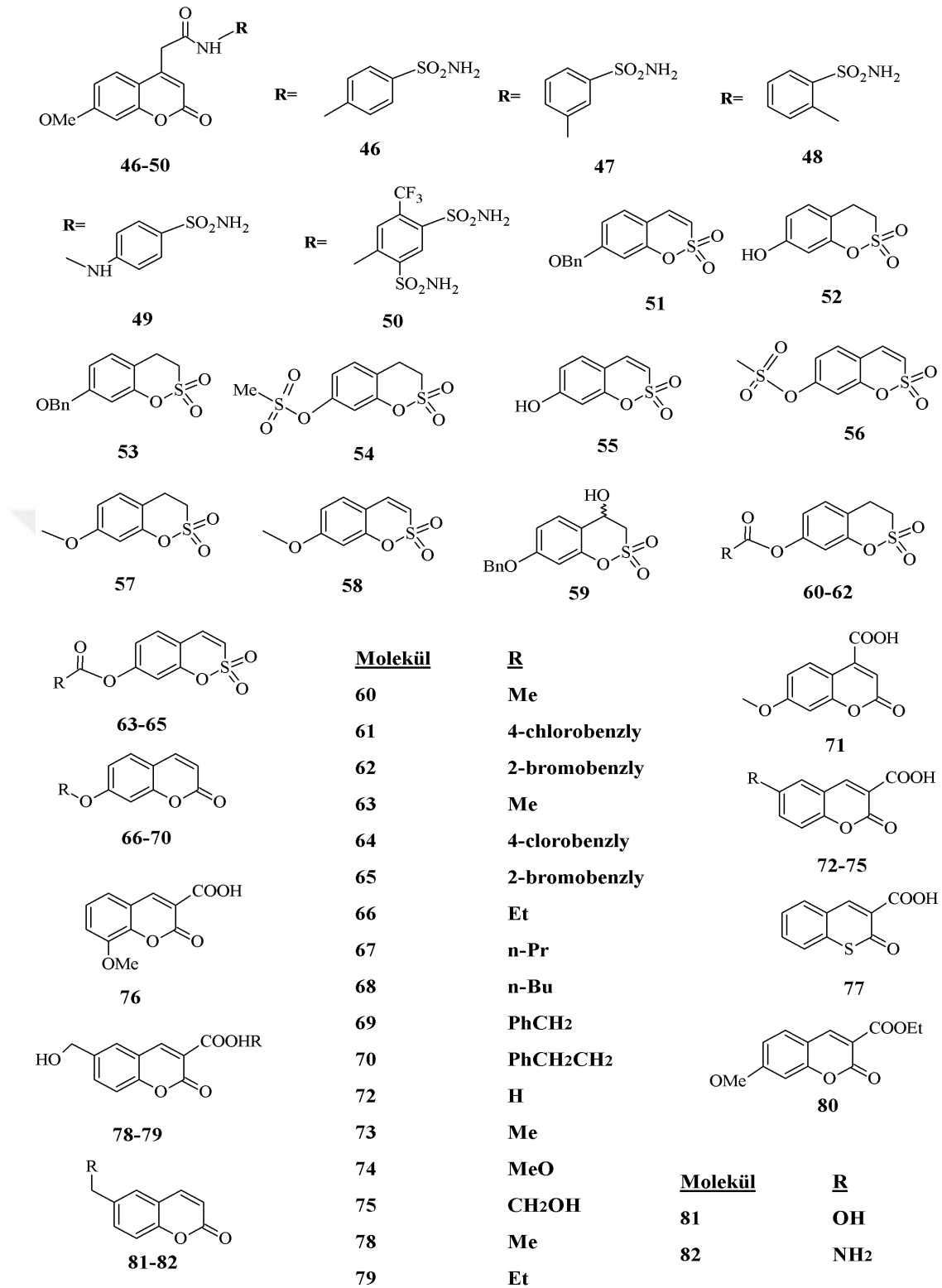
CA inhibitörü sülfonamidler aktif site içerisinde bulunan Zn atomuna bağlı su/hidroksil iyonu yerine geçerek inhibisyonu gerçekleştirirken kumarinler aktif site içerisine yerleştiğinde bulkmedia ile Zn atomu arasında oluşan proton transfer mekanizmasını felç ederek inhibisyon gerçekleştirmektedir.

3.3 Karbonik Anhidraz İnhibitörü Kumarinler

Kumarinlerin CA inhibitörü olduğu ilk defa Maresca vd. tarafından 2009 yılında keşfedilmiştir. O günden günümüze 80'den fazla kumarin türevinin CA'nın farklı izoenzimlerine karşı inhibisyon sabiti (K_i) verileri ölçülmüş ve yayınlanmıştır. Yedi farklı kaynaktan topladığımız 82 iki adet kumarin türevinin iki boyutlu yapıları Şekil 3.7'de gösterilmektedir (Maresca vd. 2009; Maresca ve Supuran 2010; Maresca vd. 2010a; Tars vd. 2013; Wagner vd. 2010; Tanc vd. 2013; Maresca vd. 2010b). Bu yapılara ait deneysel yolla ölçülmüş inhibisyon katsayısı değerleri Çizelge 3.1'de listelenmiştir. Çizelge 4.1'de verilen inhibisyon katsayısı değerleri alıntı yapılan makalelerde nano-mol veya mikro-mol olarak verilmektedir. Çizelgeye yerleştirilirken mikro-mol olanları da nano-mol'a çevirildi ve sonra logaritmaları alındı. Çizelge 3.1'de birçok kumarin için CAII veya CA IX için aktivite verisi yoktur. Çizelgede mevcut olmayan veriler genelde 100 mili-mol'den daha büyük olduğundan verilerin alındığı makalelerde nümerik bir değer verilmemektedir. Bir başka deyişle bir kumarin türevinin CAII ya da CAIX izoenzimine karşı K_i değeri ile Çizelge 3.1'de yoksa o ölçülmediği anlamına gelmiyor.

1	2	3	4
5	6	7	8
			Molekül No
9	10	11-18	R
		11	6-OH
19-28	29-30	12	6-MeSO₃
		13	6-BnO
		14	5,6-benzo
		15	6-Br
		16	6-NO₂
		17	6-NH₂
		18	6,8-Cl₂
		19	Ph
		20	COOMe
		21	COOEt
		22	SiMe₃
		23	HOCH₂
		24	Et₂NCH₂
		25	4F₃CO-C₆H₄
		26	4MeO-C₆H₄
		27	3F₃C-C₆H₄
		28	3MeO-C₆H₄
		29	Me
		30	Et
		31	Me
		32	Et

Şekil 3.7: 82 adet Karbonik anhidraz inhibitörü kumarin türevinin yapısı.



Şekil 3.7 (devam): 82 adet Karbonik anhidraz inhibitörü kumarin türevinin yapısı.

Çizelge 3.1: Yapıları şekil 3. 7’de verilen kumarin türevlerinin deneysel ölçülmüş inhibisyon katsayısı verileri

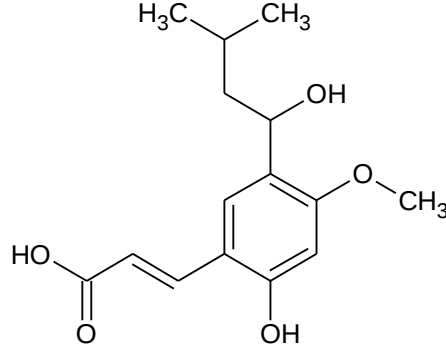
Molekül No	Log K _i (nanoMol)	
	CA II	CAIX
1	1,778151	4,736397
2	3,963788	#
3	1,845098	5,884795
4	#	2,705864
5	#	2,621176
6	#	2,296665
7	#	2,683047
8	#	2,679428
9	#	2,555094
10	#	3,691965
11	#	2,477121
12	#	2,510545
13	#	2,439333
14	#	2,574031
15	#	3,834421
16	#	3,576341
17	3,948902	1,662758
18	#	3,513218
19	3,889862	1,462398
20	3,801404	1,977724
21	3,96426	1,934498
22	3,857332	1,778151
23	3,968016	1,763428
24	3,97174	1,39794
25	3,984077	1,869232
26	3,970812	1,255273
27	3,887617	1,681241
28	3,935003	1,690196
29	#	3,904716
30	#	3,903904
31	#	1,919078
32	#	1,764923
33	#	3,892095
34	#	3,869232
35	#	3,879669
36	#	#
37	#	#
38	#	1,893762
39	#	1,850033
40	#	1,753583
41	#	1,786751
42	#	1,859138
43	#	1,577492
44	#	1,669317
45	#	0,716003
46	0,954243	1,939519
47	1,20412	1,770852
48	1,897627	2,053078

Çizelge 3.1 (devam): Yapıları şekil 3. 7’de verilen kumarin türevlerinin deneysel ölçülmüş inhibisyon katsayısı verileri

Molekül No	Log K _i (nanoMol)	
	CA II	CAIX
49	1,079181	1,959041
50	1,556303	1,146128
51	0,39794	#
52	0,880814	#
53	0,30103	#
54	0,255273	#
55	0,924279	#
56	0,322219	#
57	0,380211	#
58	0,176091	#
59	0,361728	#
60	0,079181	#
61	0,342423	#
62	0,531479	#
63	0,342423	#
64	0,414973	#
65	0,278754	#
66	4,093422	3,20412
67	4,161368	3,518514
68	5,32838	3,653213
69	5,350248	3,230449
70	5,385606	3,591065
71	1,995635	#
72	4,650308	3,672098
73	4,628389	3,886491
74	4,184691	3,819544
75	4,64836	3,748188
76	4,478566	3,869232
77	3,792392	1,672098
78	1,50515	1,653213
79	4,703291	1,672098
80	4,49693	3,812913
81	4,510545	1,968483
82	3,568202	3,826075

Şekil 3.7’de verilen kumarin türevleri için moleküler tanımlayıcılar hesaplanırken iki farklı yapı üzerinden hesaplamalar gerçekleştirildi. Çünkü kumarin türevleri CA ile reaksiyona Şekil 3.7’de gösterildiği yapılarında başlamasına rağmen CA aktif sitesi içerisinde hidrolize olarak kumarinlerin lakton halkasının açıldığına dair PDB’de üç adet (3F8E, 5BNL ve 4WL4) yapı mevcut. 1 nolu kumarin türevinin karboksilik asit formundaki yapısı Şekil 3.8’de gösterilmektedir. Dolayısıyla tüm kumarin

türevlerinin lakton halkaları açılarak karboksilik asit formunda da moleküler tanımlayıcılarını hesaplayarak QSAR modelleri oluşturuldu.



Şekil 3.8: Şekil 4.7’de verilen 1 no’lukumarin türevinin karboksilik asit formu.

Karboksilik asit formundan elde edilen QSAR modellerinin genel olarak daha iyi sonuçlar verdiği Bulgular ve Tartışma bölümünde görülecektir.

3.4 QSAR (Kantitatif Yapı-Etki İlişkisi) Analizleri

İngilizce bir terim olan QSAR ifadesi, “Quantitative Structure-Activity Relationship” baş harflerinin bir araya gelmesiyle oluşmuş ve Kantitatif Yapı-Etki İlişkilerini açıklaması ile dünyada kullanılır hale gelmiştir (Stanford, 2008).

QSAR analizi; bir grup molekülün kimyasal bileşik yapıları ile bu yapıların daha önceden gözlemlenmiş olan aktivite değerleri (örneğin; inhibisyon katsayısı, K_i) arasındaki ilişkiyi matematiksel ve istatistiksel yöntemlerle tayin ederek elde edilen modellerin yeni bileşiklerindeki aktivite değerlerini tahmin edebilme öngörüsü sağlar. Moleküllerin yapıları moleküler tanımlayıcı denem nümerik rakamlar ile temsil edilir. Molekül tanımlayıcılar çok farklı yollardan elde edilebilir. Bunlar pek çok sınıfa ayrılmakla beraber en yaygın ve kabaca bir sınıflama yapılacak olsa başlıca sayılacaklar arasında topolojik, geometrik, elektronik, elektrostatik ve orbital enerjisi ilişkili moleküler tanımlayıcılar sayılabilir. Her türlü moleküler tanımlayıcı fiziksel yorumlamaya uygun değildir. Yorumlamaya uygun moleküler tanımlayıcılar ile QSAR modellerinin oluşturulması aktivite mekanizmalarının açıklanmasının da büyük öneme sahiptir.

Başlangıç tarihi net olarak bilinmemesine karşın 1860’lardan 1960’lara kadar gelişim sürecinde olan QSAR çalışmalarının kökeni Cros’un, (1863) toxoloji alanında yapmış olduğu primer alifatik alkollerin toxilojisi ile suda çözünabilirliği arasındaki ilişkiyi önermesi ile olmuştur. Benzer şekilde, Cum Brown ve Froser (1868)’de kimyasal

yapı ve fizyolojik etki arasındaki bağlantıyı öne sürmüşlerdir. Richet, (1893) yılı, Meyer, (1899) yılı ve Overton, (1901) yılında ayrı ayrı lipofilik ve biyolojik etki arasında doğrusal bir bağlantı keşfetmişlerdir.

QSAR analizlerinin yaygın olarak kullanıldığı pek çok amaç var ancak bunlardan bazıları (Leszczynski, 2010);

- Rasyonel yöntemlerle biyolojik aktivite, fiziko-kimyasal gibi özellikleri tahmin edebilme
- Kimyasal bileşik içinde yer alan döngülerin rasyonelize edilmesi ve kavranması
- Kimyasal bileşiklerin hedef (reseptör, enzim) ile arasındaki etkileşmelerde rol oynayan dinamiklerin belirlenmesidir.

Bu amaçlar doğrultusunda oluşturulan modellerin geliştirilme sebepleri;

- İlaç, pestisit, kişisel ürünler gibi ürünlerin geliştirilme maliyetinde tasarruf edilmesi
- Tahminlerin uzun ve pahalı olan hayvan testlerindeki gereksinimi azaltabilir olmasıdır.

3.5 Çoklu Doğrusal Regresyon (ÇDR) Analizi

QSAR modeli oluşturmada yaygın olarak kullanılan ÇDR analiz yöntemi, matematiksel tekniklerle değişkenler arasındaki lineer ilişkiyi bulmak için kullanılan en yaygın analizdir (Leach ve Gillet, 2007). Bu analizdeki en basit eşitlik;

$$y = mx + c \quad (3.2)$$

Burada y; bağımlı değişken, x ise; bağımsız değişkendir. Lineer regresyonun amacı; denklemde tahmin edilen değerler ile gerçekte gözlemlenen değerler arasındaki farkı minimize eden c sabiti ve m katsayısını bulmaktır. m ve c değerlerinin eşitlikleri aynı zamanda şu şekildedir;

$$m = \frac{\sum_{i=1}^N (x_i - \langle x \rangle)(y_i - \langle y \rangle)}{\sum_{i=1}^N (x_i - \langle x \rangle)^2} \quad (3.3a)$$

$$c = \langle y \rangle - m\langle x \rangle \quad (3.3b)$$

Regresyon denklemi ile tanımlanan bu grafik, sırasıyla bağımsız ve bağımlı değişkenler olan x ve y ; (x,y) noktalarıyla açıklanır.

$$\langle x \rangle = \frac{1}{N} \sum_{i=1}^N x_i \quad (3.4.a)$$

$$\langle y \rangle = \frac{1}{N} \sum_{i=1}^N y_i \quad (3.4.b)$$

Denklem 3.4 de belirtilen bu durum birden fazla bağımsız değişken olduğunda kullanılan çoklu lineer regresyon analizi için geçerli olup QSAR çalışmalarında yaygın olarak kullanılan matematiksel tekniklerden biridir. Yapılan bu QSAR modelinde y bir özellik olup (biyolojik aktivite gibi); x ise moleküler tanımlayıcı olacaktır. QSAR modelini çözümleyen bu analizin istatistikî açıdan kabul edilebilir ve güvenilir olması gerekir.

Korelasyon Katsayı Karesi, R^2

Basit ve çoklu lineer regresyon analizlerinde ‘nitelik’ bulunduran en yaygın değerlendirmelerden birisi korelasyon katsayısı yani R^2 ’dir. R^2 , 0 ile 1 aralığında değere sahip olurken varyasyon oranını regresyon denklemi ile tanımlanan bağımlı değişken açıklar. Regresyon denklemindeki bağımsız değişkenler i tane değere sahip ise hesaplanan y_{calc} değeri; y_i deneysel değerine karşılık gelir ve bu değer ile hesaplanan eşitlikler şu şekildedir;

Toplam Kareler Toplamı,

$$TSS = \sum_{i=1}^N (y_i - \langle y \rangle)^2 \quad (3.5)$$

Açıklanan Kareler Toplamı,

$$ESS = \sum_{i=1}^N (y_{calc,i} - \langle y \rangle)^2 \quad (3.6)$$

Rezidüel Kareler Toplamı,

$$RSS = \sum_{i=1}^N (y_i - y_{calc,i})^2 \quad (3.7)$$

Sonuç olarak ise;

$$TSS = ESS + RSS \quad (3.8)$$

Bu ifadelere bağlı olarak R^2 eşitliği ise;

$$R^2 = \frac{ESS}{TSS} = \frac{TSS-RSS}{TSS} = 1 - \frac{RSS}{TSS} \quad (3.9)$$

R^2 değerinin 0'a karşılık gelmesi değişkenler arasında korelasyonun zayıf olduğuna tekabül ederken, değer 1'e karşılık gelmesi korelasyonun mükemmel olduğu anlamına gelir. Pek çok istatistiksel uygulamalarda kullanılan bu korelasyon katsayısının QSAR modellemesinde geçerliliği ≥ 0.6 olup tek başına yanıltıcı olabilir.

Korelasyon Belirleme Katsayısı, R_{cv}^2

Çapraz geçerlilik metodu regresyon analizi sonucunda korelasyon denklemindeki doğruluk ve güvenilirliği saptamak için kullanılır. Çapraz geçerlilik, bir veri setindeki bazı değerleri kaldırarak kalan değerlerle QSAR modelini türetir daha sonra bu modelde kaldırdığı değerleri uygulama da tahmin eder. Bunun en kolay uygulaması içerisinden tek bir veri çıkarılan leave-one-out (LOO) yaklaşımıdır. Veri setindeki her bir değer için bu metod uygulandığında R_{cv}^2 (ya da Q^2) değeri açığa çıkar. İstatistiksel analizler sonucunda açığa çıkan Q^2 değeri R^2 değerinden küçük olmasına rağmen regresyon denklemini tahmin etmede daha elverişlidir. Sonuç olarak R^2 uygunluğu ölçerken, Q^2 tahmin edilebilirliği ölçer. Basit bir LOO yaklaşımı için dahi Q^2 değeri hesaplanırken veri seti dört ya da beş gruba ayrılıp her defasında farklı bir grubun dışarıda bırakılması ile bu işlem birçok kez tekrarlanır.

Bazı durumlarda model oluşturulduktan sonra sentezlenen moleküllerle modelin doğruluğu test edebilmek PRESS (Tahmin Edilen Residüel Kareler Toplamı) ile olur. RSS ile benzerdir ancak burada kullanılan değerler oluşturulan modeldeki $y_{calc,i}$ değil PRESS kullanılarak bulunan $y_{pred,i}$ değerleridir ve Q^2 ile bağlantılı olan eşitliği şu şekildedir;

$$Q^2 = 1 - \frac{PRESS}{\sum_{i=1}^N (y_i - \bar{y})^2}; \quad PRESS = \sum_{i=1}^N (y_i - y_{pred,i})^2 \quad (3.10)$$

Diğer Regresyon Parametreleri

Standart sapma

Regresyon analizi neticesinde regresyon denkleminde gözlenen verilerle korelasyon mertebesini belirlememize yardımcı olan bir başka parametre de standart sapma (s)'dır. Türetilen QSAR modelindeki gözlemlenen verilerin ne kadar iyi tahmin edilebildiğini ölçer. Bu durum s değeri ne kadar küçük ise QSAR modelinin o denli iyi olduğunun bir göstergesidir.

$$s = \sqrt{\frac{RSS}{N-p-1}} \quad (3.11)$$

Bir lineer regresyonu değerlendirmek için kullanılan istatistiki testler; F (Fisher) ve t testleridir.

Fisher Testi

$$F = \frac{ESS}{p} \cdot \frac{N-p-1}{RSS} \quad (3.12)$$

Bu denklem formu her bir parametre ile ilişkili olan serbestlik derece sayısını yansıtır. ESS, p serbestlik derecesi ile ilişkili olup RSS, N-p-1 serbestlik derecesi ile ilişkilidir.

F için hesaplanan bu değer ile istatistiksel Çizelgede yer alan değerler karşılaştırılır. Çizelgede yer alan F değerleri her bir değişken için farklı bir güvenilirlik seviyesini belirtir. Model oluşturulurken hesaplanan denklem 3.12'teki F değeri Çizelgede yer alan F değerlerinden büyük ise söz konusu olan modelin güven düzeyinin istatistiki açıdan anlamlı olduğunu gösterir.

t Testi

$$t = \left| \frac{k_i}{s(k_i)} \right|; s(k_i) = \sqrt{\frac{RSS}{N-p-1} \cdot \frac{1}{\sum_{j=1}^N (x_{ij} - \bar{x}_i)^2}} \quad (3.13)$$

t istatistiği doğrusal regresyon denklemindeki her bir katsayının (s; standart hata gibi) önemini gösterir. F istatistiğinde de olduğu gibi denklem 3.13'deki t değeri, istatistiksel olarak listelenen Çizelgedeki değerlerle karşılaştırılır; hesaplanan bu değer Çizelgede yer alan değerlerden büyük ise katsayının anlamlı olduğunu gösterir.

3.6 Moleküler Tanımlayıcı Hesaplamada Yoğunluk Fonksiyonel Teorisi

QSAR modeli oluşturabilmek için moleküler tanımlayıcı hesaplamada DFT yaygın olarak kullanılan kuantum kimyasal bir yöntemdir (Foresman ve Frisch, 1996). Bu yöntem ile optimize edilen bir yapıdan yine aynı yöntem kullanılarak pek çok fiziksel yorumlamaya müsait moleküler tanımlayıcı hesaplanabilir. Bunlar arasında elektrostatik potansiyel yüzeylerinden elde edilen iso-yoğunluk tanımlayıcıları, orbital enerjisi değerleri (Homo, Lumo vs.) ve orbital yoğunluk haritaları sayılabilir.

Schrödinger Denklemi

Kuantum mekaniği, elektron gibi yapıların hem parçacık hemde dalga gibi nasıl davrandığını açıklar. Schrödinger denklemi bir parçacığın dalga fonksiyonunu tanımlar ve eşitliği şu şekildedir:

$$\left\{ \frac{-\hbar^2}{8\pi^2 m} \nabla^2 + V \right\} \Psi(\vec{r}, t) = \frac{i\hbar}{2\pi} \frac{\partial \Psi(\vec{r}, t)}{\partial t} \quad (3.14)$$

Bu denklemde Ψ ; dalga fonksiyonu, m; parçacık kütlesi, h; Plank sabiti ve V; hareket eden parçacığın potansiyel fonksiyonudur. $\Psi^* \Psi$ ($|\Psi|^2$) parçacığın olasılık dağılımıdır. Bir molekül gibi parçacıklar toplamı için Schrödinger denklemi çok benzerdir. Bu durumda Ψ , t zamanda sistemdeki tüm parçacıkların koordine bir fonksiyonu olur.

Parçacığın enerjisi ve birçok diğer özelliği, uygun sınır koşullarına bağlı olarak Ψ için Schrödinger denkleminin çözümü ile elde edilebilir. Dalga fonksiyonu sistemdeki birçok farklı duruma göre çözümlenebilir.

V zamanın bir fonksiyonu değilse; Schrödinger denklemi değişkenlerine ayrılarak basitleştirilebilir. Buna bağlı olarak uzaysal ve zamana bağlı dalga fonksiyonu;

$$\Psi(\vec{r}, t) = \Psi(\vec{r})\tau(t) \quad (3.15)$$

gibi iki denklem elde edilir; birincisi zamandan bağımsız parçacığın durum fonksiyonu, diğeri ise sadece zamana bağlı olan fonksiyondur.

$$H\Psi(\vec{r}) = E\Psi(\vec{r}) \quad (3.16)$$

E, parçacığın enerjisidir. H, Hamiltonyen operatörü olup, eşitliği şu şekildedir:

$$H = \frac{-\hbar^2}{8\pi^2 m} \nabla^2 + V \quad (3.17)$$

Denklem 3.15, parçacığın farklı durumlarına bağlı olarak birkaç farklı şekilde çözümlenir. İlki en düşük enerjiyle taban durumudur. Denklem 3.17 sistemin rölativistik olmadığı durumlarda ışık hızına yakın hızlarda hareket eden parçacıklarda geçerli değildir.

Denklem 3.16 aynı zamanda bir özdeğer eşitliğidir. Özdeğer eşitliği; bir eşitlikteki operatör fonksiyonu etkilerse fonksiyonun kendi katlarını üretir ve bunun genel formu,

$$Opf = cf \quad (3.18)$$

Op; operatör, f; fonksiyon ve c; sabit. Denklemde bulunan fonksiyon grubunun her bir öz fonksiyonlarının c sabiti ile ilişkili olarak bir öz değerleri vardır. Bu durumda Schrödinger denklemindeki öz değerler, moleküler sistemdeki farklı sabit durumlardaki enerjilerdir.

Moleküler Hamiltonyen

Bir moleküler sistem için Ψ , bir moleküldeki çekirdek ve elektronların konumu sırasıyla r ve R olarak gösterilen bir fonksiyondur. Bu semboller her bir parçacığın konumunu gösteren bileşke vektörleridir. Belirli bir elektron ya da çekirdeğe karşılık gelen vektör indis ile temsil edilir; r_i ve R_i .

Hamiltonyen, kinetik ve potansiyel terimlerinden oluşur;

$$H = T + V \quad (3.19)$$

Kinetik enerji, molekül içindeki tüm parçacıkların ∇^2 üzerinde toplamıdır:

$$T = \frac{-\hbar^2}{8\pi^2} \sum_j \sum_{k < j} \frac{1}{m_k} \left(\frac{\partial^2}{\partial x_k^2} + \frac{\partial^2}{\partial y_k^2} + \frac{\partial^2}{\partial z_k^2} \right) \quad (3.20)$$

Potansiyel enerji bileşeni, her bir yüklü çift arasında Coluomb itmesinin olmasıdır:

$$V = \frac{1}{4\pi\epsilon_0} \sum_j \sum_{k < j} \frac{e_j e_k}{\Delta r_{jk}} \quad (3.21)$$

Δr_{jk} ; iki parçacık arasındaki mesafe, e_j ve e_k ; j ve k parçacıklarının yükleri. Elektron için yük -e, çekirdek için ise Ze olup Z; atom numarasıdır. Sonuç olarak,

$$V = \frac{1}{4\pi\epsilon_0} \left(- \sum_i^{\text{çekirdek elektronları}} \sum_l \left(\frac{Z_l e^2}{\Delta r_{il}} \right) + \sum_i^{\text{elektronlar}} \sum_{j < i} \left(\frac{e^2}{\Delta r_{ij}} \right) + \sum_l^{\text{çekirdek}} \sum_{j < l} \left(\frac{Z_l Z_j e^2}{\Delta r_{lj}} \right) \right) \quad (3.22)$$

Birinci terim elektron-çekirdek etkileşimi, ikinci terim elektron-elektron itmesi ve üçüncü terim ise çekirdek-çekirdek itmesidir.

Kuantum kimyasının temel denklemleri genellikle temel sabitleri ortadan kaldırarak onların formunu basitleştirmek için oluşturulmuş birimlerle ifade edilir. Bohr yarıçapının atomik birimdeki uzunluğu;

$$a_0 = \frac{\hbar^2}{4\pi^2 m_e e^2} = 0.52917725 \text{ Å} \quad (3.23)$$

Koordinatlar, Bohr'un a_0 yarıçapları tarafından bölünerek ifade edilir. Enerjiler, 1 Bohr yarıçap mesafesi kadar uzaklıkta bulunan iki elektron arasındaki Coluomb itmesi olarak tanımlanan Hartrees ile ölçülür;

$$1 \text{ hartree} = \frac{e^2}{a_0} \quad (3.24)$$

Kütleler elektronik kütle birimi tarafından ifade edilir (örneğin; $m_e=1$). Diğer bölümlerde yer alan denklemlerde bu birimler kullanılacaktır.

Born-Oppenheimer Yaklaşıklığı

Born-Oppenheimer yaklaşıklığı Schrödinger denkleminin çözümünü basitleştirmek için kullanılan birçok yaklaşıklığın ilkidir. Bu yaklaşıklık çekirdek ve elektronların hareketlerini birbirinden ayırarak moleküler problemi basite indirgeyip elektrondan binlerce kat daha büyük olan çekirdek kütleleri için geçerlidir. Çekirdek, elektronlara kıyasla daha yavaş hareket ederken buna karşılık elektronlar hareketlerinden dolayı çekirdek etrafındaki yörüngelerinde ani değişiklikler sergilerler. Bu yüzden bir moleküler sistemdeki elektronların konumu, çekirdekteki hızına değil konumuna bağlıdır.

Bir moleküler sistem için Hamiltonyen eşitliği;

$$H = T^{elekt.}(\vec{r}) + T^{çek.}(\vec{R}) + V^{çek.-elekt.}(\vec{R}, \vec{r}) + V^{elekt.}(\vec{r}) + V^{çek.}(\vec{R}) \quad (3.25)$$

Born-Oppenheimer yaklaşığı problemi bağımsız iki parça halinde çözümlediğinden çekirdek için kinetik enerji terimini ihmal ederek Hamiltonyen eşitliği elektronlar için düzenlenirse;

$$H^{elekt.} = -\frac{1}{2} \sum_i^{elektron} \left(\frac{\partial^2}{\partial x_i^2} + \frac{\partial^2}{\partial y_i^2} + \frac{\partial^2}{\partial z_i^2} \right) - \sum_i^{elektron} \sum_I^{çekirdek} \left(\frac{Z_I}{|\vec{R}_I - \vec{r}_i|} \right) + \sum_i^{elektronlar} \sum_{j < i} \left(\frac{1}{|\vec{r}_i - \vec{r}_j|} \right) + \sum_I^{çekirdek} \sum_{J < I} \left(\frac{Z_I Z_J}{|\vec{R}_I - \vec{R}_J|} \right) \quad (3.26)$$

Sabit çekirdek alanındaki elektron hareketini tanımlayan bu Hamiltonyen eşitliği Schrödinger denkleminde yazılırsa;

$$H^{elekt.} \Psi^{elekt.}(\vec{r}, \vec{R}) = E^{etkin}(\vec{R}) \Psi^{elekt.}(\vec{r}, \vec{R}) \quad (3.27)$$

Elektronik dalga fonksiyonu için bu denklemin çözümü ile E^{etkin} elde edilecektir. E^{etkin} ; sistemin potansiyel enerji yüzeyini tanımlar ve çekirdek koordinasyonlarına bağlıdır.

Çekirdek Hamiltonyen eşitliği için etkin potansiyel;

$$H^{çek.} = T^{çek.}(\vec{R}) + E^{etkin}(\vec{R}) \quad (3.28)$$

Bu eşitlik Schrödinger denkleminde çekirdeklerin titreşim, öteleme ve dönme durumları için kullanılır. Schrödinger denkleminin çözümü, moleküllerin titreşim spektrumunun bulunması için gereklidir.

Dalga Fonksiyonundaki Kısıtlamalar

Ψ^2 , parçacık ya da parçacıkların olasılıkları dağılımıdır. Eğer parçacık olasılığı uzaysal boyutta entegre ediliyorsa Ψ 'nin normalize olması gerekir.

$$\int_{-\infty}^{+\infty} |\Psi|^2 dv = n_{\text{parçacık}} \quad (3.29)$$

Schrödinger denklemi bir öz değer denklemidir ve herhangi bir c değeri için, cf kadarıyla, f gibi bir özdeğer denklemin çözümü olur. Schrödinger denklemi için $H(c\Psi)=cH(\Psi)$ ve $E(c\Psi)=cE(\Psi)$ ifadeleri kullanılırsa; sonuç olarak cΨ kadarıyla denklemin çözümü Ψ olur.

Ψ, aynı zamanda antisimetrik de olabilir yani bu durum iki özdeş parçacık yer değiştirdiği zaman işaretle değiştirebilir anlamını verir.

$$f(i,j) = -f(j,i) \quad (3.30)$$

Elektronik dalga fonksiyonu için antisimetri durumu, elektronların fermiyon olmasından kaynaklanır. Ψ, fizik deneyleri sonucunda herhangi bir dalga fonksiyonunu oluşturan durum olarak da tanımlanabilir;

$$\Psi(\vec{r}_1, \dots, \vec{r}_i, \dots, \vec{r}_j, \dots, \vec{r}_n) = -\Psi(\vec{r}_1, \dots, \vec{r}_j, \dots, \vec{r}_i, \dots, \vec{r}_n) \quad (3.31)$$

Hartree-Fock Teorisi

Schrödinger denkleminin tam olarak çözümü en küçük moleküler sistem için bile mümkün değildir. Fakat varsayımları ve işlemleri daha basite indirmek, moleküllerin büyük bir oranda çözömlenme olasılığı bir yaklaşıklık ile yapılabilir.

Moleküler Orbitaler

İlk yaklaşıklık, bir sistemdeki elektronlar için yoğunluk olasılığı $|\Psi|^2$ nin yorumlanmasıdır. Moleküler orbital teorisi Ψ yi moleküler bileşenlerine ayırır: $\phi_1, \phi_2, \dots$ Bu ortogonal setin Ψ şartlarını sağlaması için öncelikle normalize olmalıdır;

$$\iiint \phi_i^* \phi_j d_x d_y d_z = 1 \quad (3.32.a)$$

$$\iiint \phi_i^* \phi_j d_x d_y d_z = 0 ; i \neq j \quad (3.32.b)$$

Bu moleküler orbitallerin bir araya gelerek Ψ oluşturmaları en basit olası yol olan Hartree yöntemi ile olur;

$$\Psi(\vec{r}) = \phi_1(\vec{r}_1) \phi_2(\vec{r}_2) \dots \phi_n(\vec{r}_n) \quad (3.33)$$

Ancak bu tip fonksiyonlar antisimetrik olmadıkları için dalga fonksiyonu oluşturmak için yetersizdir.

Elektron Spini

Moleküler orbital bileşenlerini antisimetrik hale getirmek için öncelikle elektron spini faktörüne gereksinim vardır. Elektronların spin aşağı (-1/2) ve spin yukarı (+1/2) olmak üzere iki spini vardır. Denklem 3.33'de her bir moleküler orbitale bir elektron eşlik eder. Ancak çoğu hesaplamalarda kapalı kabuk (yani çiftleşmiş orbitaller) hesaplamaları olduğundan zıt spinleri bir arada tutar. Bu durumda sınırlar iki spin fonksiyonu olan α ve β olarak belirlenir;

$$\alpha(\uparrow) = 1 \quad \alpha(\downarrow) = 0 \quad (3.34.a)$$

$$\beta(\uparrow) = 0 \quad \beta(\downarrow) = 1 \quad (3.34.b)$$

α fonksiyonu bir spin yukarı elektron için 1, β fonksiyonu spin aşağı elektron içinde 1 dir. i tane elektron için α ve β değerleri $\alpha(i)$ ve $\beta(i)$ notasyonlarıdır.

Ψ , elektron dalga fonksiyonunun bir parçası olan elektron spini α ve β moleküler orbital fonksiyonlarını içerir. n elektronlu bir sistem için n/2 tane moleküler orbitalin elektronları zıt ve çift spinli olarak tanımlanmış bir kapalı kabuk dalga fonksiyonu oluşturulabilir;

$$\Psi(\vec{r}) = \frac{1}{\sqrt{n!}} \begin{vmatrix} \phi_1(\vec{r}_1)\alpha(1) & \phi_1(\vec{r}_1)\beta(1) & \phi_2(\vec{r}_1)\alpha(1) & \phi_2(\vec{r}_1)\beta(1) & \dots & \phi_{n/2}(\vec{r}_1)\alpha(1) & \phi_{n/2}(\vec{r}_1)\beta(1) \\ \phi_1(\vec{r}_2)\alpha(2) & \phi_1(\vec{r}_2)\beta(2) & \phi_2(\vec{r}_2)\alpha(2) & \phi_2(\vec{r}_2)\beta(2) & \dots & \phi_{n/2}(\vec{r}_2)\alpha(2) & \phi_{n/2}(\vec{r}_2)\beta(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_1(\vec{r}_i)\alpha(i) & \phi_1(\vec{r}_i)\beta(i) & \phi_2(\vec{r}_i)\alpha(i) & \phi_2(\vec{r}_i)\beta(i) & \dots & \phi_{n/2}(\vec{r}_i)\alpha(i) & \phi_{n/2}(\vec{r}_i)\beta(i) \\ \phi_1(\vec{r}_j)\alpha(j) & \phi_1(\vec{r}_j)\beta(j) & \phi_2(\vec{r}_j)\alpha(j) & \phi_2(\vec{r}_j)\beta(j) & \dots & \phi_{n/2}(\vec{r}_j)\alpha(j) & \phi_{n/2}(\vec{r}_j)\beta(j) \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \phi_1(\vec{r}_n)\alpha(n) & \phi_1(\vec{r}_n)\beta(n) & \phi_2(\vec{r}_n)\alpha(n) & \phi_2(\vec{r}_n)\beta(n) & \dots & \phi_{n/2}(\vec{r}_n)\alpha(n) & \phi_{n/2}(\vec{r}_n)\beta(n) \end{vmatrix} \quad (3.35)$$

Her satır i tane elektronun tüm spin-yörünge etkileşimlerinin olası karşılaştırmalarıdır. Bu eşitlikteki faktör ifadesinin normalize olması gerekir.

Yukarıdaki ifade sadece antisimetrik bir dalga fonksiyonu için geçerli bir eşitlik değildir. Kuantum mekaniği bir elektronun konumunu belirlemekten ziyade herhangi bir yerdeki yoğunluk olasılığını belirler. Bu durum dalga fonksiyonu oluşturmak için bir moleküler sistemde bulunan elektronların olası orbitallerinin karşılaştırılmasıdır.

Temel Setler

Bir sonraki yaklaşıklık temel fonksiyonlar olarak bilinen tek elektronlu fonksiyon setlerinde önceden tanımlanmış olan moleküler orbitallerdir. Bu fonksiyonlar genellikle atomik çekirdek merkezindedirler, bundan dolayı bazı atomik orbitaller benzerdir.

Bir moleküler orbital tanımı;

$$\Phi_i = \sum_{\mu=1}^N c_{\mu i} \chi_{\mu} \quad (3.36)$$

$c_{\mu i}$, moleküler orbital genişleme katsayısıdır. $x_1, x_2, \dots, x_n$ 'e kadar olan temel fonksiyonların normalize olmaları gerekir. Genellikle moleküler orbitaller roman simgeleriyle, temel fonksiyonlarda Yunan simgeleri ile temsil edilir. Bu yüzden x_{μ} temel fonksiyonu da ϕ_i moleküler orbital de keyfi seçilmiş ifadelerdir.

Gaussian ve diğer abinitio elektronik yapı programlarında temel fonksiyonlar olarak gaussian tipi atomik fonksiyonlar kullanılır. Gaussian fonksiyonlarının genel ifadesi;

$$g(\alpha, \vec{r}) = c x^n y^m z^l e^{-\alpha r^2} \quad (3.37)$$

r ; x , y ve z koordinatlarının oluşumudur. α , fonksiyonun radyal bileşenidir. Gaussian fonksiyonu $e^{-\alpha r^2}$, normalizasyon sabitidir ve

$$\int g^2 = 1 \quad (3.38)$$

c ise α , l , m ve n değerlerine bağlıdır.

Gaussian fonksiyonları s , p_y , ve d_{xy} türlerine bağlı olarak üç şekilde gösterilir;

$$g_s(\alpha, \vec{r}) = \left(\frac{2\alpha}{\pi}\right)^{3/4} e^{-\alpha r^2} \quad (3.39.a)$$

$$g_y(\alpha, \vec{r}) = \left(\frac{128 \alpha^5}{\pi^3} \right)^{1/4} y e^{-\alpha r^2} \quad (3.39.b)$$

$$g_{xy}(\alpha, \vec{r}) = \left(\frac{2048 \alpha^5}{\pi^3} \right)^{1/4} x y e^{-\alpha r^2} \quad (3.39.c)$$

Bunun gibi Gaussian fonksiyonlarının lineer birleşimleri temel fonksiyonlarda sınırlandırılmış gaussian olarak kullanılır;

$$\chi_\mu = \sum_p d_{\mu p} g_p \quad (3.40)$$

Buna bağlı olarak moleküler orbital eşitliği yeniden yazılır;

$$\phi_i = \sum_{\mu=1}^N c_{\mu i} \chi_\mu = \sum_{\mu=1}^N c_{\mu i} \left(\sum_p d_{\mu p} g_p \right) \quad (3.41)$$

Varyasyon İlkesi

Burada $c_{\mu i}$, katsayılarının nasıl çözümleneceğine değinilir. Hartree-Fock teorisinde bir moleküler sistemde taban durumunda olan herhangi bir normalize olmuş antisimetrik fonksiyonun E ifadesine karşılık gelen varyasyon ilkesinden yararlanılır ve bu beklenen enerji değeri ifadesi, bir tam dalga fonksiyonunun enerjisinden daha büyüktür;

$$E(\Xi) > E(\Psi); \Xi \neq \Psi \quad (3.42)$$

Bir başka ifadeyle, tam dalga fonksiyonunun enerjisi herhangi bir normalize antisimetrik fonksiyonun hesaplanmış daha düşük enerjisine tekabül edebilir. Sonuç olarak katsayılar kümesinde bulunan ifadelerden biri dalga fonksiyonundaki enerjinin minimize edilmesiyle çözümlenmiş olur.

Roothan-Hall Eşitlikleri

Varyasyon ilkesi Roothan-Hall tarafından türetilmiş c_{vi} katsayısı eşitliğini açıklar;

$$\sum_{v=1}^N (F_{\mu v} - \varepsilon_i S_{\mu v}) c_{vi} = 0 \quad \mu = 1, 2, \dots, N \quad (3.43)$$

Yukarıdaki eşitlik matris formuyla yeniden yazılacak olursa;

$$FC = SC\varepsilon \quad (3.44)$$

Buradaki her bir ifade matris olup; ε , enerji orbitallerinin köşegen matrisidir ve her bir ε_i ifadesi X_i moleküler orbitalin bir elektrondaki orbital enerjisidir.

F, Fock matrisidir ve her bir orbitaldeki bütün elektron alanlarının ortalama etkinliğini gösterir. Kapalı alan sistemleri için ifadeleri;

$$F_{\mu\nu} = H_{\mu\nu}^{\text{çekirdek}} + \sum_{\lambda=1}^N \sum_{\sigma=1}^N P_{\lambda\sigma} \left[(\mu\nu|\lambda\sigma) - \frac{1}{2}(\mu\nu|\lambda\sigma) \right] \quad (3.45)$$

$H_{\mu\nu}$, bir çekirdekteki tek elektronun alan enerjisinin matris gösterimi, P ise olasılık matrisidir;

$$P_{\lambda\sigma} = 2 \sum_{i=1}^{\text{dolü}} c_{\lambda i}^* c_{\sigma i} \quad (3.46)$$

s, orbitaller arasındaki örtüşmeyi gösteren matris ifadesidir. Hem Fock matrisi hem de orbitaller, moleküler genişleme katsayısına bağlı olduğundan Denklem 31 de yer alan eşitlik lineer değildir ve Öz Tutarlılık Kapsamı (SCF) metodu ile tekrarlamalı olarak çözümlenmelidir. Uygulanan bu çözüm sonucunda temel fonksiyonlardaki denklem sayısına eşit olacak şekilde hem dolu orbitaller ($\phi_{i,j}, \dots$) hem de boşorbitaller ($\phi_{a,b}, \dots$) üretilir.

Denklem 3.45 de yer alan $(\mu\nu|\lambda\sigma)$ ifadesi, Hartree-Fock işlemi içerisinde yer alan bir elektronun diğer tüm elektronlara olan ortalama uzaklığını gösteren; iki elektronun itme türevini açıklar.

Açık Kabuk Methodlar

Hartree-Fock metodu sadece sınırlandırılmış durumlarda kullanılan bir metoddur ancak Açık-Kabuk sistemleri, çiftleşmemiş elektronların çözümünde kullanılan sınırlandırılmamış bir metoddur. Bu durumda elektron spin orbitalleri α ve β yı iki farklı sette bulunan farklı moleküler orbital genişleme katsayıları olarak ele alırız;

$$\phi_i^\alpha = \sum_{\mu} c_{\mu i}^\alpha \chi_{\mu} \quad (3.47.a)$$

$$\phi_i^\beta = \sum_{\mu} c_{\mu i}^\beta \chi_{\mu} \quad (3.47.b)$$

Elektron Korelasyon Methodları

Hartree-Fock teorisi bir moleküler sistemdeki zıt ve çift spinli elektronların hareketleri arasındaki korelasyonunu tanımlamakta yetersizdir. Ancak Hartree-Fock teorisinin değişim korelasyonu olarak nitelendirilen yüksek değerlikte korelasyon şartını gerçekleştirebilmesi için antisimetrik dalga fonksiyonundaki herhangi iki aynı spinli elektronun karşılıklı alışverişi ile $|\Psi^2|$ 'nin değişmez olması gerekir.

SCF özelliği ile işlev gösteren herhangi bir metod *elektron korelasyon metodu* veya *post-SCF metodu* olarak bilinir.

Yoğunluk Fonksiyonel Teorisi (DFT)

Yoğunluk fonksiyonel teori tabanlı methodlar 1920'lerde özellikle Thomas-Fermi-Dirac modeli olan kuantum mekaniği araştırmalarında ve 1950'de Slater'in kuantum kimyasındaki temel çalışması sonucunda türetilmiştir. DFT yaklaşımı, elektron yoğunluğunun genel fonksiyonu aracılığıyla bir elektron korelasyonu modellemesi esasına dayanır.

Kökeni 1964'te yayınlanan Hohenberg-Kohn teoremine dayanan bu tür metotlar taban durum enerjisini ve yoğunluğunu tam olarak belirleyen bir tek fonksiyonun varlığını göstermesiyle yorumlanır. Ancak bu teorem fonksiyonelliğin eşitliğini sağlamaz.

Kohn ve Sham (1965)'in bunu takiben yaptıkları çalışmada yaklaşım fonksiyonelleri, mevcut DFT metodları tarafından elektronik enerjiyi birkaç terime bölerek sağladı:

$$E = E^T + E^V + E^J + E^{xc} \quad (3.48)$$

E^T , kinetik enerji (elektronların hareketi sonucu açığa çıkan enerji), E^V , potansiyel enerji (çekirdek çiftleri arasındaki itme ve çekirdek-elektron arasındaki çekmeden açığa çıkan enerji), E^J , elektron-elektron itme terimi (elektron yoğunluğunun Coulomb etkileşimi) ve E^{xc} , korelasyon değişim terimi ve elektron-elektron etkileşiminin geri kalan kısmını içerir.

Elektron yoğunluğundaki bütün bu terimlerin dışındaki çekirdek-çekirdek itmesi ρ fonksiyonlarıdır. E^J denklemi ile verilir:

$$E^J = \frac{1}{2} \iint \rho(\vec{r}_1) (\Delta r_{12})^{-1} \rho(\vec{r}_2) d\vec{r}_1 d\vec{r}_2 \quad (3.49)$$

$E^T + E^V + E^C$ terimi, ρ yük dağılımının toplam enerjisidir. E^{XC} terimi ise; dalga fonksiyonunun antisimetrik olmasından kaynaklanan enerjisi ve elektronların dinamik hareketidir. Bu terim Hohenberg ve Kohn tarafından elektron yoğunluğu fonksiyonunu ifade etmek için belirlenmiş olsa da pratikte sadece spin yoğunluklarının gradyen olasılıklarını içeren integral ifadelerinde kullanılır;

$$(E)^{XC}(\rho) = \int f(\rho_\alpha(\vec{r}), \rho_\beta(\vec{r}), \nabla \rho_\alpha(\vec{r}), \nabla \rho_\beta(\vec{r})) d^3\vec{r} \quad (3.50)$$

ρ_α , α spini yoğunluğu, ρ_β , β spini yoğunluğu ve ρ ise toplam elektron yoğunluğu ($\rho_\alpha + \rho_\beta$) ifadesidir.

E^{XC} ifadesini aynı spinlerin ve karma spinlerin etkileşimi ile iki parça şeklinde ele alacak olursak;

$$(E)^{XC}(\rho) = (E)^X(\rho) + (E)^C(\rho) \quad (3.51)$$

Belirtilmiş olan bu ifadeler sırasıyla değişim fonksiyonları ve korelasyon fonksiyonlarıdır. Kısmi değişim fonksiyonları sadece ρ (elektron yoğunluğu) bağlı iken, korelasyon fonksiyonları hem ρ ya hem de (elektron yoğunluğunun gradyeni) Δ_ρ ya bağlıdır. Kısmi değişim fonksiyonları genel olarak tanımlanacak olursa;

$$E_{LDA}^X = -\frac{3}{2} \left(\frac{3}{4\pi} \right)^{1/3} \int \rho^{4/3} d^3\vec{r} \quad (3.52)$$

ρ , r 'nin bir fonksiyonudur. Bu eşitlik serbest bir elektron enerjisini yeniden geliştirebilir ancak bir moleküler sistemdeki elektron için yetersizdir.

LDA değişim fonksiyonunun gradyen kısmına bağlı olarak Becke 1988 yılında bir formül geliştirmiştir;

$$E_{Becke88}^X = E_{LDA}^X - \gamma \int \frac{\rho^{4/3} x^2}{(1 + 6\gamma \sinh^{-1} x)} d^3\vec{r} \quad (3.53)$$

$x = \rho^{4/3} |\nabla \rho|$, γ durağan atom gazlarının enerji parametreleridir ve Becke tanımıyla değerleri 0,0042 Hartree'dir.

Benzer şekilde Perdew ve Wang, (1992) formüllerinin korelasyon fonksiyonlarının temel eşitliği;

$$E^C = \int \rho \varepsilon_C(r_s(\rho\vec{r}), \zeta) d^3\vec{r} \quad (3.54.a)$$

$$r_s = \left[\frac{3}{4\pi\rho} \right]^{1/3} \quad (3.54.b)$$

$$\zeta = \frac{\rho_\alpha - \rho_\beta}{\rho_\alpha + \rho_\beta} \quad (3.54.c)$$

$$\varepsilon_c(r_s, \zeta) = \varepsilon_c(\rho, 0) + a_c(r_s) \frac{f(\zeta)}{f''(0)} (1 - \zeta^4) + [\varepsilon_c(\rho, 1) - \varepsilon_c(\rho, 0)] f(\zeta) \zeta^4 \quad (3.54.d)$$

$$f(\zeta) = \frac{[(1+\zeta)^{4/3} + (1-\zeta)^{4/3} - 2]}{(2^{4/3} - 2)} \quad (3.54.e)$$

r_s , yoğunluk parametresi. ζ , spin polarizasyonu olup $\zeta=0$ durumu α ve β yoğunluklarının eşit olması, $\zeta=1$ durumu α yoğunluğuna ve $\zeta=-1$ durumu ise β yoğunluğuna karşılık gelir.

$\varepsilon_c(r_s, 0)$, $\varepsilon_c(r_s, 1)$ ve $-a_c(r_s)$ değerlerini hesaplamak için kullanılan G fonksiyonu kullanılır ve hesaplama yeniden düzenlenir;

$$G(r_s, A, \alpha_1, \beta_1, \beta_2, \beta_3, \beta_4, P) = -2A(1 + \alpha_1 r_s) \ln \left(1 + \frac{1}{2A(\beta_1 r_s^{1/2} + \beta_2 r_s + \beta_3 r_s^{3/2} + \beta_4 r_s^{P+1})} \right) \quad (3.55)$$

DFT methodları değişim fonksiyonları ile korelasyon fonksiyonlarının eşleşmesi olarak tanımlanır ki en iyi örneği BLYP fonksiyon çifti olan Becke değişim fonksiyonu ile Lee, Yang ve Parr, (1988) korelasyon fonksiyonlarının eşleşmesidir.

Hibrit Fonksiyoneller

Kohn ve Sham tarafından Hartree-Fock teorisine benzer şekilde uygulanan DFT hesaplamaları, analog bir şekilde tekrarlamalı SCF hesaplamaları ile gerçekleştirilir. E^{XC} olarak tanımlanan bu DFT hesaplamaları;

$$E_{hibrit}^{XC} = c_{HF} E_{HF}^X + c_{DFT} E_{DFT}^{XC} \quad (3.56)$$

c , sabit bir değerdir. Bu formülizasyonu 3 adet parametresi bulunan Becke fonksiyon parametreleri ile tanımlarsak;

$$E_{B3LYP}^{XC} = E_{LDA}^X + c_0(E_{HF}^X - E_{LDA}^X) + c_X \Delta E_{B88}^X + E_{VWN3}^C + c_C(E_{LYP}^C - E_{VWN3}^C) \quad (3.57)$$

Burada c_0 parametresi, Hartree-Fock ve LDA'nın konumsal değişimiyle hesaba katılır. c_X parametresi de LDA değişimi ile Becke'ningradyen bileşeni olarak ölçülür. c_C parametresi ise kullanılan VWN3 local fonksiyonunun LYP korelasyonu ile ölçülür. B3LYP fonksiyonunda G1 molekül setindeki atomizasyon enerjileri; iyonizasyon potansiyelleri, proton afiniteleri ve birinci düzey atom enerjileri olarak Becke tarafından belirlenen bu parametrelerin değerleri; $c_0=0.20$, $c_X=0.72$ ve $c_C=0.81$ dir.

Birçok farklı fonksiyonun da fonksiyon bileşenlerini değiştirerek aynı yöntemle parametre değerlerini bulabiliriz.

3.7 Moleküler Tanımlayıcılar

Bu bölümde, moleküler yapısı Şekil 3.7'de, CAII ve CAIX'a karşı inhibisyon aktiviteleri Çizelge 3.1'de verilen kumarin türevlerinden elde edilen, bir sonraki Bulgular ve Tartışma bölümünde verilen QSAR modellerinde yer alan moleküler tanımlayıcıların tanımları verilmektedir. Moleküler tanımlayıcılar birçok kategoriye ayrılabilir. Bunların hepsi fiziksel yorumlamaya uygun değildir. Bir sonraki Bulgular ve Tartışma bölümünde verilen QSAR modellerinin seçiminde özellikle tüm moleküler tanımlayıcılar yorumlamaya uygun olmasa da her bir modelde en az bir moleküler tanımlayıcının yorumlanabilir olmasına dikkat edildi.

Moleküler elektrostatik potansiyel tabanlı moleküler tanımlayıcılar

Bir molekül için elektrostatik potansiyel yüzeyini elde etmek elektrik ve manyetizma alanının temel yasası olan Coulomb yasasının moleküler sisteme uygulanmasına dayanır. Bu elektrostatik potansiyel yüzeyleri fiziksel gözlenebilir kuantiti olup hem deneysel hem de hesaplama yoluyla elde edilebilir. Bir molekül için elektrostatik potansiyel, $V(r)$ aşağıdaki denklem ile verilir (Murray ve Politzer, 2011).

$$V(r) = \frac{1}{4\pi\epsilon_0} \left[\sum_A \frac{Z_A e}{|Z_A - r|} - e \int \frac{\rho(r') dr'}{|r' - r|} \right] \quad (3.58)$$

Bu denklem, uzayda r konum vektörü ile verilen bir noktadaki elektrostatik potansiyeli ifade eder. Denklemde ki Z_{Ae} , konumu Z_A da yer alan A atomunun

çekirdek yükünü, $|Z_A - r|$ ise A atomunun çekirdeğinin r noktasına olan mesafesi gösterir. Denklemdaki $|r' - r|$ ise r noktasından her bir elektronik yük artırımındaki ($-e\rho(r')dr'$) mesafe olarak verilir. Burada Born–Oppenheimer yakınlığı kabul edilir. Bu yakınlıkta bir molekülde elektronların hızı çekirdeklerin hızından görece çık daha hızlı olduklarından molekülün elektrostatik enerjisi hesaplanırken çekirdekler durağan kabul edilir. Burada elektronların hızlı hareketinden bahsetmemize rağmen $V(r)$ ‘ye elektrostatik diyebilmemiz $\rho(r)$ ifadesinin bir hacim elamanındaki ortalama elektron yükü temsil etmesine dayanır. $V(r)$ uzayın herhangi bir noktasında pozitif ya da negatif olabilir. Bu uzayın o noktasında elektronların mı yoksa çekirdeklerin mi dominant olduğuna bağlıdır. $\rho(r)$ deneysel olarak x-ışınları veya nötron kırınımından elde edilebilir. Hesaplama yolu ile ise moleküler dalga fonksiyonundan elde edilir. Biz bu çalışmada her bir molekül için DFT yöntemi ile B3LYP fonksiyoneli ve 6-311 G (d,p) temel seti kullanarak moleküler dalga fonksiyonlarını elde ettik.

Yukarıdaki denklem 3.58’e göre $V(r)$ istenilen her noktada hesaplanabilir. Örneğin iki atom arasındaki mesafe (kimyasal bağ) boyunca veya iki boyutlu bir düzlemde hesaplamak mümkündür. Eğer hesaplama molekülün dış yüzeylerine yakın bölgelerde sık olarak seçilmiş gridler boyunca üç boyutlu olarak hesaplanırsa molekülün reaktivitesi hakkında çok değerli bilgilere ulaşmamızı sağlar. Bir molekülün reaktivitesi konusunda etkili olan elektronların ne olduğunun temsili konusunda bilgi edinebilmek için $\rho(r)=0.001$ elektron/bohr³ eş-yoğunluk yüzey haritaları yeterli görülmektedir. Bu eş-yoğunluk yüzey haritalarının bir molekülün toplam yükünün % 96’sını temsil ettiği bildirilmektedir (Bader vd. 1987).

Bu çalışmada yer alan her kumarin türevleri için DFT yöntemi ile B3LYP fonksiyoneli ve 6-311 G (d,p) temel seti kullanarak moleküler dalga fonksiyonlarını Gaussian 09 yazılımı kullanılarak elde edildi. Gaussian 09 çıktı (log) dosyaları Codessa 3 yazılımına giriş dosyası olarak atandı. Cadessa 3 yazılımı Gaussian 09 log dosyalarından aldığı moleküler dalga fonksiyonunu kullanarak her bir molekül için denklem 3.58’e dayalı olarak eş-yoğunluk yüzey haritalarını hesapladı. Bu eş-yoğunluk yüzeyi haritaları pozitif ve negatif yüzeyler içermektedir. Her bir molekülde bulunan bu yüzeylerin alan büyüklükleri, ektrim noktaları (maksimum pozitif, maksimum negatif) ve bu yüzeylerin gradyanları gibi birçok parametre ve bu parametrelerinin kareleri, birinin diğerine oranı gibi farklı kombinasyonlar hesaplanarak moleküler tanımlayıcı olarak kullanıldı. Cadessa3 bu parametrelerin

kombinasyonunu kullanarak 20 adet moleküler tanımlayıcı hesaplamaktadır. Bu moleküler tanımlayıcılardan bu çalışmada seçilen QSAR modellerinde yer alan eş-yoğunluk yüzey grubu moleküler tanımlayıcılar aşağıda verilmektedir.

Iso Density Surface, Surface Integral of (ESP-)(ESP-)*

Bir molekülün elektrostatik potansiyel eş-yoğunluk yüzeyinde negatif olan alanın karesini ifade eder.

*Iso Density Surface, Surface Integral of ESP*ESP*

Bir molekülün elektrostatik potansiyel eş-yoğunluk yüzeyinde pozitif olan alanın karesini ifade eder.

Iso Density Surface, Surface Integral of ESP-

Bir molekülün elektrostatik potansiyel eş-yoğunluk yüzeyinde negatif olan alanıdır.

Iso Density Surface, Area, ESP+:

Bir molekülün elektrostatik potansiyel eş-yoğunluk yüzeyinde pozitif olan alanın iki boyutlu yüzeye izdüşümünün alanıdır.

Iso Density Surface, Mean Absolute Deviation of ESP+:

Bir molekülün elektrostatik potansiyel eş-yoğunluk yüzeyinde pozitif olan alanın gradyantını ifade eder.

Iso Density Surface, Minimum ESP:

Bir molekülün elektrostatik potansiyel eş-yoğunluk yüzeyinde negatif yada pozitif olan alanlarda bulunan mutlak değeri maksimum olan noktadır.

Öncü-orbital dalga fonksiyonu tabanlı moleküler tanımlayıcılar

HOMO Energy: En yüksek dolu moleküler orbital enerjisini ifade etmekte olup eşitliği şu şekildedir;

$$\varepsilon_{HOMO} = \langle \phi_{HOMO} | \hat{F} | \phi_{HOMO} \rangle \quad (3.59)$$

ϕ_{HOMO} = En yüksek dolu moleküler orbitalin dalga fonksiyonu

$\hat{F}$ = Fock operatörü

HOMO-1Energy: En yüksek dolu moleküler orbitalin bir altındaki moleküler orbitalin enerjisi;

$$\varepsilon_{HOMO-1} = \langle \phi_{HOMO-1} | \hat{F} | \phi_{HOMO-1} \rangle \quad (3.60)$$

ϕ_{HOMO-1} = En yüksek dolu moleküler orbitalin bir altındaki moleküler orbitalin dalga fonksiyonu

$\hat{F}$ = Fock operatörü

$LUMO+1Energy$: En düşük dolu olmayan moleküler orbitalin bir üstündeki moleküler orbitalin enerjisi;

$$\varepsilon_{LUMO+1} = \langle \phi_{LUMO+1} | \hat{F} | \phi_{LUMO+1} \rangle \quad (3.61)$$

ϕ_{LUMO} = En düşük dolu olmayan moleküler orbital

$\hat{F}$ = Fock operatörü

Bu verilen üç orbital tabanlı moleküler tanımlayıcının her bir molekül için bir enerji özdeğeri olduğu gibi, her bir orbitalin dalga fonksiyonundan elektron eş-yoğunluk haritalarını oluşturmak mümkündür. Bu haritalar moleküllerin reaktiviteleri hakkında bize yorumlamaya uygun veriler sunar.

Atomik kısmi yük tabanlı moleküler tanımlayıcılar

Bu araştırmada kullandığımız Codessa3 yazılımı bir molekülde yer alan her bir atom için üç farklı yöntemle kısmi yük hesaplamaktadır. Bunlardan birincisi Gaussian 03/09 çıktı dosyalarından okunan kuantum mekaniksel hesaplama tabanlı Mulliken kısmi yükleridir. Mulliken popülasyon analizi Gaussian hesaplamalarında programın normal ayarları ile hesaplanan değerler olduğundan Codessa3 her bir moleküldeki atom için Gaussian çıktı dosyalarından Mulliken kısmi yükleri standart olarak alır. Diğer iki kısmi yük hesaplaması deneysel yöntemlere dayanır. Bunlar Zefirov ve Gasteiger-Marsili yük hesaplama yöntemleridir. Bu çalışma elde edilen QSAR modellerinde bulunan Gasteiger-Marsili yükü yer almamaktadır.

Mulliken kısmi yükü hesaplaması:

$$Q_A = Z_A - \left(\sum_{k \in A} P_{kk} + \frac{1}{2} \sum_{l \neq k} P_{kl} + \frac{1}{2} \sum_{l \neq k} P_{lk} \right) = Z_A - P_{kl} \quad (3.62)$$

Z_A = Atomik çekirdek yükü

P_{kl} = Atomik popülasyonun matris elemanları.

Zefirov kısmi yükü hesaplaması:

$$Q_i = f(x_i) \quad (3.63)$$

χ_i = Atomik elektronegativite (Sanderson elektronegativite ölçeği her bir atomun için)

$$\chi_i = (\chi_i^0 \prod_{k=1}^n \chi_k)^{1/(n+1)} \quad (3.64)$$

χ_i^0 = izole olmuş atomların elektronegativitesi

n= verilen bir atomun ilk koordinasyonlarındaki atomları

Mulliken yüklerine dayalı aşağıdaki moleküler tanımlayıcılar modellerde yer almaktadır.

Net Mulliken Charge of All S Atoms: Kükürt (S) atomları için net Mulliken yükü değerini tanımlar;

$$Q_S = Z_S - P_{kl} \quad (3.65)$$

Maximum Mulliken Charge: Denklem 3.62’de yer alan eşitlikten yararlanarak Mulliken yükünün maksimum değeri şu şekilde ifade edilir;

$$Q_{max} = \max(Q^+) = Z - P_{kl} \quad (3.66)$$

Minimum Mulliken Charge: Denklem 3.62’de yer alan eşitlikten yararlanarak Mulliken yükünün minimum değeri şu şekilde ifade edilir:

$$Q_{min} = \min(Q^-) = Z - P_{kl} \quad (3.67)$$

Average Mulliken Charge for a S Atom: Kükürt atomu için ortalama Mulliken yükünü ifade eder.

$$\overline{Q_S} = \frac{\sum_{i=1}^N Q_{Si}}{N_S} \quad (3.68)$$

N_S =Moleküldeki kükürt atomlarının sayısı

Q_{Si} = i. kükürt atomunun Mulliken kısmi yükü

Zefirov yüklerine dayalı aşağıdaki moleküler tanımlayıcılar modellerde yer almaktadır.

Dipol Moment (Zefirov kısmi yükü tabanlı): Zefirov kısmi yükleri ile hesaplanan dipol moment büyüklüğünü ifade eder.

Zefirov yükünün değeri;

$$\mu = \sum Q_i \cdot r_i \quad (3.69)$$

r_i =Yüklerin koordinat vektörleri

Minimum Zefirov Charge: Verilen bir molekül için Zefirov yükünün en negatif değerli kısmi yükü ifade eder.

$$Q_{min} = \min(Q^-) = f(\chi_{min}) \quad (3.70)$$

Zefirov Charge Polarity/ R2: Zefirov yüküne göre hesaplanmış polarite değerinin bu yükler arasındaki mesafeye oranını ifade eder.

$$P'' = \frac{Q_{max} - Q_{min}}{R_{mm}^2} \quad (3.71)$$

Q_{max} = Atomik Zefirov kısmi yükünün en pozitif değeri

Q_{min} =Atomik Zefirov kısmi yükünün en negatif değeri

R_{mm}^2 = Yük değerleri arasındaki mesafenin karesi

Yüklenmiş kısmi moleküler yüzey alanlarına dayalı moleküler tanımlayıcıları:

Bu bölümdeki moleküler tanımlayıcılar Stanton vd, (1990) ve Jurs vd, (1992) tarafından ilk defa tanımlanan atomik kısmi yükler (Mulliken ve/veya Zefirov) ile geometrik moleküler tanımlayıcılar olan çözücü erişebilir atomik/moleküler yüzey alanlarının kombinasyonlarından elde edilen moleküler tanımlayıcılardır. Yüzey alanından kasıt Van der Waals yüzey alanıdır.

Fractional Minimum Mulliken Negative Charge:

$$\frac{Q_{min}}{PNSA} = \frac{Z_{min} - P_{kl}}{PNSA} \quad (3.72)$$

Q_{min} =Moleküldeki en negatif Mulliken yükü

$PNSA$ = Kısmi negatif moleküler yüzey alanı

PSASA-3, Mulliken: Mulliken popülasyon analizine göre hesaplanmış atomik pozitif yük değerinin erişilebilir çözücü yüzey alanını ile çarpımı üzerinden toplamı ifade eder.

$$PSASA - 3 = \sum_A Q_A \cdot S_A S_A \quad A \in \{\delta_A > 0\} \quad (3.73)$$

$S_A S_A$ = Pozitif yüklü erişilebilir çözücü moleküler yüzey alanı

Q_A = Pozitif Mulliken atomik kısmi yükü

NSASA-3, Mulliken: Mulliken popülasyon analizine göre hesaplanmış atomik negatif yük değerinin erişilebilir çözücü yüzey alanını ile çarpımı üzerinden toplamı ifade eder.

$$NSASA - 3 = \sum_A Q_A \cdot S_A S_A \quad A \in \{\delta_A > 0\} \quad (3.74)$$

$S_A S_A$ = Pozitif yüklü erişilebilir çözücü moleküler yüzey alanı

Q_A = Negatif Mulliken atomik kısmi yükü.

Fractional Minimum Zefirov Negative Charge Times ASASA:

Bir molekülde parçalı negatif Zefirov yükünün değeri ile ASASA'nın çarpımını ifade eder:

$$\frac{Q_{min}}{PNSA} * ASASA = \frac{\min(Q^-)}{PNSA} * ASASA \quad (3.75)$$

Yapısal-geometrik tabanlı moleküler tanımlayıcılar:

Bu bölümdeki moleküler tanımlayıcılar molekülün geometri optimizasyonu yapılırken kullanılan kuantum mekaniksel yöntem ve seçilen temel setin hesaplama hassasiyetine bağlıdır. Bu araştırmada tüm kumarin türevleri için hesaplamalar DFT yöntemi B3LYP fonksiyoneli ile 6-311G (d,p) temel seti kullanılarak hesaplanmıştır.

Maximum Bond Length for a O Atom: Bir molekülde Oksijen (O) atomunun en yakın komşu atom için ölçülmüş olan maksimum bağ uzunluğunu ifade eder.

Minimum Bond Length for a C-O Bond: Molekülde yer alan (Karbon-Oksijen) C-O bağları için minimum bağ uzunluğunu ifade eder.

Average Bond Length for a C-C Bond: Molekülde yer alan (Karbon-Karbon) C-C bağlarının ortalama bağ uzunluğunu ifade eder. C-C bağ uzunluklarının toplamının C-C bağ sayısına bölümü ile bulunur.

Maximum Bond Length for a H-O Bond: Molekülde yer alan (Hidrojen-Oksijen) H-O bağları içinde maksimum bağ uzunluğunu ifade eder.

Maximum Bond Length for a H-O Bond: Molekülde yer alan (Hidrojen-Oksijen) H-O bağları içinde maksimum bağ uzunluğunu ifade eder.

Eylemsizlik Momenti (orta), Principal moments of inertia of a molecule:

$$I_k = \sum_i m_i r_{ik}^2 \quad (3.76)$$

m_i = Molekölü oluşturan atomların atomik ağırlıkları

r_{ik}^2 = i. atomik çekirdeğin k. ana dönme eksenlerinden uzaklığı (k=X, Y veya Z)

Dipol Moment (Kuantum mekanik tabanlı): Bir molekülün dipol momenti elektronların orbital dalga fonksiyonları ve atomların çekirdek yükünden hesaplanır. Gaussian09 hesaplamasında standart olarak hesaplanır.

$$\mu = -\sum_{i=1}^{OCC} \int \phi_i \hat{r} \phi_i dv + \sum_{a=1}^M Z_a \vec{R}_a \quad (3.77)$$

ϕ_i = Moleküler orbital

$\hat{r}$ = Elektron konum operatör

Z_a = a. atomik çekirdek yükü

$\vec{R}_a$ = a. atomik konum vektörü

Uniform-Mass, Moment of Inertia of Mass, Average: X, Y ve Z yönündeki eylemsizlik momentlerinin ortalamasını ifade eder;

$$\bar{I}_k = \frac{I_X + I_Y + I_Z}{3} \quad (3.78)$$

Valans ve yapısal moleküler tanımlayıcılar:

Number of N Atoms: Molekülde yer alan Azot (N) atomlarının sayısını ifade eder.

Number of O Bonds: Molekülde yer alan Oksijen (O) bağlarının sayısını ifade eder.

Number of N Bonds: Modelde yer alan Azot (N) atomlarının bağ uzunluklarını ifade eder.

Sum Valence of All O Atoms: Molekülde yer alan Oksijen (O) atomlarının valanslarının toplamını ifade eder.

Average Valence for a C Atom: Modelde yer alan bir Karbon (C) atomu için ortalama valansı ifade eder.

3.8 QSAR Modellerinin Uygulanabilirlik Sınırlarının Tanımlanması (AD)

Herhangi bir regresyon analizinde olduğu gibi bir QSAR modelinin belirlenmesi eğitim setindeki moleküllerin kimyasal yapısını temsil eden moleküler tanımlayıcıların değerlerine bağlıdır (Dimitrov vd, 2005). Dolayısıyla eğitim setindeki moleküllerin yapısal çeşitliliği o QSAR modelinin tahmin edebileceği kimyasal evrenin boyutlarını tayin eder. Eğitim seti ne kadar çok yapısal açıdan farklı moleküllerden oluşuyorsa o denli farklı moleküllerin aktivitesini tahmin etmede etkili olur. Modelin hangi moleküllerin aktivitelerini güvenle tahmin etmede uygun olduğunun belirlenmesi o modelin uygulanabilirlik sınırlarının tayin edilmesi anlamına gelir (Worth vd, 2004; Sheridan vd, 2004). Bu sınırların belirlenmesinde modeli oluşturmada kullanılan eğitim setinde yer alan moleküler tanımlayıcı değerleri kullanılır.

Uygulanabilirlik sınırı belirlemede birçok farklı yöntem geliştirilmiştir (Jaworska, 2005). Bunlardan literatürde sıkça kullanılanlardan bazıları;

- Geometrik yöntemler; Sınırlayıcı Kutu (Bounding box) yöntemleri, Dışbükey Gövde (Convex Hull) yöntemleri,
- Mesafe-tabanlı yöntemler; Euclidean mesafesi, Mahalanobis mesafesi ve City Block mesafesi,
- Parametrik olmayan yöntemler,
- Olasılık yoğunluğu dağılım yöntemleridir.

Biz bu çalışmada yukarıda belirtilen mevcut yöntemler içerisinde mesafe-tabanlı yöntemlerden Euclidean mesafesi kullanarak Leverage yaklaşımı ile QSAR modellerinin uygulanabilirlik sınırlarını belirledik.

Mesafe-tabanlı yaklaşımlar veri setindeki bir noktanın diğer her noktaya olan mesafesini hesaplar.

Euclidean Mesafesi

Euclidean mesafesi, tanımlayıcı uzayındaki iki nokta arasındaki mesafedir (Stanford, 2008) ve tanımlayıcı değerlerinin arasındaki uzaklığın karesi (Pisagor teoremi) ile hesaplanır:

$$d_2 = \|x - y\|^2 = \sum_v (x_v - y_v)^2 \quad (3.79)$$

v indisi, tanımlayıcı indeksidir ve uygulama alanını küresel boyutta açığa çıkarır.

Euclidean mesafesi izotropiktir yani; ortogonal eksenler farklı bir set ile ifade edilse de değişmez. Buna dayalı olarak uygulanabilirlik alanının tahmin geçerliliği eğitim setin merkez noktası etrafında izotropik olarak dağılıma bağlıdır.

Kaldıraç (Leverage) Yöntemi

Bir QSAR modeli için uygulanabilirlik alanı tayin edilme yöntemlerinden biri de kaldıraç gücü yöntemidir. Bu yöntem ile modeldeki bir bileşik kaldıraca göre değerlendirilir. Eğitim setine yer alan bileşiklerin kaldıracı h (etki matrisi veya yüzey matrisi), onun modeldeki etkisini ölçerek uygulanabilirlik alanındaki asıl değişkenini tanımlar (Eriksson vd, 2003).

$$h(\mathbf{x}) = N^{-1} + \mathbf{x}^T (X^T X)^{-1} \mathbf{x} \quad (3.80)$$

$\mathbf{x}$ vektörü; bileşiğin tanımlayıcı değerleri, X ; N satırlı eğitim set bileşiklerinin matris ifadesini ve X^T ; ise X matrisinin transpozunu temsil eder. Uyarıcı kaldıraç olarak tanımlanan h^* ise aşağıdaki gibi ifade edilir.

$$h^* = 3 \times h = 3 \times \sum_i \frac{h_i}{n} = 3 \times \frac{p'}{n} \quad i = (1, \dots, n) \quad (3.81)$$

n ; eğitim bileşiklerinin sayısı ve p' ; model parametrelerinin (moleküler tanımlayıcı) sayısını temsil eder.

Kaldıraç gücü değerleri hem eğitim bileşikleri hem de oluşturulan yeni bileşikler için hesaplanabilir. İlk durum açısından, güvenilir olmayan bir model oluşturulduğunda, model parametrelerini büyük ölçüde etkileyen eğitim bileşiklerinin hesaplanmasında yararlanılır. İkinci durum açısından ise oluşan modelin uygulanabilirlik alanının sınır tayini kontrolünde yararlanılır.

Bir bileşiğin kaldıraç gücü değerinin uyarıcı kaldıraçtan daha büyük olması tahmin edilen bu bileşiğin modelde extrapole olduğu anlamına gelir. Bu yüzden tahmin edilen bu değerin dikkatle kullanılması gerekir.

3.9 Kullanılan Donanım ve Yazılımlar

Gaussian 09

Gaussian Inc. (<http://www.gaussian.com/index.htm>) tarafından üretilen içinde birçok kuantum kimyasal hesaplama yönteminin bulunduğu yazılım. Bu çalışmada moleküllerin üç boyutlu yapılarının geometri optimizasyonun da ve kuantum kimyasal tabanlı moleküler tanımlayıcıların hesaplamasında kullanıldı. Gaussian 09 çıktı dosyaları Codessa3 yazılımına giriş dosyası yapılarak tüm moleküler tanımlayıcılar hesaplandı.

GausView 5

Gaussian Inc. tarafından (http://www.gaussian.com/g_tech/gv5ref/gv5ref_toc.htm) üretilen, Gaussian hesaplamalı için giriş dosyası hazırlamaya ve çıktı dosyalarındaki bilgilerin görselleştirilmesinde (spektrum, orbital, elektrostatik potansiyel vb.) kullanılan bir grafik ara yüz yazılımıdır. Bu çalışmada Gaussian giriş dosyalarının hazırlanmasında, moleküler orbitallerin görselleştirilmesinde kullanıldı.

Codessa 3

Semichem Inc. (<http://www.semichem.com/codessa/default.php>) tarafından üretilen moleküler tanımlayıcı hesaplama ve QSAR modeli oluşturmada içinde birçok istatistik algoritmanın yer aldığı yazılım. Bu çalışmada Gaussian çıktı dosyalarını kullanarak moleküler tanımlayıcı hesaplamada ve çoklu doğrusal regresyon algoritması ile QSAR modellerini oluşturmada kullanıldı.

Matlab

The Math Works Inc. firması tarafından (<http://www.mathworks.com>) üretilen çok yönlü bilimsel hesaplama yazılımıdır. Bu çalışmada Milano Chemometrics and QSAR Research Group (<http://michem.disat.unimib.it/chm/>) tarafından yazılan ve grup internet sitesinden indirilebilen bir Matlab yazılımı QSAR modellerinin uygulanabilirlik sınırlarını hesaplamada kullanıldı. (<http://michem.disat.unimib.it/chm/download/addownload.php>).

DiscoveryStudio Visualizer

Biovia (<http://accelrys.com/>) tarafından üretilen makro moleküllerin ve ligantların etkileşimini göstermede de kullanılabilen çok amaçlı ve çok fonksiyonlu görsel

arayüz yazılımıdır. Biz bu çalışmada kumarin bileşiklerinin CA II izoenzimi aktif sitesi içerisindeki yerleşimini göstermek için kullandık.



4. BULGULAR VE TARTIŞMA

4.1 Giriş

Bu bölümde moleküler yapıları bölüm 3 içerisinde yer alan Şekil 3.7’de, biyolojik aktivite verileri (K_i) Çizelge 3.1’de verilen moleküllerle elde edilen QSAR modelleri verilmektedir. Çizelge 3.1’de de görüldüğü gibi, çalışmada kullanılan 82 adet molekülün tamamının hem CA II’ye hem de CA IX’a karşı ölçülmüş deneysel aktivite verileri olmadığından ve az sayıda molekül için aktivite verileri olduğu halde DFT hesaplamaları başarı ile tamamlanamadığından QSAR modeline alınan molekül sayısı 60 adet altına inmiştir. CA II ve CA IX için QSAR modelleri oluşturmada aktivite verileri olan tüm molekül seti kullanılarak “global” modeller oluşturma yanında moleküller kükürt içeren kumarinler ve kükürt içermeyenler diye iki alt molekül setine ayrılarak her bir set için “lokal” QSAR modelleri oluşturuldu. Bu alt molekül setlerinin kendi içerisinde yapısal benzerlikleri yüksek olduğundan bu setlerden elde edilen QSAR modellerinin istatistik değerleri de global molekül setinden elde edilen QSAR modellerinin istatistik değerlerinden oldukça yüksek düzeyde olduğu gözlemlendi.

Bölüm 3’de belirtildiği üzere (Şekil 3.2 ve 3.3), kumarinlerin CA II aktif sitesi içerisinde lakton halkalarının hidrolize olarak açılıp karboksilik asit formuna dönüştüğü deneysel olarak tespit edilmiştir. Bu nedenle bu çalışmada tüm kumarinlerin karboksilik asit formu yapıları üzerinden moleküler tanımlayıcıları hesaplanarak QSAR modelleri de oluşturulmuştur. Aşağıda, önce CA II’nin kapalı formundan üretilen moleküler tanımlayıcılar ile elde edilen QSAR modelleri global, kükürt içeren ve kükürt içermeyen kumarin setleri için ayrı ayrı verilecek daha sonra aynı setlerin karboksilik asit formu için QSAR modelleri yine ayrı ayrı verilecek. Takip eden bölümde aynı yaklaşım CA IX için de benzer şekilde verilecek.

4.2 Karbonik Anhidraz II İzoenzimi QSAR Modelleri

4.2.1 CA II İzoenzimi tüm kumarinler (kapalı form) global QSAR modeli

CA II kapalı form QSAR modeli 42 kumarin molekülü ile oluşturulmuştur. Modelde yer alan moleküller Çizelge 4.1’de gösterilmektedir. Model 42 molekül üzerinden Codessa 3 yazılımı ile En İyi Çoklu Doğrusal Regresyon (BMLR) analizi algoritması kullanılarak elde edilmiştir. Sonra modelin geçerliliğinin tasdiki için üç alt kümeye ayrılmıştır. Ayırma işlemi tüm moleküllerin Ki değerleri sıraya dizildikten sonra 1, 4, 7,...40’a kadar olanlar A seti, 2, 5, 8,...41’e kadar olanlar B seti ve 3, 6, 9,...42’ye kadar olanlar C seti olarak belirlenmiştir. Bu ayırma yaklaşımı bu çalışmada bundan sonra verilen tüm QSAR modelleri için aynı şekilde uygulanmıştır. Bu ayrıştırma yönteminin amacı eğitim ve test setlerindeki moleküllerin yapısal ve aktivite açısından çeşitliliğini maksimize etmektir.

CA II kapalı form QSAR modelinde yer alan moleküllerin A, B ve C alt setlerindeki moleküller Çizelge 4.1’de görülmektedir. QSAR modelinin sağlamlığının tasdiki açısından rotasyona tabi tutularak üç alt setten ikisi eğitim seti geriye kalan birisi test seti olarak seçilerek modelin performansı gözlenmiştir. Çizelgedan da görüleceği üzere eğitim setleri ((A+B) , (A+C) veya (B+C)) ile oluşturulan modellerin istatistiki manada “uygunluğunun” oldukça yüksek olan R^2 , Q^2 ve F le oldukça düşük olan s^2 değerinden görülmektedir. Modelin uygunluğunun yüksekliği yanında tahmin gücünün de oldukça iyi olduğu yüksek R_{ext}^2 ve düşük SE değerinden görülmektedir.

Çizelge 4.1: CA II İzoenzimi kapalı form tüm kumarinler QSAR modeli

CA II İnhibitörleri (Toplam Set N=42)	1, 2, 3, 17, 19, 20, 21, 22, 23, 24, 25, 26, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 79, 80, 81, 82						
CA II Set A (N=14)	3, 19, 21, 24, 52, 57, 60, 61, 65, 67, 68, 75, 77, 80						
CA II Set B (N=14)	17, 20, 23, 25, 51, 53, 55, 58, 63, 69, 71, 72, 74, 81						
CA II Set C (N=14)	1, 2, 22, 26, 54, 56, 59, 64, 66, 70, 73, 76, 79, 82						
Literatürde CA II aktivitesi (Ki) olan molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.854	0.780	33.653	0.625	C	0.820	0.804
Set (B+C)	0.842	0.767	30.655	0.664	A	0.857	0.749
Set (A+C)	0.859	0.795	35.106	0.596	B	0.818	0.825
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	8.482	6.875	5.547	4.462	3.287	2.044	
Number of N Atoms	1.172	1.305	1.143	8.872	8.187	8.379	
Maximum Valence	-1.654	-1.583	-1.448	-9.076	-8.973	-7.343	
IsoDensity Surface, Surface Integral of (ESP-)*(ESP-)	15.651	25.118	28.569	1.766	2.432	2.689	
Fractional Minimum Mulliken Negative Charge	8.681	8.884	9.689	2.971	3.234	2.623	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

Çizelge 4.1 incelendiğinde Maximum Valence dışında QSAR modelinde yer alan tüm moleküler tanımlayıcıların (Bölüm 3.7'ye bakınız) pozitif katsayıya sahip oldukları görülmektedir. Bunun anlamı bir molekülde pozitif katsayılı bir moleküler tanımlayıcının mutlak değerindeki artış ile o molekülün biyolojik aktivite değeri ters orantılı demektir. Çizelge 4.1'deki QSAR modeline göre bir molekülün Maximum Valence değeri ile biyolojik aktivite değeri ise doğru orantılıdır. QSAR modelindeki

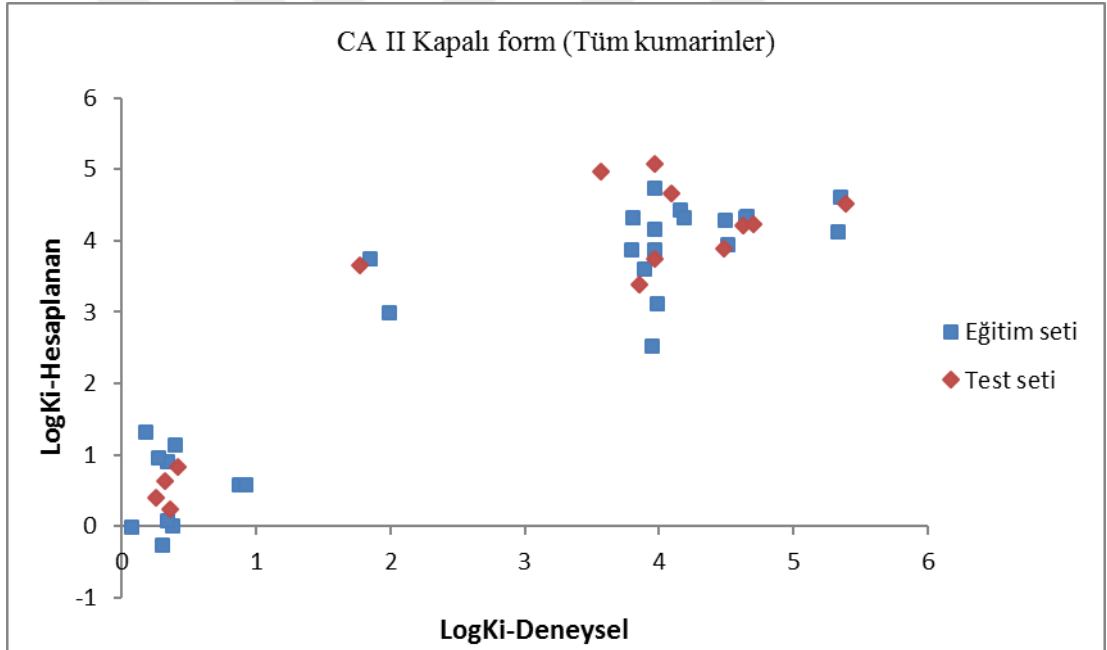
her bir moleküler tanımlayıcının t-testi değerinin mutlak değerce 1’den büyük olması bu tanımlayıcıların istatistiki manada “anlamlı” olduklarını göstermektedir.



Çizelge 4.2: CA II İzoenzimi kapalı form tüm kumarinler QSAR modeli tahmin çizelgesi

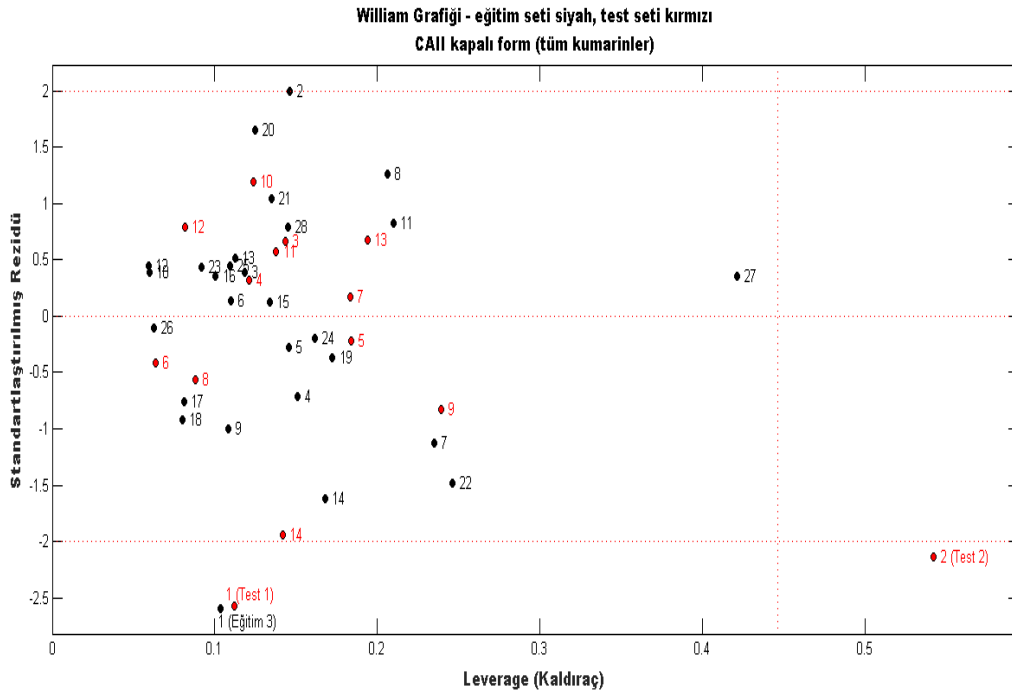
Molekül	LogKi-Deneysel	LogKi-Hesaplanan	Molekül	LogKi-Deneysel	LogKi-Hesaplanan
3	1,845	3,743	71	1,996	2,992
17	3,948	2,524	72	4,650	4,330
19	3,889	3,603	74	4,185	4,325
20	3,801	4,312	75	4,648	4,322
21	3,964	4,165	77	3,792	3,869
23	3,968	3,872	80	4,497	4,289
24	3,972	4,734	81	4,511	3,944
25	3,984	3,113	1	1,778	3,650
51	0,398	1,129	2	3,963	5,080
52	0,881	0,584	22	3,857	3,378
53	0,301	-0,265	26	3,971	3,737
55	0,924	0,588	54	0,255	0,407
57	0,380	0,008	56	0,322	0,636
58	0,176	1,318	59	0,362	0,245
60	0,079	-0,008	64	0,415	0,828
61	0,342	0,084	66	4,093	4,655
63	0,342	0,909	70	5,386	4,525
65	0,279	0,957	73	4,628	4,217
67	4,161	4,422	76	4,479	3,895
68	5,328	4,129	79	4,703	4,232
69	5,350	4,599	82	3,568	4,959

Çizelge 4.2’ de Çizelge 4.1’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.1’de Çizelge 4.2’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.2’de Çizelge 4.1’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin Kaldıraç yöntemine (Bölüm 3.8 bakınız) göre uygulanabilirlik sınırları gösterilmektedir. Bir QSAR modelinin uygulanabilirlik sınırları konusunda genel kabul, Y ekseninde verilen, molekül setindeki biyolojik aktivite (K_i) verilerine dayalı olarak ortaya çıkan “standartlaştırılmış rezidü” değerinin -2.5 ile +2.5 arasında ve X ekseninde verilen, molekül setindeki moleküler tanımlayıcıların değerlerine bağlı ortaya çıkan h^* kesmesinden küçük olan bölgedir. Bu kabule bağlı kalarak Şekil 4.2 incelendiğinde test setindeki 2 nolu molekülün yapısal tesirli açıdan sapan değer olduğu görülmektedir.



Şekil 4.1: CA II İzoenzimi kapalı form tüm kumarinler QSAR modeli deneysel K_i ’nin hesaplanan K_i değerine karşı grafiği (Bu grafik Çizelge 4.2’de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).

Bunun yanında, modelde yer alan moleküllerden eğitim setindeki 3, test setindeki 1 nolu moleküllerinde biyolojik aktiviteye bağlı sapan değerinin modelin uygulanabilirlik sınırının uç noktasında olduğu görülmektedir. Sapan değer özelliği gösteren moleküller için yapılan tahmin değerlerinin güvenilirliğine şüphe ile bakılmalıdır.



Şekil 4.2: CA II İzoenzimi kapalı form tüm kumarinler QSAR modeli sapan değer tayini.

4.2.2 CA II İzoenzimi Kükürt İçeren (Kapalı Form) Kumarinlerin QSAR Modeli

Çizelge 4.1’de yer alan tüm kumarinler den kükürt içeren kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içeren kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.3’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içeren kumarinler alt seti de kendi içinde A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.3’de kükürt içeren kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , R_{ext}^2 ve SE) bakıldığında modelin hem “uygunluğu” nun hem de “tahmin gücü” nün mükemmele yakın olduğu gözlenmektedir. Ayrıca modelde yer olan moleküler tanımlayıcıların mekanistik yorumlamaya uygun olması modelin değerini artıran bir unsurdur.

Çizelge 4.4’de Çizelge 4.3’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.3’de; çizelge 4.3’de

verilen deneysel K_i deęerlerinin hesaplanan K_i deęerlerine karşı grafięi gsterilmektedir. Őekil 4.4’de izelge 4.3’de eęitim seti (A+B), test seti C seildięi durumdan elde edilen modelin uygulanabilirlik sınırları gsterilmektedir.



Çizelge 4.3: CA II İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli

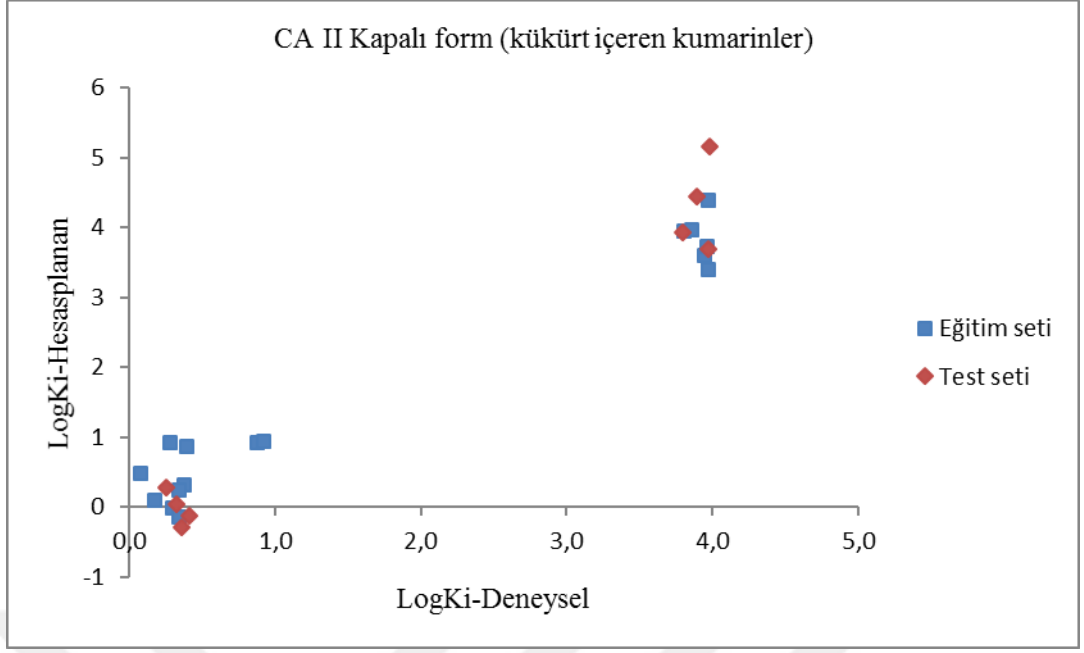
CA II inhibitörleri (Toplam Set N= 24)		17, 19, 20, 21, 22, 23, 14, 25, 26, 51, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 63, 64, 65, 77					
CA II Set A (N=8)		17, 20, 26, 52, 57, 60, 61, 65					
CA II Set B (N=8)		21, 22, 24, 51, 53, 55, 58, 63					
CA II Set C (N=8)		19, 23, 25, 54, 56, 59, 64, 77					
Literatürde CA II aktivitesi (K _i) olan kükürt içeren molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.960	0.930	66.564	0.169	C	0.963	0.567
Set (B+C)	0.973	0.928	99.302	0.121	A	0.932	0.500
Set (A+C)	0.956	0.917	60.626	0.196	B	0.954	0.412
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	1.649	0.417	0.504	1.126	0.468	0.465	
Number of O Bonds	-0.794	-0.641	-0.796	-8.169	-7.658	-7.107	
IsoDensity Surface, Surface Integral of ESP*ESP	39.544	38.109	39.156	11.835	11.368	9.603	
Net Mulliken Charge of All S Atoms	-1.612	-1.688	-1.437	-2.059	-5.368	-3.505	
Zefirov Charges' Dipole Moment Magnitude	-4.367	-2.950	-2.139	-4.197	-5.006	-2.575	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

Şekil 4.4 incelendiğinde reponse sapan değer olmamasına rağmen test setindeki 25, 54, 56 ve 77 nolu moleküllerin yapısal tesirlilik yönünden sapan değer olduğu gözlenmektedir. Çizelge 4.3 incelendiğinde “Iso Density Surface, Surface Integral of

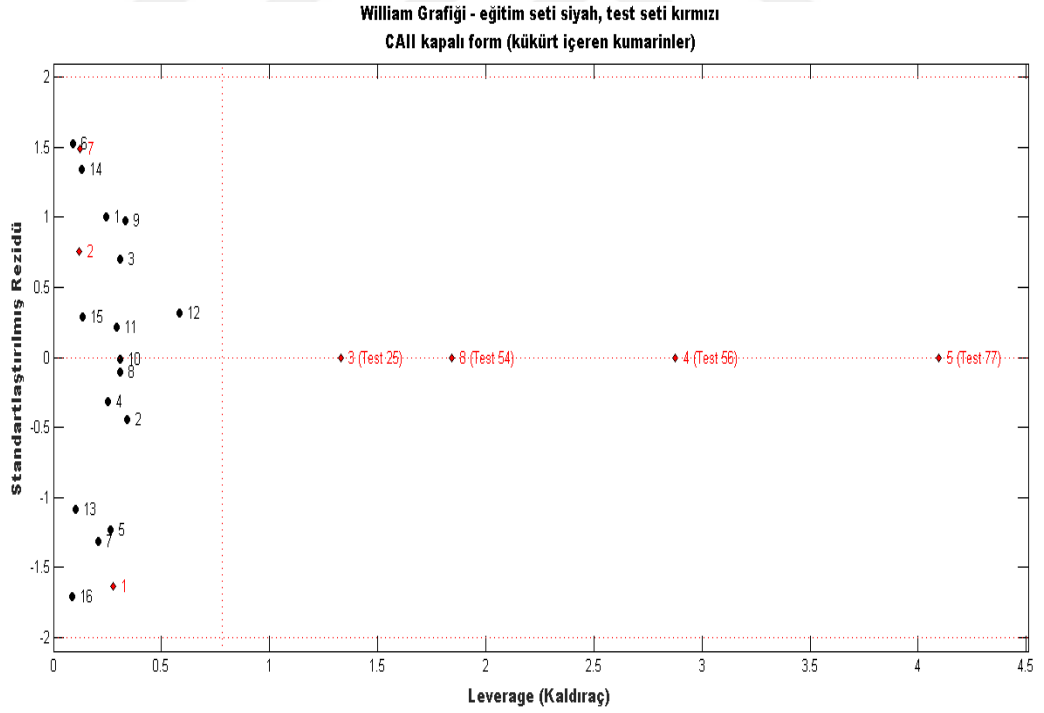
ESP*ESP” moleküler tanımlayıcısının (Bölüm 3.7’ye bakınız) değerindeki artışının aktiviteyi azaltırken modeldeki diğer moleküler tanımlayıcıların mutlak değerlerindeki artışı aktiviteyi artıran bir unsur olarak ortaya çıktığı görülmektedir. Bu modeldeki “Iso Density Surface, Surface Integral of ESP*ESP” moleküler tanımlayıcısı Çizelge 4.1 de verilen tüm kumarinler global QSAR modelinde yer alan “Surface, Surface Integral of (ESP-)*(ESP-)” gibi pozitif katsayı ya sahiptir. Bu moleküler tanımlayıcıların katsayılarının modellerdeki en büyük katsayı olması aktiviteyi belirlemede en etkin moleküler tanımlayıcı olduğunu göstermektedir.

Çizelge 4.4: CA II İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
17	3,948	3,605	60	0,079	0,483
20	3,801	3,942	61	0,342	-0,151
21	3,964	3,734	63	0,342	0,235
22	3,857	3,964	65	0,279	0,921
24	3,972	4,386	19	3,889	4,437
26	3,971	3,397	23	3,968	3,689
51	0,398	0,860	25	3,984	5,160
52	0,881	0,916	54	0,255	0,274
53	0,301	-0,013	56	0,322	0,040
55	0,924	0,929	59	0,362	-0,283
57	0,380	0,309	64	0,415	-0,134
58	0,176	0,097	77	3,792	3,934



Şekil 4.3: CA II İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.4'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.4: CA II İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli sapan değer tayini.

4.2.3 CA II İzoenzimi Kükürt İçermeyen (Kapalı Form) Kumarinlerin QSAR Modeli

Çizelge 4.1’de yer alan tüm kumarinler den kükürt içermeyen kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içermeyen kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.5’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içermeyen kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.5’de kükürt içermeyen kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , R_{ext}^2 ve SE) bakıldığında modelin hem “uygunluğu”nun hem de “tahmin gücü”nün oldukça iyi düzeyde olduğu gözlenmektedir. Ayrıca modelde yer olan moleküler tanımlayıcıların mekanistik yorumlamaya uygun olması modelin değerini artıran bir unsurdur.

Çizelge 4.6’de Çizelge 4.5’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.5’de Çizelge 4.6’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.6’da Çizelge 4.5’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

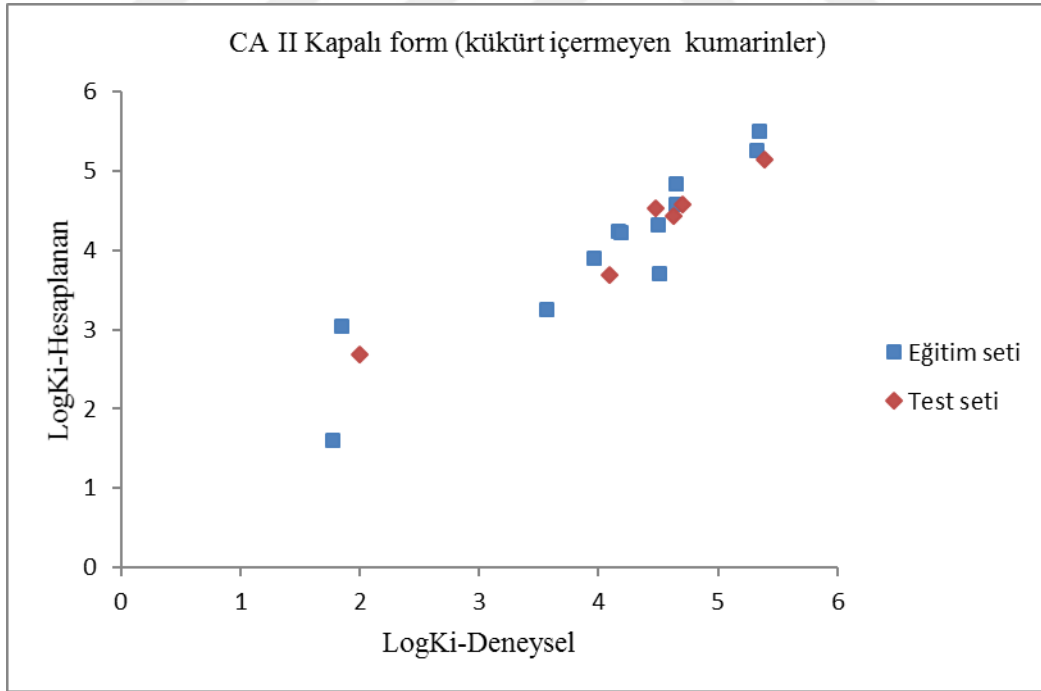
Çizelge 4.5: CA II İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli

CA II Inhibitörleri (Toplam Set N=18)		1, 2, 3, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 79, 80, 81, 82					
CA II Set A (N=6)		1, 67, 68, 75, 80, 82					
CA II Set B (N=6)		2, 3, 69, 72, 74, 81					
CA II Set C (N=6)		66, 70, 71, 73, 76, 79					
Literatürde CA II aktivitesi (K _i) olan molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.843	0.705	14.340	0.289	C	0.940	0.353
Set (B+C)	0.856	0.708	15.869	0.251	A	0.975	0.780
Set (A+C)	0.940	0.878	42.336	0.107	B	0.707	0.606
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı		Katsayı			t-testi		
		A+B	B+C	A+C	A+B	B+C	A+C
Intercept		437.730	648.467	442.985	6.010	5.514	10.326
Minimum Bond Length for a C-O Bond		-354.16	-525.50	-358.39	-5.954	-5.496	-10.244
Moment of Inertia of Mass, Middle		0.0015	0.0021	0.0016	5.755	6.237	9.556
IsoDensity Surface, Surface Integral of ESP-		3.810	5.975	3.951	5.380	4.717	8.795
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

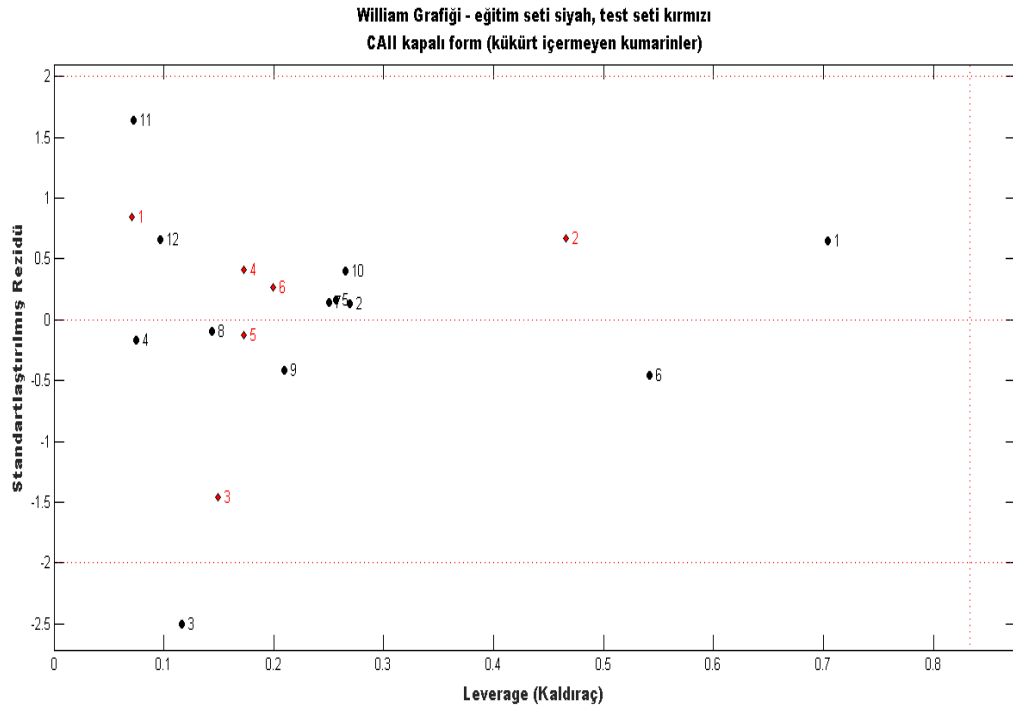
Şekil 4.6 incelendiğinde modelde hiçbir sapan değer olmadığı gözlenmektedir. Bundan önceki modellerde olduğu gibi bir adet ‘IsoDensity Surface’ moleküler tanımlayıcının (Bölüm 3.7’ye bakınız) pozitif katsayı ile modelde yer bulduğu görülmektedir.

Çizelge 4.6: CA II İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
1	1,778	1,598	80	4,497	4,323
2	3,963	3,905	81	4,511	3,710
3	1,845	3,039	82	3,568	3,249
67	4,161	4,245	66	4,093	3,682
68	5,328	5,255	70	5,386	5,139
69	5,350	5,507	71	1,996	2,679
72	4,650	4,586	73	4,628	4,438
74	4,185	4,229	76	4,479	4,537
75	4,648	4,836	79	4,703	4,583



Şekil 4.5: CA II İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.6'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.6: CA II İzoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli uygulanabilirlik sınırları ile sapan değer tayini.

4.2.4 CA II İzoenzimi Karboksilik Asit (Açık Form) Tüm Kumarinler “Global” QSAR Modeli

Bu başlıkta bölüm 3’de belirtildiği üzere kumarinlerin CA aktif sitesi içerisinde hidrolize olarak karboksilik asit formunda yerleştiğinin deneysel sonuçlardan anlaşılması dolayısıyla çalışmaya alınan tüm moleküllerin bölüm 3 Şekil 3.8’de gösterildiği biçimiyle moleküler tanımlayıcı hesaplandı. Hesaplama esnasında bazı moleküllerin optimizasyonu başarı ile sonuçlandıramadığından çalışma seti dışında tutuldu. Bu nedenle yukarıda QSAR modelleri verilen kapalı formunda bulunan moleküllerle birebir aynı olmadığı belirtilmelidir. Kumarinlerin karboksilik asit formu içinde kapalı formda olduğu gibi global, kükürt içerenler ve kükürt içermeyenler olmak üzere 3 adet model oluşturuldu.

Tüm kumarinler için QSAR modelinin istatistiği Çizelge 4.7’de gösterilmektedir. Çizelge 4.7 incelendiğinde “Maximum Bond Length for a O Atom” (Bölüm 3.7’ye bakınız) dışında QSAR modelinde yer alan tüm moleküler tanımlayıcıların pozitif katsayıya sahip oldukları görülmektedir. Bunun anlamı bir molekülde pozitif katsayılı bir moleküler tanımlayıcının mutlak değerindeki artış ile o molekülün biyolojik aktivite değeri ters orantılı demektir. Çizelge 4.7’deki QSAR modeline göre bir molekülün “Maximum Bond Length” mutlak değeri ile biyolojik aktivite değeri doğru orantılıdır. Diğer moleküler tanımlayıcılar için durum bunun tersidir.

QSAR modelindeki her bir moleküler tanımlayıcının t-testi değerinin mutlak değerce 1’den büyük olması bu tanımlayıcıların istatistiki manada “anlamlı” olduklarını göstermektedir.

Çizelge 4.7: CA II İzoenzimi açık (karboksilik asit) formu tüm kumarinler QSAR modeli

CA II İnhibitörleri (Toplam Set N=44)	1, 2, 3, 17, 19, 21, 22, 23, 24, 25, 26, 27, 28, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82						
CA II Set A (N=15)	1, 17, 23, 25, 27, 52, 57, 60, 61, 65, 69, 72, 74, 81, 82						
CA II Set B (N=15)	2, 3, 19, 26, 53, 55, 58, 63, 64, 66, 70, 73, 76, 77, 79						
CA II Set C (N=14)	21, 22, 24, 28, 54, 56, 59, 62, 67, 68, 71, 75, 78, 80						
Literatürde CA II aktivitesi (Ki) olan molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.868	0.796	31.582	0.567	C	0.757	0.911
Set (B+C)	0.806	0.715	19.228	0.822	A	0.888	0.699
Set (A+C)	0.855	0.762	27.166	0.608	B	0.790	0.888
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	36.468	34.800	32.090	9.292	7.404	7.441	
Number of N Atoms	1.343	1.291	1.325	8.592	6.559	8.824	
Maximum Bond Length for a O Atom	-13.360	-15.169	-12.050	-9.443	-7.540	-6.468	
Maximum Mulliken Charge	0.975	1.041	1.423	2.075	1.523	2.863	
Minimum Mulliken Charge	2.595	2.465	2.788	3.595	2.318	3.740	
HOMO Energy	53.284	36.007	44.941	3.861	1.961	3.121	

Çizelge 4.7’de global QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R², Q², F, s², R_{ext}² ve SE) bakıldığında modelin hem “uygunluğu” nun hem de “tahmin gücü” nün oldukça iyi düzeyde

olduğu gözlenmektedir. Ayrıca modelde yer olan moleküler tanımlayıcıların mekanistik yorumlamaya uygun olması modelin değerini artıran bir unsurdur.

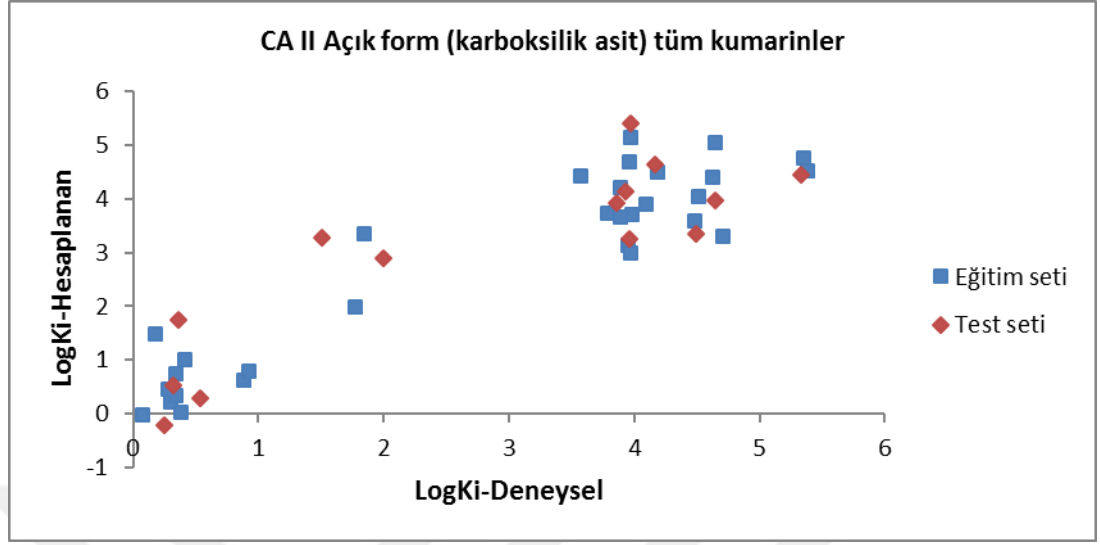
Çizelge 4.8’de Çizelge 4.7’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.5’de Çizelge 4.6’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.7’de Çizelge 4.8’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.



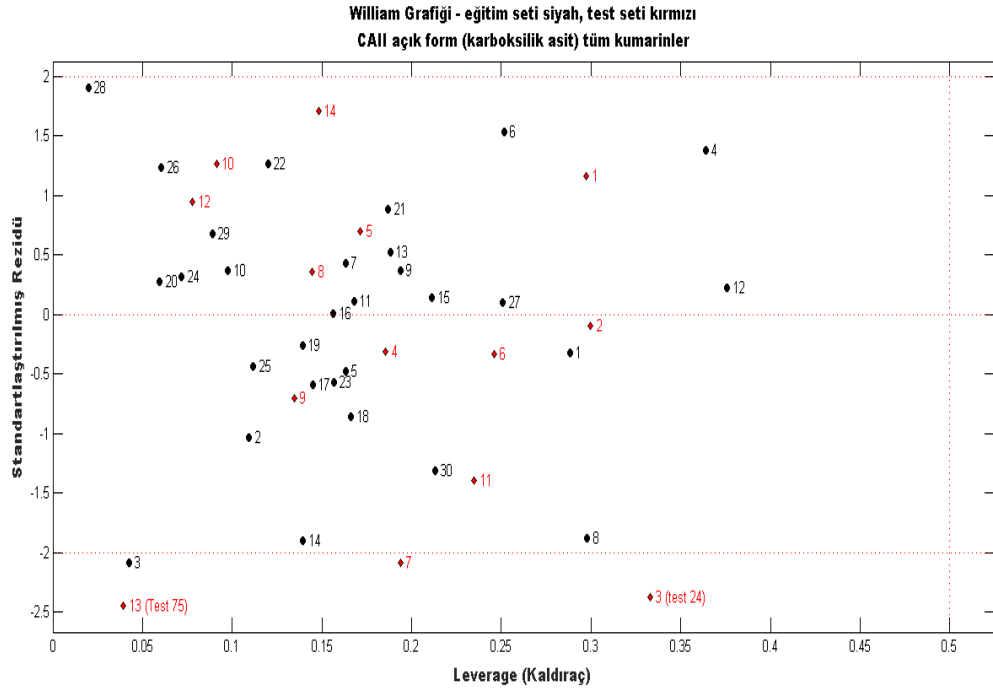
Çizelge 4.8: CA II İzoenzimi açık form (karboksilik asit) tüm kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesaplanan	Molekül	LogKi-Deneysel	LogKi-Hesaplanan
1	1,778	1,982	72	4,650	5,038
2	3,964	4,683	73	4,628	4,405
3	1,845	3,351	74	4,185	4,485
17	3,949	3,139	76	4,479	3,595
19	3,890	4,215	77	3,792	3,728
23	3,968	2,992	79	4,703	3,310
25	3,984	3,695	81	4,511	4,034
26	3,971	5,136	82	3,568	4,428
27	3,888	3,646	21	3,964	3,248
52	0,881	0,622	22	3,857	3,918
53	0,301	0,225	24	3,972	5,402
55	0,924	0,793	28	3,935	4,143
57	0,380	0,031	54	0,255	-0,214
58	0,176	1,480	56	0,322	0,537
60	0,079	-0,014	59	0,362	1,746
61	0,342	0,338	62	0,531	0,291
63	0,342	0,743	67	4,161	4,644
64	0,415	0,996	68	5,328	4,437
65	0,279	0,460	71	1,996	2,895
66	4,093	3,896	75	4,648	3,978
69	5,350	4,760	78	1,505	3,278
70	5,386	4,509	80	4,497	3,335

Şekil 4.8 incelendiğinde hem eğitim setindeki hem de test setindeki tüm kumarinlerin modelin uygulanabilirlik sınırları içinde olduğu gözlenmektedir.



Şekil 4.7: CA II İzoenzimi karboksilik asit (açık form) tüm kumarinler QSAR modeli. Deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği.



Şekil 4.8: CA II İzoenzimi karboksilik asit (açık form) tüm kumarinler QSAR modeli sapın değer tayini.

4.2.5 CA II İsoenzimi Karboksilik Asit (Açık) Formda Kükürt İçeren Kumarinler QSAR Modeli

Çizelge 4.7’de yer alan tüm kumarinler den kükürt içeren kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içeren kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.9’da gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içeren kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.9’de kükürt içeren kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , R_{ext}^2 ve SE) bakıldığında modelin hem “uygunluğu” nun hem de “tahmin gücü” nün mükemmel yakın olduğu gözlenmektedir. Ayrıca modelde yer olan moleküler tanımlayıcıların mekanistik yorumlamaya uygun olması modelin değerini artıran bir unsurdur.

Çizelge 4.9: CA II İzoenzimi karboksilik asit (açık form) kükürt içeren kumarinler QSAR modeli

CA II İnhibitörleri (Toplam Set N=24)	17, 19, 21, 22, 23, 24, 25, 26, 27, 28, 52, 53, 54, 55, 56, 57, 58, 59, 60, 61, 62, 63, 64, 65						
CA II Set A (N=8)	17, 26, 27, 52, 57, 60, 61, 65						
CA II Set B (N=8)	19, 21, 24, 53, 55, 58, 63, 64						
CA II Set C (N=8)	22, 23, 25, 28, 54, 56, 59, 62						
Literatürde CA II aktivitesi (K _i) olan kükürt içeren molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.967	0.905	82.949	0.138	C	0.939	0.455
Set (B+C)	0.963	0.874	73.034	0.163	A	0.951	0.457
Set (A+C)	0.979	0.918	133.612	0.091	B	0.933	0.555
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	74.728	47.786	32.606	4.590	2.930	2.123	
Number of N Bonds	0.760	0.679	0.770	15.178	10.382	14.018	
IsoDensity Surface, Area, ESP+	-0.011	-0.004	-0.010	-4.219	-0.999	-3.560	
Average Bond Length for a C-C Bond	-48.478	-31.578	-20.761	-4.459	-2.830	-2.022	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

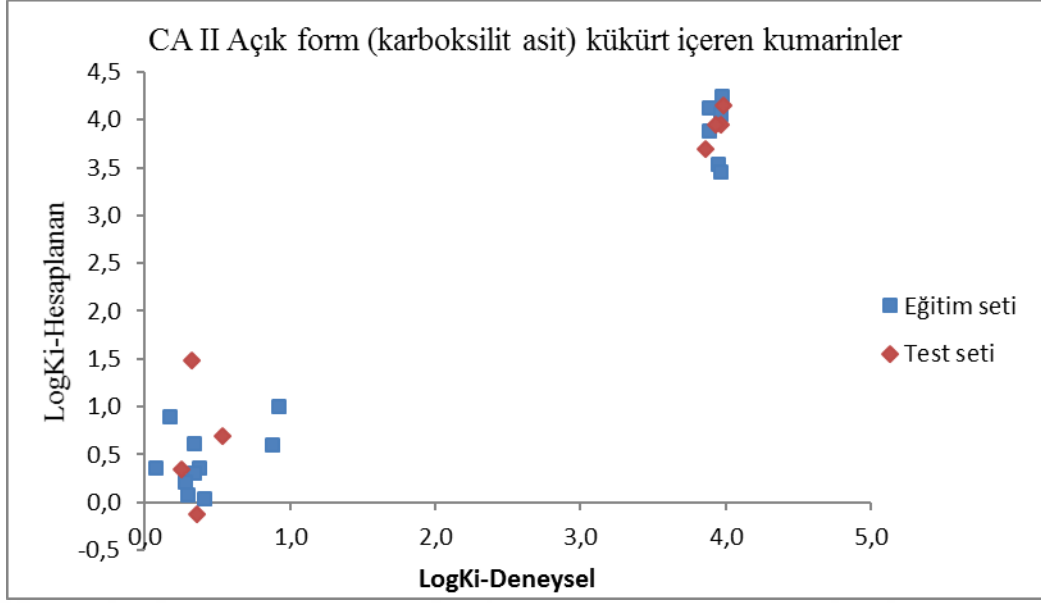
Çizelge 4.10'da ise Çizelge 4.9'da eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.9'da Çizelge 4.10'da verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği

gösterilmektedir. Şekil 4.10'da Çizelge 4.9'da eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

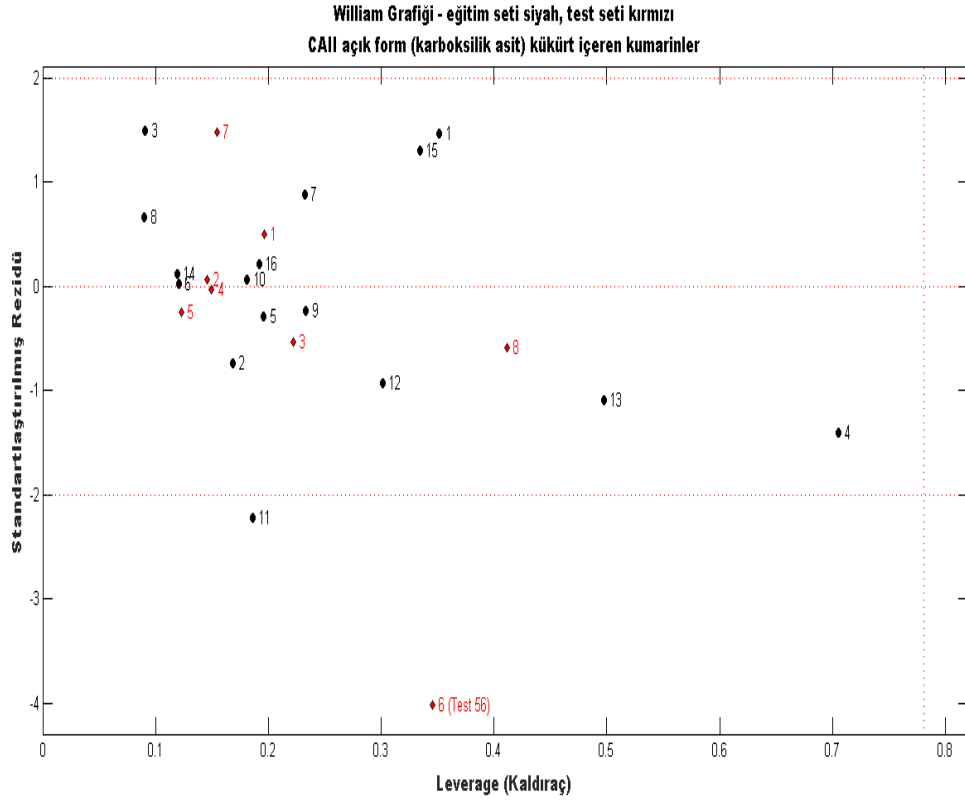
Çizelge 4.10: CA II İzoenzimi karboksilik asit (açık form) kükürt içeren kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
17	3,949	3,529	61	0,342	0,617
19	3,890	4,130	63	0,342	0,304
21	3,964	3,455	64	0,415	0,036
24	3,972	4,243	65	0,279	0,212
26	3,971	4,063	22	3,857	3,698
27	3,888	3,880	23	3,968	3,946
52	0,881	0,605	25	3,984	4,150
53	0,301	0,076	28	3,935	3,945
55	0,924	0,997	54	0,255	0,339
57	0,380	0,359	56	0,322	1,480
58	0,176	0,890	59	0,362	-0,122
60	0,079	0,355	62	0,531	0,693

Şekil 4.10 incelendiğinde eğitim setindeki tüm moleküllerin uygulanabilirlik sınırları içinde olduğu gözlenirken test setindeki 56 nolu kumarinin response sapan değeri olduğu gözlenmektedir. İstatistiki açıdan bu model 4.2.2 başlığında verilen kükürt içeren kumarinler (kapalı form) model ile kıyaslandığında benzer değerlere sahipken o modelin test setinde sekiz molekülden dördünün sapan değer olması buna karşın bu modelde sadece bir sapan değer olması bu modeli daha önemli kılmaktadır.



Şekil 4.9: CA II İzoenzimi karboksilik asit (açık form) kükürt içeren moleküller QSAR modeli. Deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.9'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.10: CA II İzoenzimi karboksilik asit (açık form) kükürt içeren kumarinler QSAR modeli sapan değer tayini.

4.2.6 CA II İzoenzimi Karboksilik Asit (Açık) Formda Kükürt İçermeyen Kumarinler QSAR Modeli

Çizelge 4.7’de yer alan tüm kumarinler den kükürt içermeyen kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içermeyen kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.11’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içermeyen kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.11’de kükürt içermeyen kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , $R(\text{ext.})^2$ ve SE) bakıldığında modelin hem “uygunluğu” nun hem de “tahmin gücü” nün oldukça iyi düzeyde olduğu gözlenmektedir.

Çizelge 4.11: CA II İzoenzimi karboksilik asit (açık form) kükürt içermeyen kumarinler QSAR modeli

CA II İnhibitörleri (Toplam Set N=20)	1, 2, 3, 66, 67, 68, 69, 70, 71, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82						
CA II Set A (N=7)	2, 69, 71, 72, 74, 78, 81						
CA II Set B (N=7)	1, 66, 70, 73, 76, 79, 82						
CA II Set C (N=6)	3, 67, 68, 75, 77, 802						
Literatürde CA II aktivitesi (K _i) olan kükürt içermeyen molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.857	0.678	20.125	0.296	C	0.646	0.799
Set (B+C)	0.713	0.309	7.488	0.484	A	0.912	0.561
Set (A+C)	0.817	0.601	13.467	0.399	B	0.809	0.656
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	11.526	9.697	10.767	9.766	7.264	8.163	
IsoDensity Surface, Mean Absolute Deviation of ESP+	-546.46	-342.77	-431.77	-4.790	-2.674	-4.379	
Zefirov Charges' Dipole Moment Magnitude	7.640	5.448	6.405	4.057	2.632	3.404	
Minimum Mulliken Charge	2.490	2.618	3.062	4.219	3.003	3.826	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

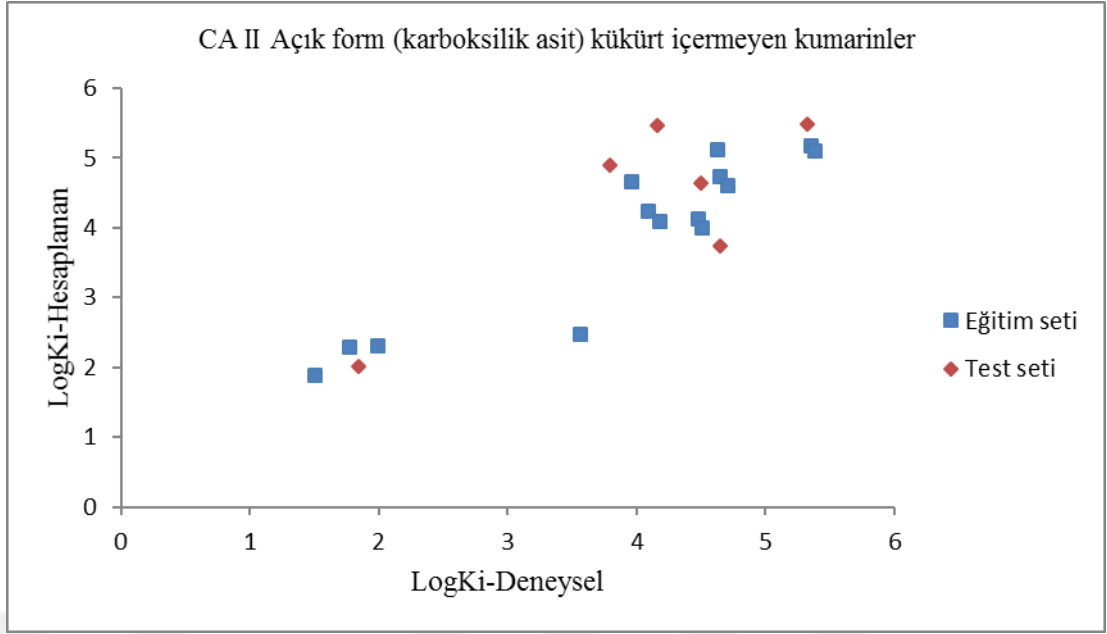
Çizelge 4.11 incelendiğinde, bu model, 4.2.3 başlığı altında verilen kükürt içermeyen kumarinler (kapalı form) modeli ile sağlamlık açısından karşılaştırıldığında daha az dengeli olduğu alt setlerin rotasyonunun istatistiğinden gözlenmektedir. (B+C) alt setleri eğitim seti seçilip C test seti seçildiği durumda R²=0.713, Q²=0.309 ve R_{ext}²=0.912 olması modelin sağlamlığını sarsmaktadır. Çizelge 4.12’de Çizelge 4.11’de

eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.10'da Çizelge 4.11'de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.11'de Çizelge 4.11'de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

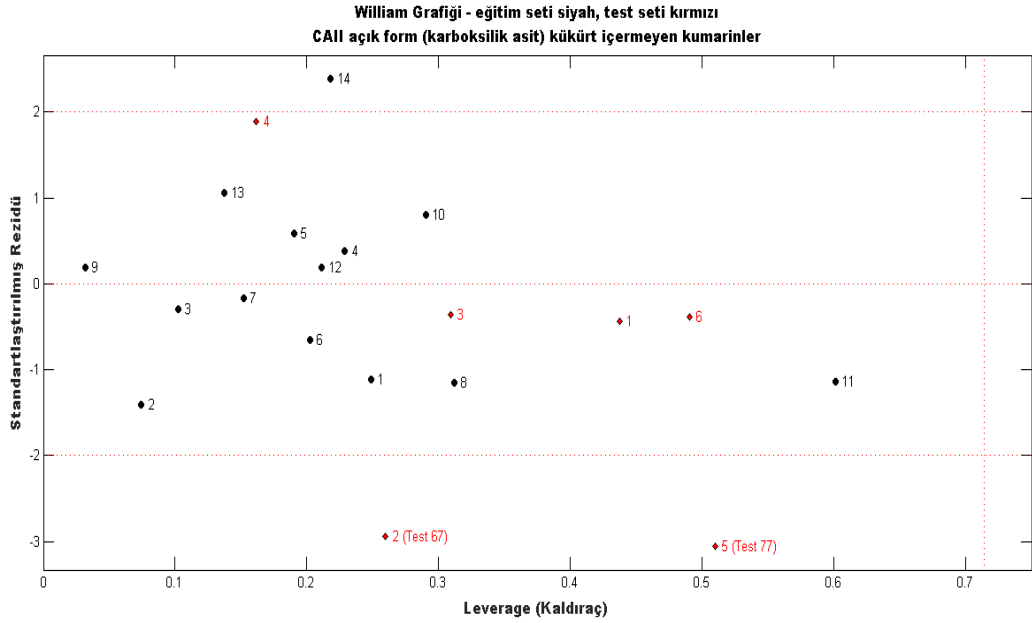
Çizelge 4.12: CA II İzoenzimi açık form (karboksilik asit) kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
1	1,778	2,281	78	1,505	1,877
2	3,964	4,668	79	4,703	4,613
66	4,093	4,236	81	4,511	4,001
69	5,350	5,173	82	3,568	2,471
70	5,386	5,109	3	1,845	2,015
71	1,996	2,300	67	4,161	5,473
72	4,650	4,733	68	5,328	5,482
73	4,628	5,121	75	4,648	3,748
74	4,185	4,085	77	3,792	4,903
76	4,479	4,127	80	4,497	4,638

Şekil 4.12 incelendiğinde eğitim setindeki tüm moleküllerin uygulanabilirlik sınırları içinde olduğu gözlenirken test setindeki 67 ve 77 nolu kumarinlerin response sapan değeri olduğu gözlenmektedir.



Şekil 4.11: CA II İzoenzimi karboksilik asit (açık form) kükürt içermeyen moleküller QSAR modeli. (Bu grafik Çizelge 4.10'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.12: CA II İzoenzimi karboksilik asit (açık form) kükürt içermeyen moleküler QSAR modeli sapan değer tayini.

4.3 Karbonik Anhidraz IX İzoenzimi QSAR Modelleri

CA II’de olduğu gibi CA IX içinde hem kapalı formda hem de karboksilik asit formunda QSAR modelleri oluşturuldu. CA II’de uygulandığı biçimde tüm kumarinlerin yer aldığı “global” QSAR modeli, kükürt içeren ve kükürt içermeyen kumarinler için lokal QSAR modelleri hesaplandı.

4.3.1 CA IX İzoenzimi Tüm Kumarinler (Kapalı Form) “Global” QSAR Modeli

CA IX kapalı form QSAR modeli 59 kumarin molekülü ile oluşturulmuştur. Modelde yer alan moleküller Çizelge 4.13’de gösterilmektedir. Model 59 molekül üzerinden Codessa 3 yazılımı ile En İyi Çoklu Doğrusal Regresyon Analizi algoritması kullanılarak elde edilmiştir. Sonra modelin geçerliliğinin tastiki için üç alt kümeye ayrılmıştır. Ayırma işlemi tüm moleküllerin K_i değerleri sıraya dizildikten sonra 1, 4, 7,...58 inci sıraya kadar olanlar A seti, 2, 5, 8,...59’e kadar olanlar B seti ve 3, 6, 9,...56’ye kadar olanlar C seti olarak belirlenmiştir. Bu ayırma yaklaşımı bu çalışmada bundan sonra verilen tüm QSAR modelleri için aynı şekilde uygulanmıştır. Bu ayırma yönteminin amacı eğitim ve test setlerindeki moleküllerin yapısal ve aktivite açısından çeşitliliğini maksimize etmektir.

CA IX kapalı form QSAR modelinde yer alan moleküllerin A, B ve C alt setlerindeki moleküller Çizelge 4.13’de görülmektedir. QSAR modelinin sağlamlığının tastiki açısından rotasyona tabi tutularak üç alt setten ikisi eğitim seti geriye kalan birisi test seti olarak seçilerek modelin performansı gözlenmiştir. Çizelgedan da görüleceği üzere eğitim setleri ((A+B) , (A+C) veya (B+C)) ile oluşturulan modellerin istatistiki manada “uygunluğunun” oldukça yüksek olan R^2 , Q^2 ve F ile oldukça düşük olan s^2 değerinden görülmektedir. Modelin uygunluğunun yüksekliği yanında tahmin gücünün de oldukça iyi olduğu yüksek R_{ext}^2 ve düşük SE değerinden görülmektedir. Ayrıca modelin sağlamlığı üst düzeyde olduğu eğitim ve test setleri rotasyona tabi tutulduğuna istatistiğin çok farklılaşmadığından gözlenmektedir.

Çizelge 4.13 incelendiğinde “Iso Density Surface, Surface Integral of ESP-“ ve “PSASA-3, Mulliken Charge” moleküler tanımlayıcıların (Bölüm 3.7’ye bakınız) QSAR modelinde pozitif katsayıya sahip iken diğer üç moleküler tanımlayıcının negatif katsayıya sahip oldukları görülmektedir. Bunun anlamı bir molekülde pozitif katsayılı bir moleküler tanımlayıcının mutlak değerindeki artış ile o molekülün biyolojik aktivite değeri ters orantılı demektir. QSAR modelindeki her bir moleküler

tanımlayıcının t-testi değerinin mutlak değerce 1'den büyük olması bu tanımlayıcıların istatistiki manada “anlamlı” olduklarını göstermektedir.

Çizelge 4.13: CA IX İzoenzimi kapalı form tüm kumarinler global QSAR modeli

CA IX İnhibitörleri (Toplam Set N=59)	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40,41, 42, 43, 44, 45, 46, 47, 48, 50, 66, 67, 68, 69, 70, 72, 73, 74, 75, 76, 77, 79, 80, 81, 82						
CA IX Set A (N=20)	1, 8, 9, 13, 17, 20, 21, 24, 25, 32, 33, 41, 45, 66, 67, 68, 75, 76, 79, 82						
CA IX Set B (N=20)	3, 7, 11, 14, 15, 16, 19, 30, 35, 38, 39, 40, 44, 46, 47, 48, 50, 69, 72, 80						
CA IX Set C (N=19)	4, 5, 6, 10, 12, 18, 22, 23, 26, 29, 31, 34, 42, 43, 70, 73, 74, 77, 81						
Literatürde CA IX aktivitesi (Ki) olan molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.658	0.546	13.132	0.493	C	0.590	0.657
Set (B+C)	0.650	0.542	12.300	0.452	A	0.657	0.704
Set (A+C)	0.670	0.536	13.405	0.384	B	0.711	0.793
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	64.061	73.86	49.444	4.090	4.082	3.353	
IsoDensity Surface, Surface Integral of ESP-	2.775	3.165	2.559	7.271	6.391	7.767	
Fractional Minimum Zefirov Negative Charge Times ASASA	-0.073	-0.137	-0.099	-1.615	-3.919	-3.098	
Net Zefirov Charge of All O Atoms	-7.013	-8.788	-6.092	-5.217	-5.280	-5.310	
Average Bond Length for a C-C Bond	-42.38	-49.178	-32.208	-3.789	-3.797	-3.022	
PSASA-3, Mulliken	0.037	0.0404	0.037	2.784	2.702	2.929	

Çizelge 4.14’ de Çizelge 4.13’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim

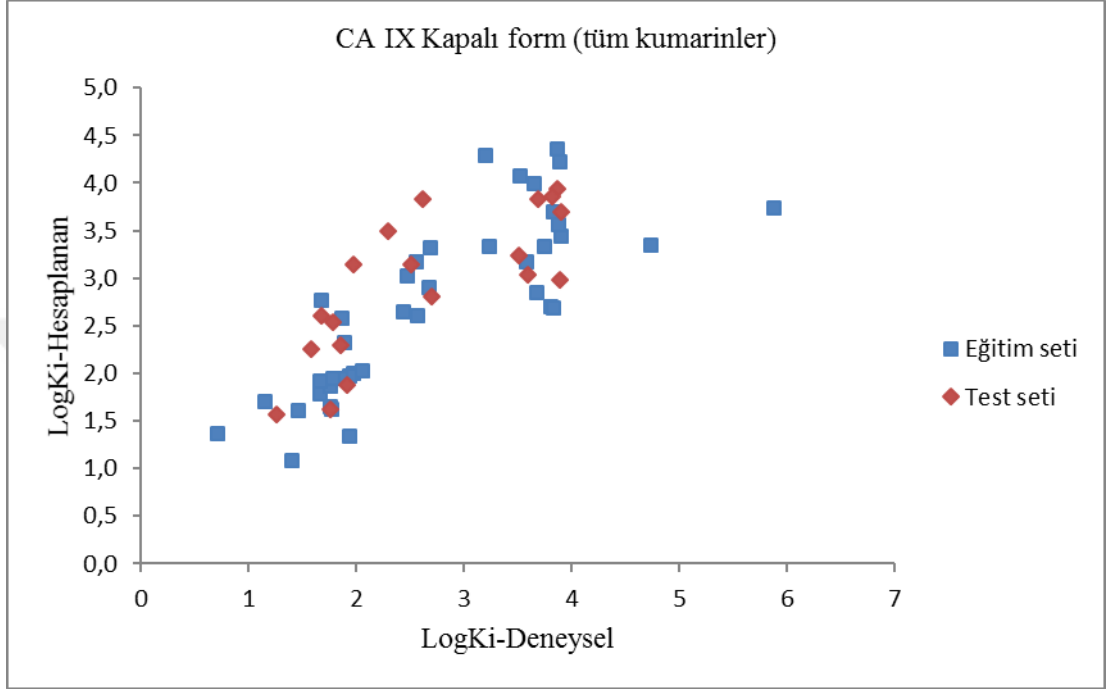
setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.13’de Çizelge 4.13’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.14’de Çizelge 4.13’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin Kaldıraç yöntemine (bölüm 3.8 bakınız) göre uygulanabilirlik sınırları gösterilmektedir.

Çizelge 4.14: CA IX İzoenzimi kapalı form tüm kumarinler QSAR modeli tahmin çizelgesi

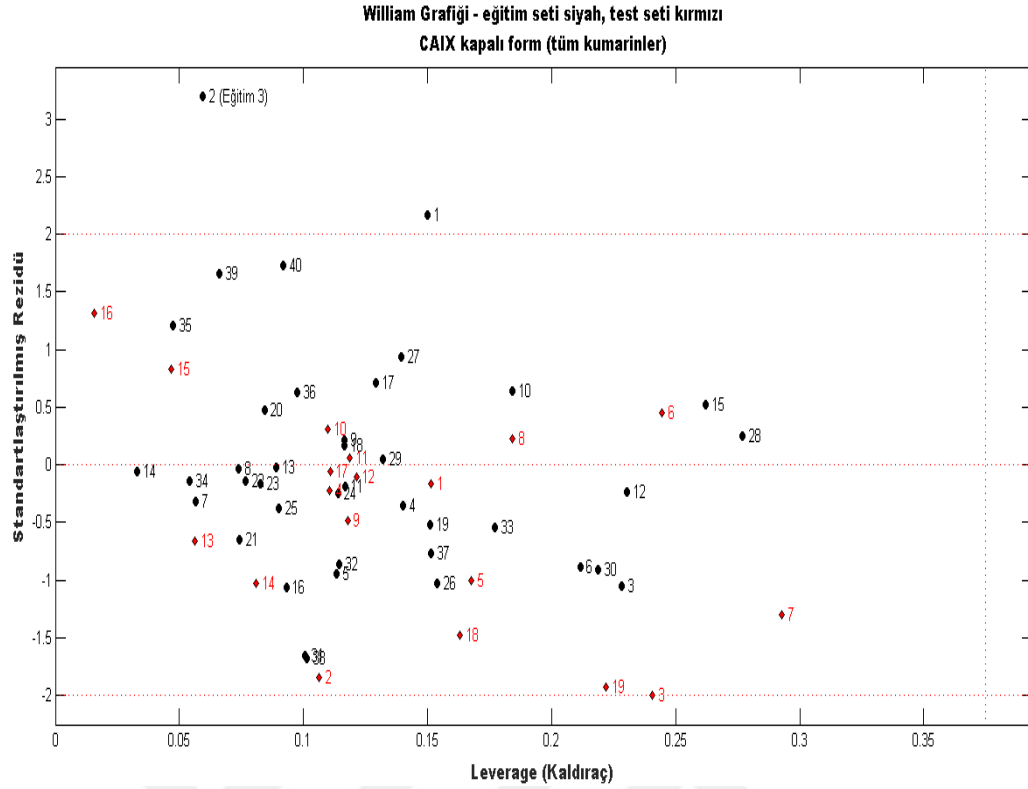
Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
1	4,736	3,349	66	3,204	4,289
3	5,884	3,739	67	3,519	4,079
7	2,683	3,320	68	3,653	3,996
8	2,679	2,908	69	3,230	3,328
9	2,555	3,168	72	3,672	2,852
11	2,477	3,018	75	3,748	3,337
13	2,439	2,651	76	3,869	4,360
14	2,574	2,600	79	1,672	2,771
15	3,834	3,695	80	3,813	2,705
16	3,576	3,176	82	3,826	2,686
17	1,662	1,783	4	2,705	2,806
19	1,462	1,607	5	2,621	3,829
20	1,977	1,994	6	2,296	3,499
21	1,934	1,970	10	3,691	3,835
24	1,398	1,086	12	2,510	3,145
25	1,869	2,572	18	3,513	3,238
30	3,904	3,443	22	1,778	2,536
32	1,765	1,653	23	1,763	1,622
33	3,892	4,224	26	1,255	1,569
35	3,880	3,564	29	3,905	3,701
38	1,894	2,325	31	1,919	1,879
39	1,850	1,940	34	3,869	3,937
40	1,754	1,860	42	1,859	2,301
41	1,787	1,946	43	1,577	2,261
44	1,669	1,915	70	3,591	3,031
45	0,716	1,370	73	3,886	2,980
46	1,940	1,340	74	3,820	3,856
47	1,771	1,621	77	1,672	2,611
48	2,053	2,022	81	1,968	3,146
50	1,146	1,703			

Bir QSAR modelinin uygulanabilirlik sınırları konusunda genel kabul, Y ekseninde verilen, molekül setindeki biyolojik aktivite verilerine dayalı olarak ortaya çıkan “standartlaştırılmış rezidü” değerinin -2.5 ile +2.5 arasında ve X ekseninde verilen, molekül setindeki moleküler tanımlayıcıların değerlerine bağlı ortaya çıkan h^* kesmesinden küçük olan bölgedir. Bu kabule bağlı kalarak Şekil 4.14 incelendiğinde eğitim setindeki 3 nolu molekülün response sapan değeri olduğu görülmektedir. Zaten Çizelge 4.14 incelendiğinde 3 nolu molekülün Log $K_i=4.884$ değeriyle diğer moleküllerde aktivitesinin oldukça farklı olduğu gözlenmektedir.

Bundan önceki CA II modellerinde olduğu gibi CA IX global modelinde de bir adet molekülün elektro statik potansiyel eş yüzey haritasına dayalı bir moleküler tanımlayıcı modele girmiştir . Buda CA II’de olduğu gibi CA IX’un inhibisyonunun QSAR modellemesinde kumarinlerin elektrostatik potansiyel eş yüzey haritalarının aktiviteyi modellemede çok etkin olduklarını göstermektedir.



Şekil 4.13: CA IX İzoenzimi kapalı form tüm kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.13’de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.14: CA IX İzoenzimi kapalı form tüm kumarinler QSAR modeli sapan değer tayini.

4.3.2 CA IX İzoenzimi Kükürt İçeren (Kapalı Form) Kumarinlerin QSAR Modeli

Çizelge 4.13’de yer alan tüm kumarinler den kükürt içeren kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içeren kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.15’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içeren kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.15’de kükürt içeren kumarinlerden elde edilen QSAR modelinin istatistikî değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , R_{ext}^2 ve SE) bakıldığında modelin hem “uygunluğu”nun hem de “tahmin gücü”nün mükemmel yakın olduğu gözlenmektedir. Ayrıca modelde yer olan moleküler tanımlayıcıların mekanistik yorumlamaya uygun olması modelin değerini artıran bir unsurdur.

Çizelge 4.16’da Çizelge 4.15’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.15’de Çizelge 4.15’de verilen deneysel Ki değerlerinin hesaplanan Ki değerlerine karşı grafiği gösterilmektedir. Şekil 4.16’da Çizelge 4.15’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

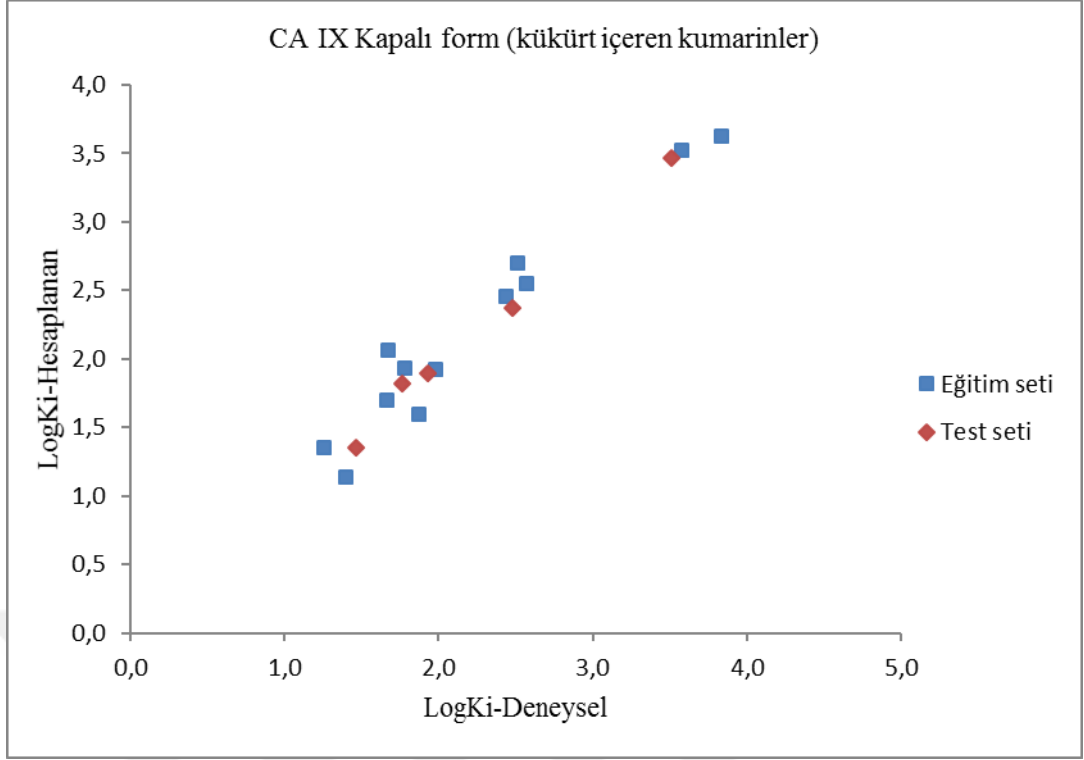
Çizelge 4.15: CA IX İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli

CA IX İnhibitörleri (Toplam Set N= 17)	11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 77						
CA IX Set A (N=6)	12, 16, 17, 20, 22, 26						
CA IX Set B (N=6)	13, 14, 15, 24, 25, 77						
CA X Set C (N=5)	11, 18, 19, 21, 23,						
Literatürde CA IX aktivitesi (Ki) olan kükürt içeren kumarinşer seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.943	0.872	44.586	0.051	C	0.993	0.075
Set (B+C)	0.954	0.875	49.091	0.041	A	0.942	0.220
Set (A+C)	0.984	0.956	152.08	0.012	B	0.912	0.248
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	-4.738	-2.858	0.227	-2.525	-1.114	-5.010	
IsoDensity Surface, Surface Integral of ESP*ESP	-20.41	-21.56	-20.27	-7.822	-9.103	-14.63	
HOMO-1 Energy	-37.40	-29.54	-39.91	-5.403	-2.960	-10.29	
Sum Valence of All O Atoms	0.154	0.227	0.133	2.738	2.861	4.423	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

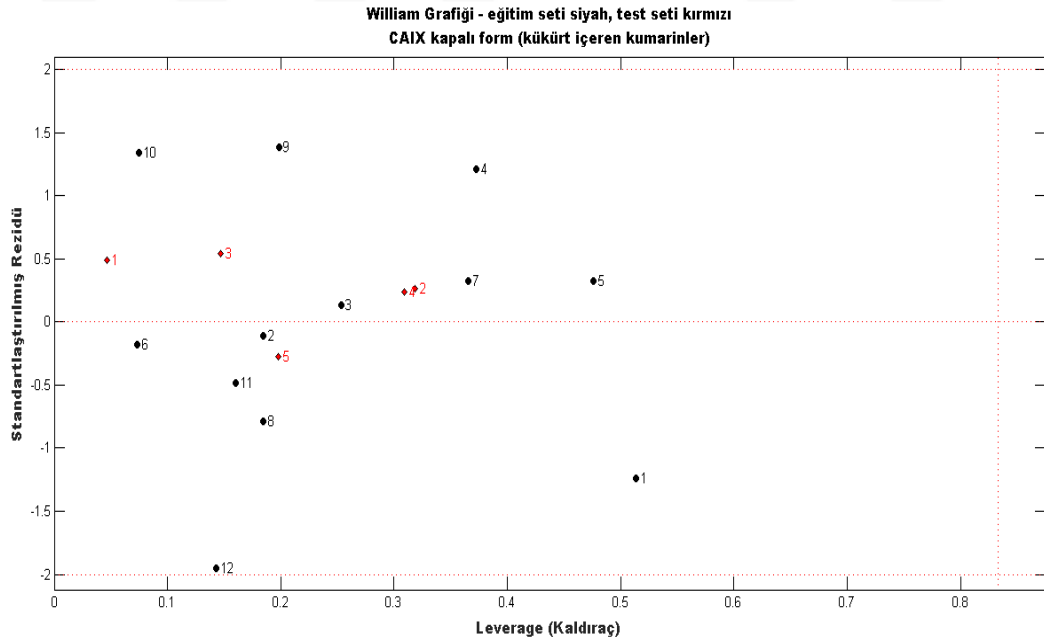
Çizelge 4.16: CA IX İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
12	2,510	2,695	25	1,869	1,593
13	2,439	2,460	26	1,255	1,350
14	2,574	2,550	77	1,672	2,059
15	3,834	3,629	11	2,477	2,374
16	3,576	3,526	18	3,513	3,467
17	1,662	1,699	19	1,462	1,355
20	1,977	1,922	21	1,934	1,892
22	1,778	1,929	23	1,763	1,816
24	1,398	1,133			

Burada verilen CA IX kükürt içeren (kapalı form) kumarinlerin QSAR modeli bu çalışmada şimdiye kadar verilen tüm modeller içerisinde istatistiki ve fiziksel yorumlanabilirlik açısından ortaya konulan en mükemmel model. Şekil 4.15 incelendiğinde deneysel aktiviteye karşı çizilen hesaplanmış aktivite değerlerinin mükemmele yakın uyum içerisinde olduğu gözlenmektedir. Ayrıca Şekil 4.16'dan görüleceği üzere modelde yer alan tüm eğitim ve test seti molekülleri modelin uygulanabilirlik sınırları içinde olup hiç sapan değer bulunmamaktadır.



Şekil 4.15: CA IX İzoenzimi açık form kükürt içeren kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.15'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.16: CA IX İzoenzimi kapalı form kükürt içeren kumarinler QSAR modeli sapma değeri.

4.3.3 CA IX İzoenzimi Kükürt İçermeyen (Kapalı Form) Kumarinlerin QSAR Modeli

Çizelge 4.13’de yer alan tüm kumarinler den kükürt içermeyen kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içermeyen kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.17’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içermeyen kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.17’de kükürt içermeyen kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F, s^2 , R_{ext}^2 ve SE) bakıldığında modelin “uygunluğu”nun ve “tahmin gücü”nün iyi düzeyde olduğu gözlenirken ‘sağlamlığının’ iyi düzeyde olmadığı rotasyon istatistiğindeki farklılıktan ortaya çıkmaktadır. Ancak modelde yer olan moleküler tanımlayıcıların tamamının mekanistik yorumlamaya uygun olması modelin önemini artıran bir unsurdur.

Çizelge 4.17: CA IX İzoenzimi kapalı form kükürt içermeyen kumarinler
QSAR modeli

CA IX İnhibitörleri (Toplam Set N=42)	1, 3, 4, 5, 6, 7, 8, 9, 10, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 50, 66, 67, 68, 69, 70, 72,73, 74, 75, 76, 79, 80, 81, 82						
CA IX Set A (N=14)	4, 5, 29, 31, 32, 34, 39, 44, 45, 48, 67, 72, 73, 80						
CA IX Set B (N=14)	1, 6, 8, 10, 33, 42, 46, 47, 50, 66, 70, 74, 76, 79						
CA IX Set C (N=14)	3, 7, 9, 30, 35, 38, 40, 41, 43, 68, 69, 75, 81, 82						
Literatürde CA IX aktivitesi (Ki) olan nonsülfo molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.750	0.641	17.320	0.330	C	0.511	0.441
Set (B+C)	0.576	0.335	7.840	0.644	A	0.889	0.654
Set (A+C)	0.729	0.626	15.536	0.406	B	0.651	1.117
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	4.371	5.417	9.755	2.640	2.151	2.373	
IsoDensity Surface, Surface Integral of (ESP-)*(ESP-)	-34.02	-31.55	-44.18	-6.962	-4.658	-7.080	
IsoDensity Surface, Minimum ESP	-117.4	-106.1	-161.7	-6.972	-3.933	-7.001	
LUMO+1 Energy	-40.50	-31.35	-42.92	-4.238	-2.269	-4.115	
HOMO Energy	32.047	32.230	62.38	4.496	3.169	3.188	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

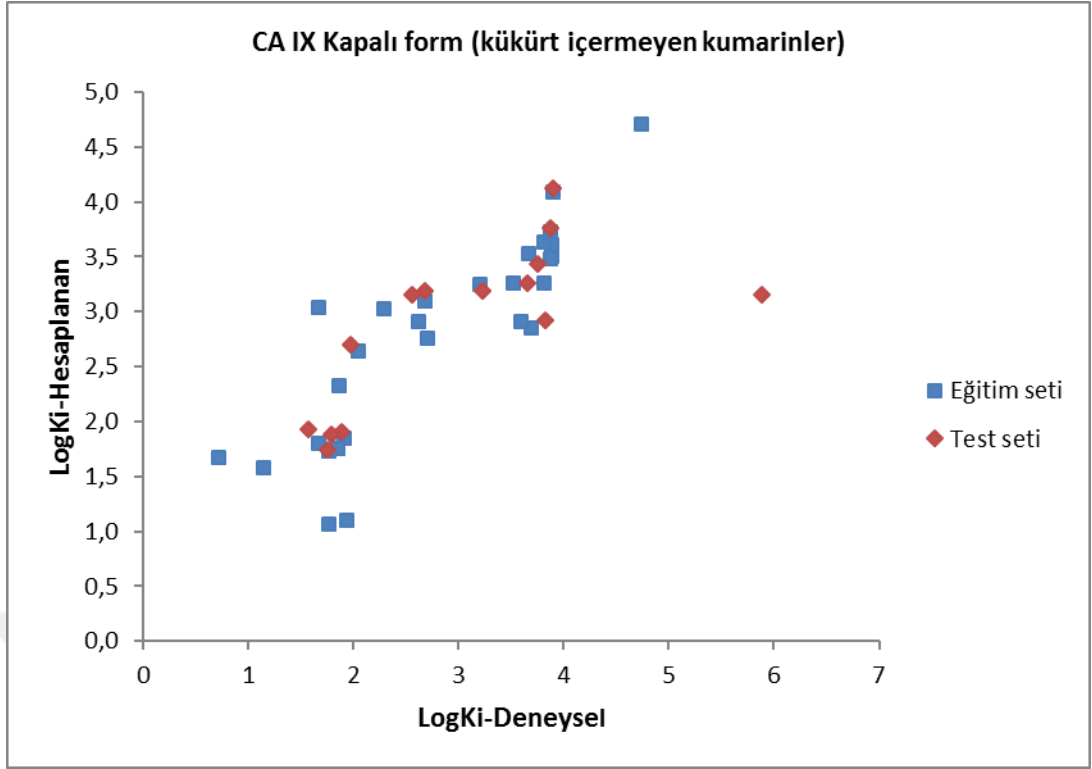
Çizelge 4.18’de ise Çizelge 4.17’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.17’de Çizelge 4.18’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği

gösterilmektedir. Şekil 4.18’da Çizelge 4.17’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

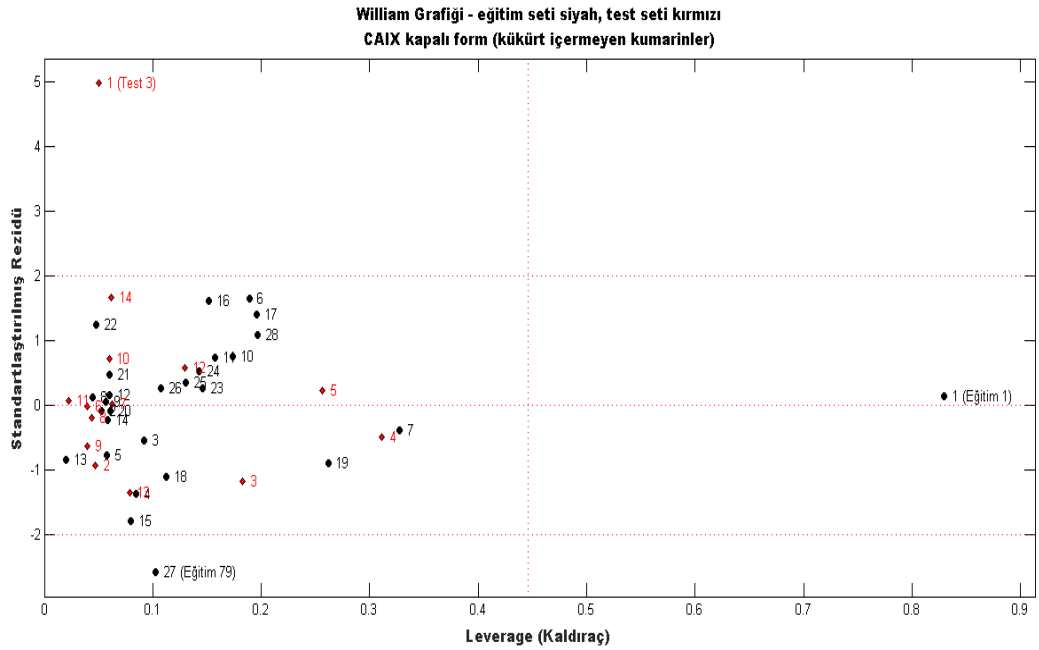
Uygunabilirlik sınırı incelendiğinde CA IX global modelinde olduğu gibi ekstrim aktivite (LogKi=5.884) sahip olan test seti molekülünün ve eğitim setindeki ekstrim aktivite değerine sahip olan (LogKi=4.736) 1 nolu molekülün sapan değer olduğu görülmektedir.

Çizelge 4.18: CA IX İzoenzimi kapalı form kükürt içermeyen kumarinler
QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
1	4,736	4,703	70	3,591	2,906
4	2,705	2,753	72	3,672	3,529
5	2,621	2,905	73	3,886	3,605
6	2,296	3,028	74	3,820	3,628
8	2,679	3,100	76	3,869	3,727
10	3,691	2,855	79	1,672	3,042
29	3,905	4,083	80	3,813	3,260
31	1,919	1,846	3	5,884	3,151
32	1,765	1,733	7	2,683	3,190
33	3,892	3,503	9	2,555	3,148
34	3,869	3,486	30	3,904	4,127
39	1,850	1,757	35	3,880	3,762
42	1,859	2,321	38	1,894	1,900
44	1,669	1,794	40	1,754	1,738
45	0,716	1,674	41	1,787	1,886
46	1,940	1,103	43	1,577	1,923
47	1,771	1,060	68	3,653	3,254
48	2,053	2,637	69	3,230	3,184
50	1,146	1,577	75	3,748	3,439
66	3,204	3,252	81	1,968	2,695
67	3,519	3,259	82	3,826	2,918



Şekil 4.17: CA IX izoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli.



Şekil 4.18: CA IX izoenzimi kapalı form kükürt içermeyen kumarinler QSAR modeli sapın değeri tayini.

4.3.4 CA IX İzoenzimi Tüm Kumarinlerin (Karboksilik Asit) “Global” QSAR Modeli

CA IX karboksilik asit formu QSAR modeli 56 kumarin molekülü ile oluşturulmuştur. Modelde yer alan moleküller Çizelge 4.19’da gösterilmektedir. Model 56 molekül üzerinden Codessa 3 yazılımı ile En İyi Çoklu Doğrusal Regresyon analizi algoritması kullanılarak elde edilmiştir. Sonra modelin geçerliliğinin tasdiki için üç alt kümeye ayrılmıştır. Ayırma işlemi tüm moleküllerin K_i değerleri sıraya dizildikten sonra 1, 4, 7,...55 inci sıraya kadar olanlar A seti, 2, 5, 8,...56’ya kadar olanlar B seti ve 3, 6, 9,...54’ye kadar olanlar C seti olarak belirlenmiştir. Bu ayırma yaklaşımı bu çalışmada bundan sonra verilen tüm QSAR modelleri için aynı şekilde uygulanmıştır. Bu ayrıştırma yönteminin amacı eğitim ve test setlerindeki moleküllerin yapısal ve aktivite açısından çeşitliliğini maksimize etmektir.

CA IX açık karboksilik asit formu QSAR modelinde yer alan moleküllerin A, B ve C alt setlerindeki moleküller Çizelge 4.19’da görülmektedir. QSAR modelinin sağlamlığının tasdiki açısından rotasyona tabi tutularak üç alt setten ikisi eğitim seti geriye kalan birisi test seti olarak seçilerek modelin performansı gözlenmiştir. Çizelgedan da görüleceği üzere eğitim setleri ((A+B) , (A+C) veya (B+C)) ile oluşturulan modellerin istatistiki manada “uygunluğunun” kabul edilebilir düzeyde olduğu R^2 , Q^2 , F ve s^2 değerinden görülmektedir. Modelin uygunluğunun kabul edilebilirliği yanında tahmin gücünün de kabul edilebilir düzeyde olduğu R_{ext}^2 ve SE değerinden görülmektedir. Ayrıca modelin sağlamlığının da ancak kabul edilebilir düzeyde olduğu eğitim ve test setleri rotasyona tabi tutulduğuna istatistiğin çok farklılaşmadığından gözlenmektedir.

**Çizelge 4.19: CA IX İzoenzimi karboksilik asit tüm kumarinler
QSAR Modeli**

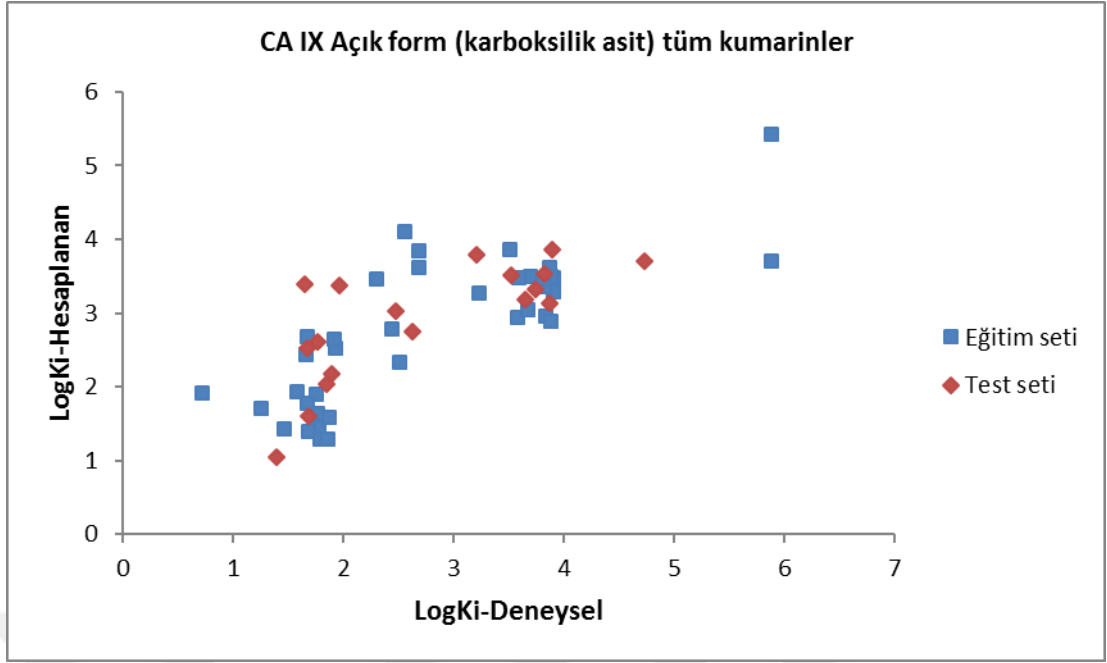
CA IX İnhibitörleri (Toplam Set N=56)	1, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40, 41, 42, 43, 44, 45, 66, 67, 68, 69, 70, 72, 73, 74, 75, 76, 77, 78, 79, 80, 81, 82							
CA IX Set A (N=19)	3, 6, 8, 12, 15, 16, 17, 19, 22, 30, 31, 35, 40, 42, 45, 69, 72, 79, 80							
CA IX Set B (N=19)	4, 7, 9, 10, 13, 18, 21, 23, 25, 26, 27, 29, 34, 41, 43, 44, 70, 73, 74							
CA IX Set C (N=18)	1, 5, 11, 24, 28, 32, 33, 38, 39, 66, 67, 68, 75, 76, 77, 78, 81, 82							
Literatürde CA IX aktivitesi (Ki) olan sülfö molekül seti ve alt setleri								
Eğitim Seti	R ²		Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.633		0.243	11.071	0.626	C	0.524	0.726
Set (B+C)	0.572		0.431	8.312	0.622	A	0.663	0.723
Set (A+C)	0.664		0.575	12.295	0.509	B	0.494	0.843
Eğitim ve test setlerinin rotasyonunun istatistiği								
Moleküler Tanımlayıcı		Katsayı			t-testi			
		A+B	B+C	A+C	A+B	B+C	A+C	
Intercept		65.97	83.93	87.36	3.801	2.643	5.726	
Uniform-Mass, Moment of Inertia of Mass, Average		-0.003	-0.004	-0.003	-4.641	-3.605	-4.018	
Dipole Moment		-0.248	-0.262	-0.230	-3.699	-3.568	-3.581	
Fractional Minimum Zefirov Negative Charge Times ASASA		-0.168	-0.202	-0.204	-3.019	-3.808	-4.478	
Maximum Bond Length for a H-O Bond		-32.05	-28.95	-41.76	-2.184	-2.217	-3.873	
Average Valence for a C Atom		-6.931	-11.97	-9.830	-1.948	-1.697	-3.439	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri								

Çizelge 4.20’de ise Çizelge 4.19’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.19’de Çizelge 4.20’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.20’de Çizelge 4.19’da eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

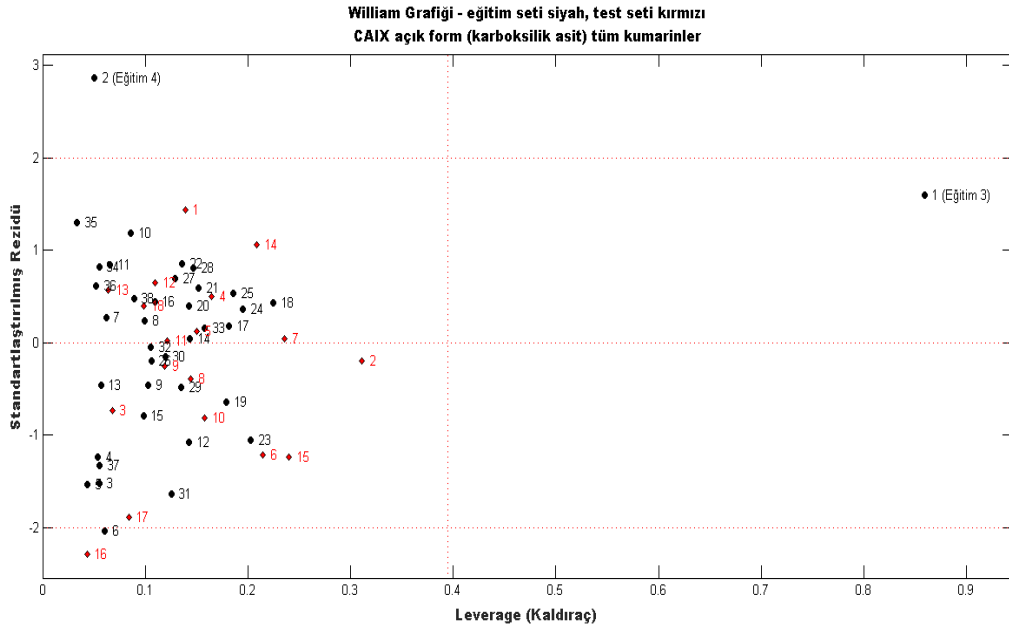
Uygunabilirlik sınırı incelendiğinde CA IX tüm kumarinler “global” QSAR modeli için eğitim setinde yer alan ekstrim aktivite ($\text{Log}K_i=5.885$) değerine sahip olan 4 nolu molekülün ve yine ekstrim aktivite değerine sahip olan ($\text{Log}K_i=5.885$) 3 nolu molekülün sapan değer olduğu görülmektedir.

Çizelge 4.20: CA IX İzoenzimi karboksilik asit tüm kumarinler QSAR modeli tahmin çizelgesi

Molekül	Log K_i -Deneysel	Log K_i -Hesap	Molekül	Log K_i -Deneysel	Log K_i -Hesap
3	5,885	5,418	43	1,577	1,926
4	5,885	3,710	44	1,669	1,785
6	2,297	3,453	45	0,716	1,910
7	2,683	3,620	69	3,230	3,272
8	2,679	3,845	70	3,591	3,479
9	2,555	4,096	72	3,672	3,052
10	3,692	3,489	73	3,886	2,891
12	2,511	2,334	74	3,820	3,353
13	2,439	2,778	79	1,672	2,680
15	3,834	2,953	80	3,813	3,462
16	3,576	2,939	1	4,736	3,698
17	1,663	2,442	5	2,621	2,753
18	3,513	3,860	11	2,477	3,035
19	1,462	1,435	24	1,398	1,039
21	1,934	2,520	28	1,690	1,600
22	1,778	1,452	32	1,765	2,608
23	1,763	1,638	33	3,892	3,866
25	1,869	1,577	38	1,894	2,175
26	1,255	1,707	39	1,850	2,035
27	1,681	1,395	66	3,204	3,785
29	3,905	3,483	67	3,519	3,506
30	3,904	3,287	68	3,653	3,179
31	1,919	2,653	75	3,748	3,323
34	3,869	3,616	76	3,869	3,135
35	3,880	3,502	77	1,672	2,517
40	1,754	1,901	78	1,653	3,398
41	1,787	1,284	81	1,968	3,382
42	1,859	1,282	82	3,826	3,535



Şekil 4.19: CA IX izoenzimi karboksilik asit tüm kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.19'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



Şekil 4.20: CA IX izoenzimi karboksilik asit tüm kumarinler QSAR modeli sapan değer tayini.

4.3.5 CA IX İzoenzimi Kükürt İçeren (Karboksilik Asit) Kumarinlerin QSAR Modeli

Çizelge 4.19’da yer alan tüm kumarinlerden kükürt içeren kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içeren kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.21’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içeren kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.21’de kükürt içeren kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , R_{ext}^2 ve SE) bakıldığında modelin hem “uygunluğu”nun hem de “tahmin gücü”nün iyi bir düzeyde olduğu gözlenmektedir. Ayrıca modelde yer olan moleküler tanımlayıcıların mekanistik yorumlamaya uygun olması modelin değerini artıran bir unsurdur.

Çizelge 4.22’de ise Çizelge 4.21’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.21’de Çizelge 4.22’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği gösterilmektedir. Şekil 4.22’de Çizelge 4.21’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

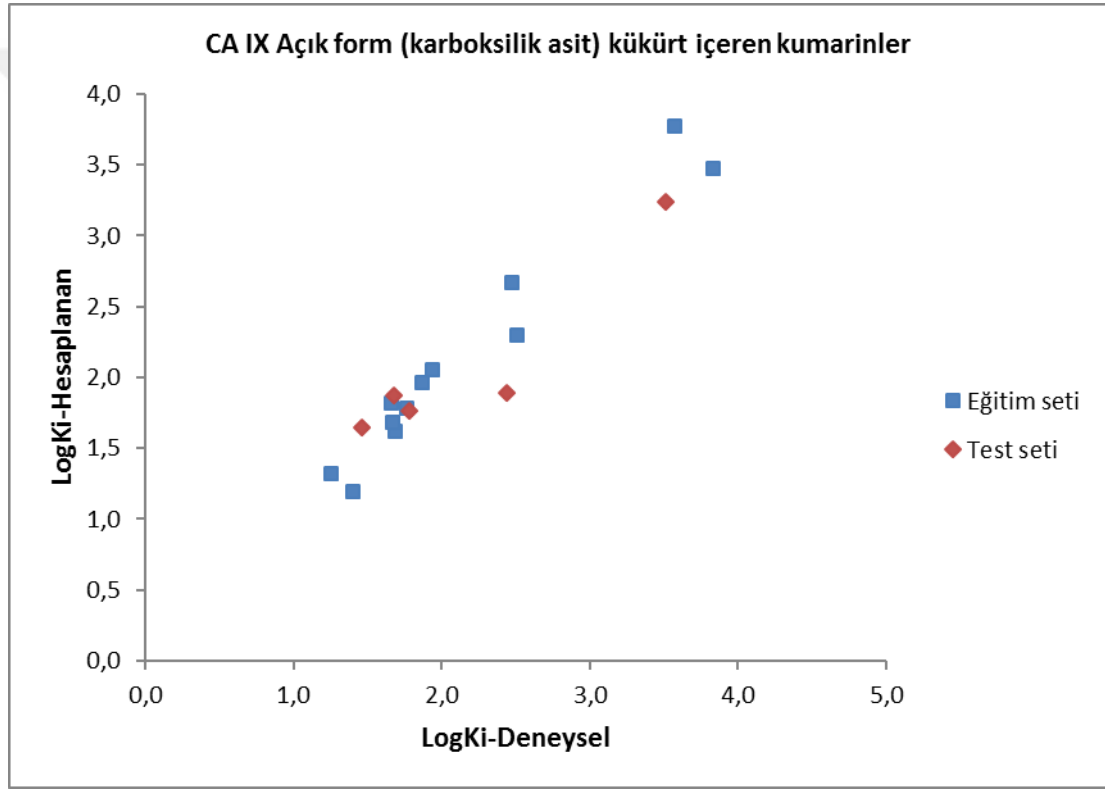
Çizelge 4.21: CA IX İzoenzimi karboksilik asit kükürt içeren kumarinler
QSAR Modeli

CA IX İnhibitörleri (Toplam Set N= 17)	11, 12, 13, 15, 16, 17, 18, 19, 21, 22, 23, 24, 25, 26, 27, 28, 77						
CA IX Set A (N=6)	11, 16, 17, 25, 26, 28						
CA IX Set B (N=6)	12, 15, 21, 23, 24, 77						
CA X Set C (N=5)	13, 18, 19, 22, 27						
Literatürde CA IX aktivitesi (K _i) olan sülfö molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.952	0.878	53.451	0.044	C	0.876	0.298
Set (B+C)	0.959	0.853	55.078	0.039	A	0.942	0.435
Set (A+C)	0.916	0.699	25.474	0.074	B	0.964	0.271
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	-13.54	-13.67	-16.46	-7.014	-5.591	-2.966	
Iso Density Surface, Surface Integralof ESP-	1.026	1.298	1.127	9.158	11.88	5.422	
HOMO Energy	-31.89	-24.62	-31.97	-8.466	-3.808	-5.241	
Minimum Zefirov Charge	-63.45	-80.03	-81.74	-6.871	-8.341	-2.826	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

Çizelge 4.22: CA IX İzoenzimi karboksilik asit kükürt içeren kumarinler

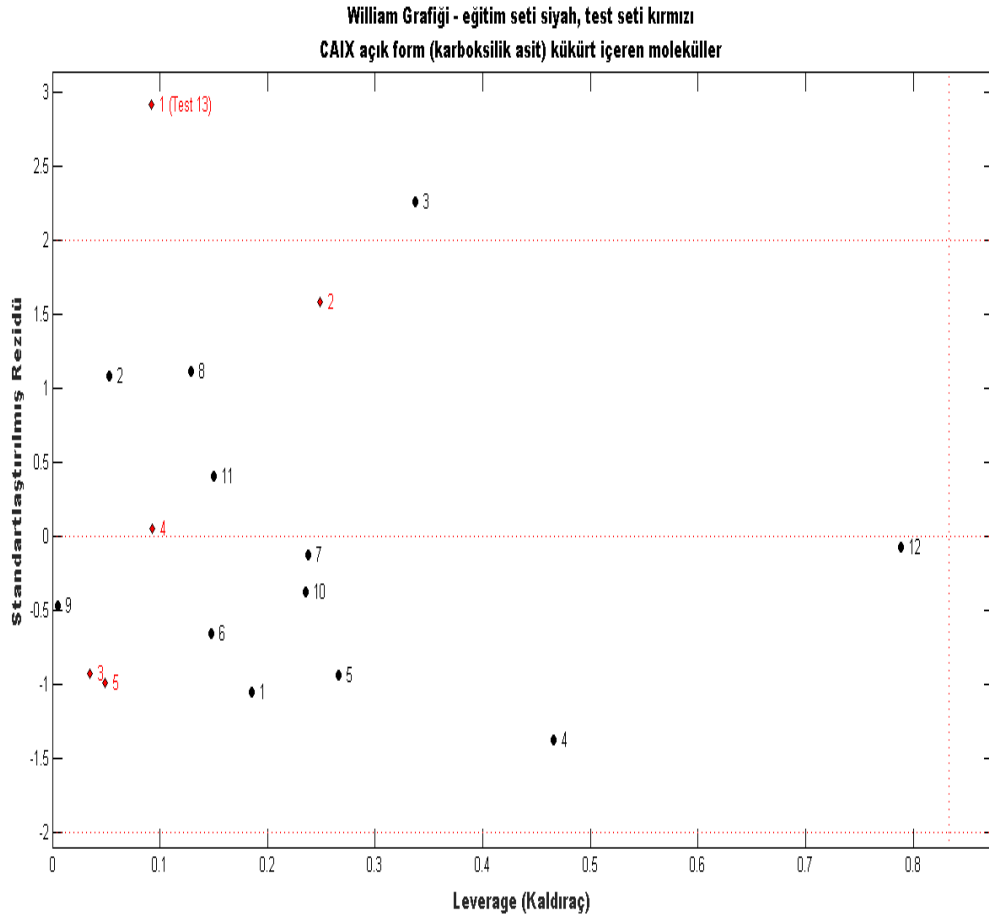
QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
11	2,477	2,665	26	1,255	1,320
12	2,511	2,301	28	1,690	1,616
15	3,834	3,470	77	1,672	1,679
16	3,576	3,776	13	2,439	1,889
17	1,663	1,823	18	3,513	3,241
21	1,934	2,055	19	1,462	1,643
23	1,763	1,785	22	1,778	1,768
24	1,398	1,191	27	1,681	1,873
25	1,869	1,962			



Şekil 4.21: CA IX izoenzimi karboksilik asit kükürt içeren kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.21'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).

CA IX karboksilik asit (açık) formu için verilen QSAR modelinin istatistiki açıdan anlamlı olduğu ($R^2=0.952$, $Q^2=0.878$, $F=53.451$ ve $R_{ext}^2=0.876$) ve fiziki olarak yorumlanabilir olduğu görülmektedir. Ayrıca Şekil 4.22'de de herhangi bir sapan değer olmadığı ve Çizelge 4.22'de aktivite değerlerinin iyi olması modelin uygulanabilir olabileceğini göstermektedir.



Şekil 4.22: CA IX izoenzimi karboksilik asit kükürt içeren kumarinler QSAR modeli sapan değer tayini.

4.3.6 CA IX İzoenzimi Kükürt İçermeyen (Karboksilik Asit) Kumarinlerin QSAR Modeli

Çizelge 4.19’da yer alan tüm kumarinler den kükürt içermeyen kumarinler (Şekil 3.7’ye bakınız) ayrıştırılarak kükürt içermeyen kumarinler alt seti oluşturulmuştur. Bu alt sette yer olan moleküller Çizelge 4.23’de gösterilmektedir. Tüm kumarinler setinde olduğu gibi kükürt içermeyen kumarinler alt seti de kendi içine A, B ve C alt setlerine ayrılarak bu alt setlerden ikisi eğitim seti biri test seti seçilerek elde edilen model güvenilirliğine bakılmıştır.

Çizelge 4.23’de kükürt içermeyen kumarinlerden elde edilen QSAR modelinin istatistiki değerleri gösterilmektedir. Elde edilen QSAR modelinin istatistik verilerine (R^2 , Q^2 , F , s^2 , R_{ext}^2 ve SE) bakıldığında modelin “uygunluğu”nun ve “tahmin gücü”nün orta düzeyde olduğu gözlenirken “sağlamlığının” iyi düzeyde olmadığı rotasyon istatistiğindeki farklılıktan ortaya çıkmaktadır. Ancak modelde yer olan

moleküler tanımlayıcıların tamamının mekanistik yorumlamaya uygun olması modelin önemini artıran bir unsurdur.

Çizelge 4.23:CA IX İzoenzimi Karboksilik Asit Kükürt İçermeyen Kumarinler QSAR Modeli

CA IX İnhibitörleri (Toplam Set N=39)	1, 3, 4, 5, 7, 8, 9, 10, 29, 30, 31, 32, 33, 34, 35, 38, 39, 40, 41, 42, 43, 44, 45, 66, 67, 68, 69, 70, 72, 73, 74, 75, 76, 78, 79, 80, 81, 82						
CA IX Set A (N=13)	1, 5, 10, 32, 33, 42, 44, 45, 66, 70, 74, 76, 81						
CA IX Set B (N=13)	3, 6, 8, 30, 35, 38, 41, 43, 68, 69, 75, 79, 82						
CA IX Set C (N=13)	4, 7, 9, 29, 31, 34, 39, 40, 67, 72, 73, 78, 80						
Literatürde CA IX aktivitesi (Ki) olan nonsülfo molekül seti ve alt setleri							
Eğitim Seti	R ²	Q ²	F	s ²	Test Seti	R _{ext} ²	SE
Set (A+B)	0.746	0.658	15.445	0.432	C	0.442	0.887
Set (B+C)	0.590	0.186	7.555	0.714	A	0.805	0.597
Set (A+C)	0.611	0.411	8.262	0.653	B	0.719	0.704
Eğitim ve test setlerinin rotasyonunun istatistiği							
Moleküler Tanımlayıcı	Katsayı			t-testi			
	A+B	B+C	A+C	A+B	B+C	A+C	
Intercept	123.0	109.2	137.8	7.530	4.983	3.276	
Maximum Bond Length for a H-O Bond	-77.10	-61.67	-76.40	-5.903	-3.291	-4.883	
Average Valence for a C Atom	-12.54	-12.61	-16.19	-4.154	-3.092	-1.789	
NSASA-3, Mulliken	-0.053	-0.043	-0.049	-2.677	-2.601	-2.993	
Zefirov Charge Polarity/ R ²	438.9	422.4	436.1	3.067	3.544	3.504	
QSAR denkleminde yer alan moleküler tanımlayıcıların katsayı ve t-testi değerleri							

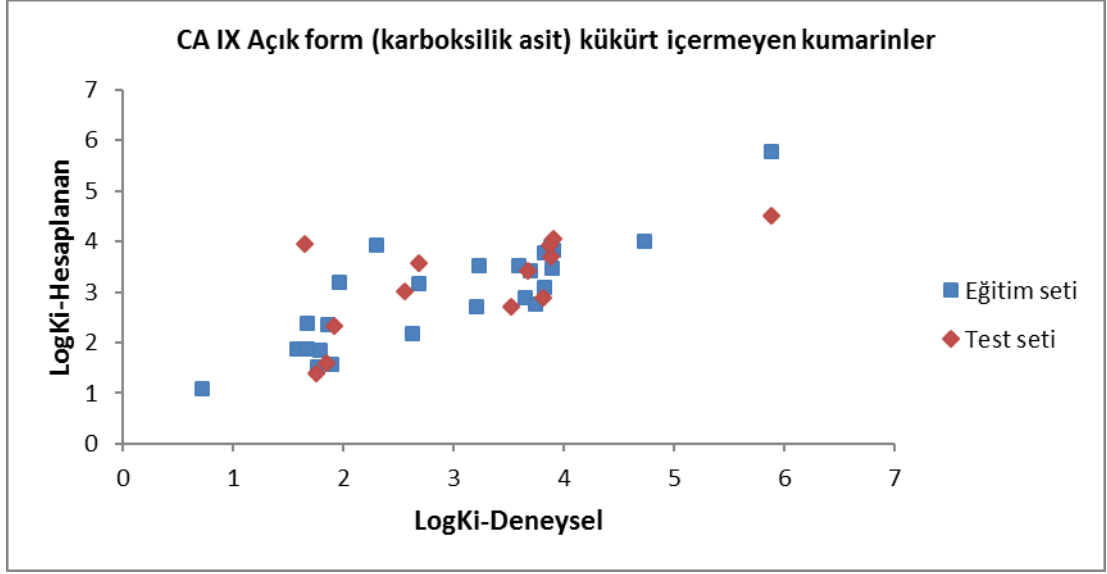
Çizelge 4.24’de ise Çizelge 4.23’de eğitim seti (A+B), test seti C seçildiği durumda ortaya çıkan tahmin değerleri verilmektedir. Mavi ile gösterilen moleküller eğitim setini kırmızı ile gösterilenler test setini ifade etmektedir. Şekil 4.23’de Çizelge 4.24’de verilen deneysel K_i değerlerinin hesaplanan K_i değerlerine karşı grafiği

gösterilmektedir. Şekil 4.24’da Çizelge 4.23’de eğitim seti (A+B), test seti C seçildiği durumdan elde edilen modelin uygulanabilirlik sınırları gösterilmektedir.

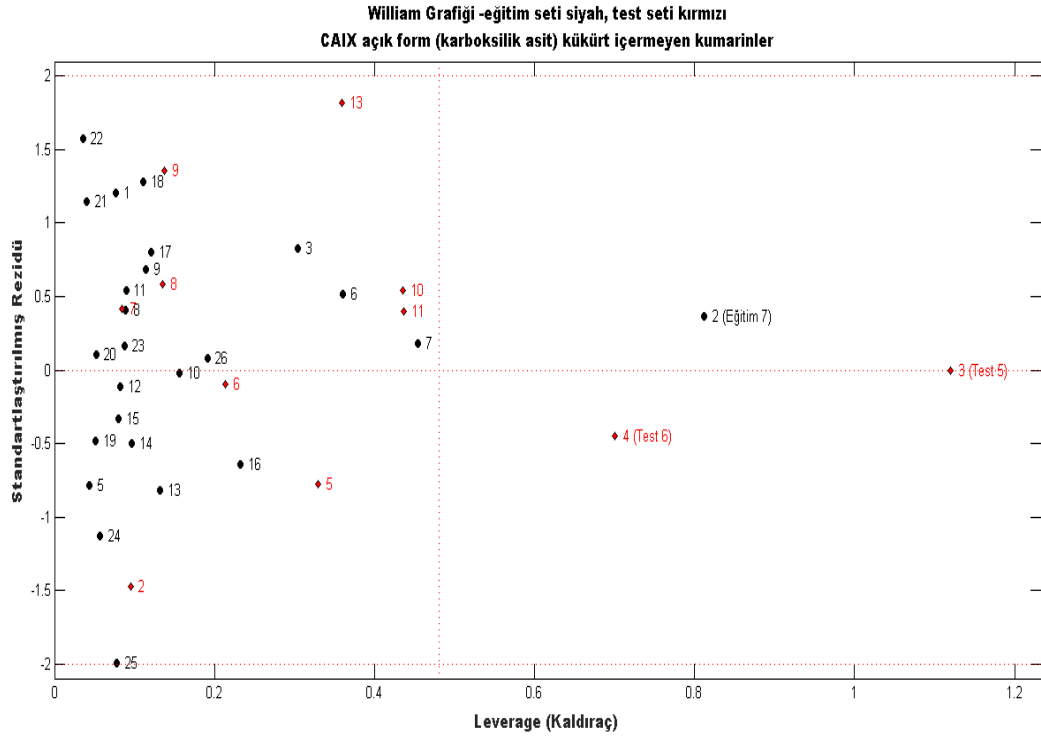
Uygulanabilirlik sınırı incelendiğinde CA IX karboksilik asit global modelinde olduğu gibi ekstrim aktiviteye sahip olan 2 adet test seti molekülünün ve eğitim setindeki ekstrim aktivite değerine sahip olan 1 adet molekülün sapan değer olduğu görülmektedir.

Çizelge 4.24: CA IX İzoenzimi karboksilik asit kükürt içermeyen kumarinler QSAR modeli tahmin çizelgesi

Molekül	LogKi-Deneysel	LogKi-Hesap	Molekül	LogKi-Deneysel	LogKi-Hesap
1	4,736	3,994	74	3,820	3,099
3	5,885	5,783	75	3,748	2,753
5	2,621	2,178	76	3,869	3,769
6	2,297	3,937	79	1,672	2,379
8	2,679	3,172	81	1,968	3,200
10	3,692	3,426	82	3,826	3,779
30	3,904	3,819	4	5,885	4,518
32	1,765	1,515	7	2,683	3,585
33	3,892	3,480	9	2,555	3,015
35	3,880	3,893	29	3,905	4,063
38	1,894	1,564	31	1,919	2,326
41	1,787	1,856	34	3,869	3,926
42	1,859	2,351	39	1,850	1,595
43	1,577	1,884	40	1,754	1,404
44	1,669	1,873	67	3,519	2,709
45	0,716	1,077	72	3,672	3,410
66	3,204	2,722	73	3,886	3,693
68	3,653	2,877	78	1,653	3,960
69	3,230	3,533	80	3,813	2,879
70	3,591	3,523			



Şekil 4.23: CA IX izoenzimi karboksilik asit kükürt içeren kumarinler QSAR modeli deneysel K_i 'nin hesaplanan K_i değerine karşı grafiği. (Bu grafik Çizelge 4.23'de verilen eğitim setinin A+B, test setinin C olduğu veriden elde edilmiştir).



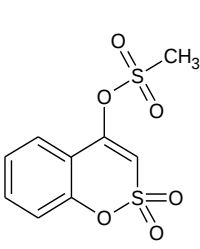
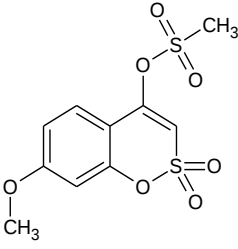
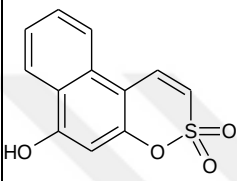
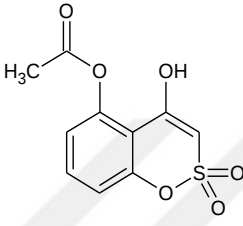
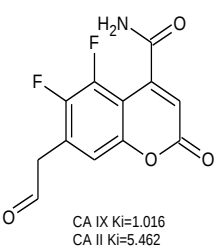
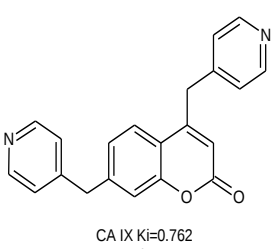
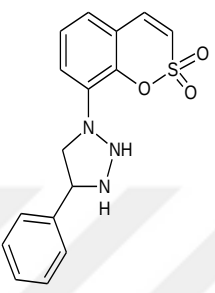
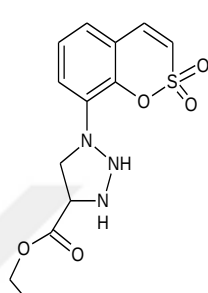
Şekil 4.24: CA IX izoenzimi karboksilik asit kükürt içermeyen kumarinler QSAR modeli sapma değer tayini.

4.4. QSAR Modelleri İle Seçiciliği Yüksek Sanal Kumarin Türevleri

Yukarıda sunulan QSAR modelleri oluşturulduktan sonra temel kumarin yapısı üzerine kimyasal çeşitleme (fonksiyonel grupların eklenmesi/çıkarılması) yoluyla 250'den fazla sanal kumarin türevi oluşturuldu. Bu sanal moleküllerin tamamına QSAR modelleri oluşturmada kullandığımız teori seviyesi (DFT, B3LYP fonksiyoneli ve 6-311 G (d.p) temel seti) ile hesaplamalar yaparak moleküler tanımlayıcıları hesaplandı. Sanal moleküllerin hesaplanan moleküler tanımlayıcıları QSAR modellerinin Codessa III dosyalarına yüklenerek CA II ve CA IX için aktiviteleri hesaplandı. Aktivite verileri seçicilik gösterenler (K_i -CA II/ K_i -CA IX değeri 100'den büyük olanlar) ayrılarak aktiviteyi hesaplayan QSAR modelinin uygulanabilirlik sınırları içinde olup olmadığı kontrol edildi. Uygulanabilirlik sınırları dışında kalan çok sayıda sanal molekül ekarte edildi.

Aşağıdaki Çizelge 4.25'de ekarte edilen moleküllerden arta kalan seçiciliği en yüksek moleküller verilmektedir. Çizelge 4.25'de verilen CA II aktif olan moleküllerin CA IX'a karşı aktiviteleri en az yüz kat daha azdır. Durum CA IX aktif moleküller içinde benzer şekilde CA II'ye karşı aktiviteleri en az yüz kat daha azdır. Benzer durum inaktif moleküller için de geçerli. CA II ya da CA IX için inaktif olan bir molekülün aktivitesi diğeri için en az yüz kat daha fazla olanlar seçici inhibitor olarak bu Çizelgeye alındı.

Çizelge 4.25: QSAR modelleri ile tasarlanan sanal kumarin türevleri

CA II Aktif / CA IX İnaktif	CA IX Aktif / CA II İnaktif
 CA II Ki=0.140 CA IX Ki=2.720  CA II Ki= -0.139 CA IX Ki=2.773  CA II Ki= -0.555 CA IX Ki=2.847  CA II Ki=0.271 CA IX Ki=3.069	 CA IX Ki=1.016 CA II Ki=5.462  CA IX Ki=0.762 CA II Ki=6.430  CA IX Ki=0.073 CA II Ki=4.875  CA IX Ki=1.307 CA II Ki=5.166

5. SONUÇ VE ÖNERİLER

Bu çalışma kapsamındaki sonuçlarda kumarinler ile hem CA II'nin hem de CA IX'un inhibisyonunu modelleyen 12 adet istatistiki açıdan yeterli uygulanabilirlik sınırları belirlenmiş model elde edilmiştir. Bu modeller ile hedeflenen sayıda seçiciliği yüksek sanal kumarin türevinin tasarımı yapılmıştır. Ayrıca tasarlanan bu moleküllerin tasarlamada kullanılan modelin uygulanabilirlik sınırları içerisinde olduğu teyit edilmiştir. Daha fazla sayıda sanal molekül üretilip, bunların aktiviteleri elde edilen modellerle belirlenip, seçiciliği yüksek olanların uygulanabilirlik sınırları içinde olup olmadığı araştırılarak tasarlanmış molekül sayısı artırılabilir.

Bu alanda gelecekte çalışılması gereken bir başka başlık da kumarin türevlerinin aktif site içerisinde lakton halkalarının açılma mekanizmasının modellenmesi olabilir. Bu açılma çok büyük bir olasılıkla aktif site içerisindeki bir amino asitin asisti ile gerçekleşmekte olabilir. Beklide bu açılma reaksiyonunun hızı ile bağlanma sabiti arasında bir ilişki söz konusu olabilir.

6. KAYNAKLAR

- Agrawal, V.K., Sharma, R. ve Khadikar, P.V., 2002. QSAR studies on carbonic anhydrase inhibitors: a case of ureido and thioureido derivatives of aromatic/heterocyclic sulfonamides, *Bioorganic & Medicinal Chemistry*, 10, 9, 2993-2999.
- Alterio, V., Di Fiore, A., D'Ambrosio, K., Supuran, C.T. ve De Simone, G., 2012. Multiple binding models of inhibitors carbonic anhydrases: how to design specific drugs targeting 15 different isoforms?, *Chemical Review*, 112, 8, 4421-4468.
- Atkins, P.W., 1991. *Quanta: A Handbook of Concepts*, Oxford University Press, Oxford.
- Bader, R.F., Carroll, M.T., Cheeseman, J.R. ve Chang, C., 1987. Properties of atoms in molecules: atomic volumes, *Journal of the American Chemical Society*, 109, 26, 7968-7979.
- Becke, A.D., 1988. Density-functional exchange energy approximation with correct asymptotic behavior, *Physical Review A*, 38, 3098-3100.
- BioVia, Dassault Systèmes BIOVIA, 2016. Discovery Studio Modeling Environment, Release 2017, San Diego, Dassault Systèmes, 2016. Adres: <http://accelrys.com/products/collaborative-science/biovia-discovery-studio/visualization/download.php>
- Brahimi-Horn, M.C. ve Pouyssegur, J., 2007. Hypoxia in cancer cell metabolism and pH regulation, *Essays Biochemistry*, 43, 165-178.
- Breiman, L. ve Friedman, J.H., 1985. Estimating Optimal Transformations for Multiple Regression and Correlation, *Journal of the American Statistical Association*, 80, 391, 580-619.
- Clare. B.W., 1994. Frontier orbital energies in quantitative structure-activity relationship: a comparison of quantum chemical methods, *Theoretica Chimica Acta*, 87, 415-430.

- Clare, B.W. ve Supuran, C.T., 1994. Carbonic-Anhydrase Activators. Part 3. Structure-Activity Correlations of a Series of Isozyme-II Activators, *Journal of Pharmaceutical Sciences*, 83, 6, 768-773.
- Clare, B.W. ve Supuran, C.T., 1997. Carbonic anhydrase inhibitors. Part 41. Quantitative structure-activity correlations involving kinetic rate constants of 20 sulfonamide inhibitors from a non-congeneric series, *European Journal of Medicinal Chemistry*, 32, 4, 311-319.
- Clare, B.W. ve Supuran, C.T., 1999. Carbonic anhydrase inhibitors. Part 61. Quantum chemical QSAR of a group of benzenedisulfonamides, *European Journal of Medicinal Chemistry*, 34, 6, 463-474.
- Clare, B.W. ve Supuran, C.T., 2000. Carbonic anhydrase inhibitors. Part 86. A QSAR study on some sulfonamide drugs which lower intra-ocular pressure, using the ACE non-linear statistical method, *European Journal of Medicinal Chemistry*, 35, 9, 859-865.
- Clare, B.W. ve Supuran, C.T., 2005. A physically interpretable quantum-theoretic QSAR for some carbonic anhydrase inhibitors with diverse aromatic rings, obtained by a new QSAR procedure, *Bioorganic & Medicinal Chemistry*, 13, 2197-2211.
- Clare, B.W. ve Supuran, C.T., 2005. Predictive flip regression: A technique for QSAR of derivatives of symmetric molecules, *Journal of Chemical Information and Modeling*, 45, 5, 1385-1391.
- CODESSATMIII, Semichem, 2002. 7204, Mullen, Shawnee, KS, Copyright© Semichem and University of Florida, USA.
- Cros, A., 1863. Action de l'alcoholamylique sur l'organisme. Strasbourg, Thesis, University of Strasbourg, Strasbourg.
- Crum-Brown, A. ve Froser, T.R., 1868. On the connection between chemical constitution and physiological action. Pt 1. On the physiological action of the salts of the ammonium bases, derived from Strychnia, Brucia, Thebia, Codeia, Morphia, and Nicotia. *T Roy Soc Edin.* 25, 151-203.
- Csizmadia, I. G., 1976. Theory and Practice of MO Calculations on Organic Molecules, Elsevier, Amsterdam.
- De Benedetti, P.G., Menziani, M.C. ve Frassinetti, C., 1985. A Quantum Chemical QSAR Study of Carbonic Anhydrase Inhibition by Sulfonamides Sulfonamide Carbonic Anhydrase Inhibitors: Quantum Chemical QSAR, Quantitative Structure-Activity Relationships, 4, 1, 23-28.

- De Benedetti, P.G., Menziani, M.C. ve Frassinetti, C., 1987. A Quantum Chemical QSAR Study of Carbonic Anhydrase Inhibition by Sulfonamides Sulfonamide Carbonic Anhydrase Inhibitors: Quantum Chemical QSAR, Quantitative Structure-Activity Relationships, 6, 51-53.
- De Simone, G., Scozzafava, A. ve Supuran, C.T., 2009. Which carbonic anhydrases are targeted by the antiepileptic sulfonamides and sulfonamates?, Chemical Biology and Drug Design, 74, 3, 317-321.
- Dimitrov, S., Dimitrova, G., Pavlov, T., Dimitrova, N., Patlewicz, G., Niemela, J., Mekenyan, O., 2005. A stepwise approach for defining the applicability domain of SAR and QSAR models. Journal of chemical information and modeling, 45, 4, 839-849.
- Eriksson, L., Jaworska, J., Worth, A.P., Cronin, M.T., McDowell, R.M., Gramatica, P., 2003. Methods for reliability and uncertainty assessment and for applicability evaluations of classification and regression-based QSARs, Environmental Health Perspectives, 111, 10, 1361-1375.
- Esposito, E.X., Baran, K., Kelly, K. ve Madura, J.D., 2000. Docking of sulfonamides to carbonic anhydrase II and IV, Journal of Molecular Graphics and Modelling, 18, 3, 283-289.
- Ferraroni, M., Carta, F., Scozzafava, A. ve Supuran, C.T., 2016. Thioxocoumarins Show an Alternative Carbonic Anhydrase Inhibition Mechanism Compared to Coumarins, Journal of Medicinal Chemistry, 59, 1, 462-473.
- Foresman, J., Frish, E., 1996. Exploring with chemistry. Gaussian Inc., Pittsburg, USA.
- Gao, B.B., Clermont, A., Rook, S., Fonda, S.J., Srinivasan, V.J., Wojtkowski, M., Fujimoto, J.G., Avery, R.L., Arrigg, P.G., Bursell, S.E., Aiello, L.P. ve Feener, E.P., 2007. Extracellular carbonic anhydrase mediates hemorrhagic retinal and cerebral vascular permeability through prekallikrein activation, Nature Medicine, 13, 2, 383-392.
- Gaussian 09, Revision E.01, Frisch, M.J., Trucks, G.W., Schlegel, H.B., Scuseria, G.E., Robb, M.A., Cheeseman, J.R., Scalmani, G., Barone, V., Mennucci, B., Petersson, G.A., Nakatsuji, H., Caricato, M., Li, X., Hratchian, H.P., Izmaylov, A.F., Bloino, J., Zheng, G., Sonnenberg, J.L., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Vreven, T., Montgomery, J.A., Jr., Peralta, J. E., Ogliaro, F., Bearpark, M., Heyd, J.J., Brothers, E., Kudin, K.N., Staroverov, V.N., Kobayashi, R., Normand, J., Raghavachari, K., Rendell, A.,

- Burant, J.C., Iyengar, S.S., Tomasi, J., Cossi, M., Rega, N., Millam, J.M., Klene, M., Knox, J.E., Cross, J.B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Stratmann, R.E., Yazyev, O., Austin, A.J., Cammi, R., Pomelli, C., Ochterski, J.W., Martin, R.L., Morokuma, K., Zakrzewski, V.G., Voth, G. A., Salvador, P., Dannenberg, J.J., Dapprich, S., Daniels, A.D., Farkas, Ö., Foresman, J.B., Ortiz, J.V., Cioslowski, J., Fox, D.J., 2009. Gaussian, Inc., Wallingford CT.
- Gohlke, H. ve Klebe, G., 2002. Drug Score meets CoMFA: adaptation of fields for molecular comparison (AFMoC) or how to tailor knowledge-based pair-potentials to a particular protein, *Journal of Medicinal Chemistry*, 45, 19, 4153-4170.
- Golden, J.B., 2003. Prioritizing the human genome: knowledge management for drug discovery, *Current Opinion in Drug Discovery and Development*, 6, 3, 310-316.
- Hen, N., Bialer, M., Yagen, B., Maresca, A., Aggarwal, M., Robbins, A.H., McKenna, R., Scozzafava, A. ve Supuran, C.T., 2011. Anticonvulsant 4-aminobenzenesulfonamide derivatives with branched-alkylamide moieties: X-ray crystallography and inhibition studies of human carbonic anhydrase isoforms I, II, VII, and XIV, *Journal of Medicinal Chemistry*, 54, 11, 3977-3981.
- Hillebrecht, A., Supuran, C.T. ve Klebe, G., 2006. Integrated approach using protein and ligand information to analyze selectivity- and affinity-determining features of carbonic anhydrase isozymes *Chemical Medicinal Chemistry*, 1, 8, 839-853.
- Hohenberg, K. ve Kohn, W., 1964. In homogeneous Electron Gas, *Physical Review*, 136, B864.
- Huang, H., Pan, X., Tan, N., Zeng, G. ve Ji, C., 2007. 3D-QSAR study of sulfonamide inhibitors of human carbonic anhydrase II, *European Journal of Medicinal Chemistry*, 42, 3, 365-372.
- Imming, P., Sinning, C. ve Meyer, A., 2006. Drugs, their targets and the nature and number of drug targets, *Nature Reviews. Drug Discovery*, 5, 10, 821-834.
- Jaiswal, M., Khadikar, P.V., Scozzafava, A. ve Supuran, C.T., 2004. Carbonic Anhydrase inhibitors: the first QSAR study on inhibition of tumor-associated isoenzyme IX with aromatic and heterocyclic sulfonamides, *Bioorganic & Medicinal Chemistry Letters*, 14, 12, 3283-3290.
- Jaworska, J., Nikolova-Jeliazkova, N. ve Aldenberg, T., 2005. QSAR applicability

domain estimation by projection of the training set descriptor space: a review
Atla- Nottingham, 33, 5, 445.

Khadikar, P. V., Sharma, V., Karmarkar, S. ve Supuran, C.T., 2005. QSAR studies on benzenesulfonamide carbonic anhydrase inhibitors: need of hydrophobic parameter for topological modeling of binding constants of sulfonamides to human CA-II, *Bioorganic & Medicinal Chemistry Letters*, 15, 4, 923-930.

Keha, E. ve Kührevioğlu, İ., 2004. *Biyokimya* (1.Basım). Aktif Yayınevi, Erzurum

Kirpichenok, M.A. ve Zefirov, N.S., 1987. Electronegativity and molecular geometry. I. General principles of the method and analysis of the effect of short-range electrostatic interactions on bond lengths in organic molecules, *Zhurnal Organicheskoi Khimii*, 23, 4, 673–703.

Kohn, W. ve Sham, L.J., 1965. Self-consistent equations including exchange and correlation effects, *Physical Review*, 140, 4A, A1133.

Landry, Y. ve Gies, J.P. 1965. Drugs and their molecular targets: an updated overview, *Fundamental and Clinical Pharmacology*, 22, 1, 1-18.

Leach, A.R. ve Gillet, V.J., 2007. *An introduction to chemoinformatics*, Springer Science and Business Media, London, New York.

Leszczynski, J., 2010. *Recent Advances in QSAR Studies Methods and Applications*, Editörler: Puzyn, T., Leszczynski, J. ve Cronin, T. D. Springer Science and Business Media, London, New York.

Liljas, A., Kannan, K.K., Bergsten, P.C., Waara, I., Fridborg, K., Strandberg, B., Carlbom, U., Jarup, L., Lejvgren, S. ve Petef, M., 1972. Crystal structure of human carbonic anhydrase, *Nature-New Biology*, 235, 131–137.

Maresca, A., Scozzafava, A., ve Supuran, C.T., 2010. 7, 8-Disubstituted-but not 6, 7 disubstituted coumarins selectively inhibit the transmembrane, tumor-associated carbonic anhydrase isoforms IX and XII over the cytosolic ones I and II in the low nanomolar/subnanomolar range, *Bioorganic & Medicinal Chemistry Letters*, 20, 24, 7255-7258.

Maresca, A., ve Supuran, C. T., 2010. Coumarins incorporating hydroxy-and chloro-moieties selectively inhibit the transmembrane, tumor-associated carbonic anhydrase isoforms IX and XII over the cytosolic ones I and II, *Bioorganic & Medicinal Chemistry Letters*, 20, 15, 4511-4514.

Maresca, A., Temperini, C., Vu, H., Pham, N.B., Poulsen, S.A., Scozzafava, A., Quinn, R.J. ve Supuran, C.T., 2009. Non-zinc mediated inhibition of carbonic

- anhydrases: coumarins are a new class of suicide inhibitors, *Journal of the American Chemical Society*, 131, 8, 3057-3062.
- Maresca, A., Temperini, C., Pochet, L., Masereel, B., Scozzafava, A. ve Supuran, C.T., 2010. Deciphering the mechanism of carbonic Anhydrase inhibition with coumarins and thiocoumarins, *Journal of Medicinal Chemistry*, 53(1), 335-344. Erratum in: 2015.*Journal of Medicinal Chemistry*, 58, 14, 5689.
- Meldrum, N.U. ve Roughton, F.J.W., 1933. Carbonic anhydrase, its preparation and properties, *The Journal of Physiology*, 80, 2, 113-142.
- Meyer, H., 1899. Zur Theorie der Alkohol narkose. *Archives Experiments Pathol Pharmacol*, 42, 109-18.
- Mulliken, R.S., 1955. Electronic Population Analysis on LCAO–MO Molecular Wave Functions. I, *The Journal of Chemical Physics*, 23, 1833-1840.
- Osmialowski, K., Halkiewicz, J., Radecki, A. ve Kaliszan, R., 1985. Quantum chemical parameters in correlation analysis of gas—liquid chromatographic retention indices of amines, *Journal of Chromatography*, 346, 53-60.
- Overington, J.P., Al-Lazikani, B. ve Hopkins, A.L., 2006. How many drug targets are there?, *Nature Review. Drug Discovery*, 5, 12, 993-996.
- Overton, C. E., 1901. *Studienüberdie Narkose*. Jena: Fischer.
- Pastorekova, S., Parkkila, S. ve Zavada, J., 2006. Tumor-associated carbonic anhydrases and their clinical significance, *Advances in Clinical Chemistry*, 42, 167-216.
- Pastorekova, S., Barathova, M., Kopacek, J. ve Pastorek, J., 2009. Carbonic Anhydrase Inhibitors Targeting Cancer: Therapeutic, Immunologic, and Diagnostic Tools Targeting Isoforms IX and XII, *Drug Design of Zinc-Enzyme Inhibitors*. Editörler: Supuran, C. T. ve Winum, J.-Y. New Jersey: John Wiley and Sons.
- Perdew, J.P. ve Wang, Y., 1992. Accurate and Simple Analytic Representation of the Electron Gas Correlation Energy, *Physical Review*, 45, 13244.
- Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. ve Ferrin, T.E., 2013. UCSF Chimera a visualization system for exploratory research and analysis, *Journal of Computational Chemistry*, 25, 13, 1605-1612.

- Renzi, G., Scozzafava, A. ve Supuran, C.T., 2000. Carbonic anhydrase inhibitors: Topical sulfonamide antiglaucoma agents incorporating secondary amine moieties, *Bioorganic and Medicinal Chemistry*, 10, 673-676.
- Richet, M.C., 1893. Note sur le rapport entre la toxicité et les propriétés physiques des corps. *Compt Rend Society Biology, Paris*, 45, 775-776.
- Sanniraghi, A.B., 1992. AB Initio Molecular Orbital Calculations of Bond Index and Valency, *Advances in Quantum Chemistry*, 23, 301-351.
- Scholz, T.H., Sondey, J.M., Randall, W.C., Schwam, H., Thompson, W.J., Mallorga, P.J., Sugrue, M.F. ve Graham, S.L., 1993. Sulfonyl methane sulfonamide inhibitors of carbonic Anhydrase, *Journal of Medicinal Chemistry*, 36, 2134-2141.
- Stanforth, R.W., 2008. Extending K-Means Clustering for Analysis of Quantitative Structure Activity Relationships (QSAR), School of Computer Science and Information Systems, Birkbeck College, University of London ,Submitted for the degree of Doctor of Philosophy of the University of London.
- Stanton, D.T., Egolf, L.M., Jurs, P.C. ve Hicks, M.G., 1992. Computer-assisted prediction of normal boiling points of pyrans and pyrroles, *Journal of Chemical Information and Modeling*, 32, 306-316.
- Stanton, D.T. ve Jurs, P.C., 1990. Development and use of charged partial surface area structural descriptors in computer-assisted quantitative structure-property relationship studies, *Analytical Chemistry*, 62, 2323-2329.
- Supuran, C.T. ve Clare, B.W., 1995. Carbonic-Anhydrase Inhibitors.24. A Quantitative Structure-Activity Relationship Study of Positively Charged Sulfonamide Inhibitors, *European Journal of Medicinal Chemistry*, 30, 687-696.
- Supuran, C.T. ve Clare, B.W., 1998. Carbonic anhydrase inhibitors - Part 47: Quantum chemical quantitative structure-activity relationships for a group of sulfanilamide Schiff base inhibitors of carbonic anhydrase, *European Journal of Medicinal Chemistry*, 33, 489-500.
- Supuran, C.T. ve Clare, B.W., 1999. Carbonic anhydrase inhibitors - Part 57: Quantum chemical QSAR of a group of 1, 3, 4-thiadiazole- and 1, 3, 4-thiadiazoline disulfonamides with carbonic anhydrase inhibitory properties, *European Journal of Medicinal Chemistry*, 34, 41-50.
- Supuran, C.T. ve Clare, B.W., 2001. Orbital symmetry in QSAR: some Schiff's base inhibitors of carbonic anhydrase, SAR and QSAR in Environmental Research, 12, 1, 17-29.

- Supuran, C.T., Scozzafava, A. ve Conway, J., 2004. Carbonic Anhydrase: Its Inhibitors and Activators, Boca Raton: CRC Press.
- Supuran, C.T., 2008. Carbonic anhydrases: novel therapeutic applications for inhibitors and activators, *Nature Review Drug Discovery*, 7, 2, 168-181.
- Supuran, C.T., Di Fiore, A. ve De Simone, G., 2008. Carbonic anhydrase inhibitors as emerging drugs for the treatment of obesity, *Expert Opinion on Emerging Drugs*, 13, 2, 383-392.
- Supuran, C.T., 2009. Carbonic Anhydrases as Drug Targets: General Presentationin, *Drug Design of Zinc-Enzyme Inhibitors*. Editörler: Supuran C. T.ve Winum J.-Y. New Jersey: John Wiley and Sons.
- Supuran C.T.ve Winum J.Y., 2009. Introduction to Zinc Enzymes as Drug Targets. *Drug Design of Zinc-Enzyme Inhibitors*. Editörler: Supuran C. T.ve Winum J.-Y. New Jersey: John Wiley and Sons.
- Supuran, C.T., 2016. How many carbonic anhydrase inhibition mechanisms exist?, *Journal of Enzyme Inhibition and Medicinal Chemistry*, 31, 3, 345-360.
- Tanc, M., Carta, F., Bozdag, M., Scozzafava, A., ve Supuran, C.T., 2013. 7-Substituted-sulfocoumarins are isoform-selective, potent carbonic anhydrase II inhibitors, *Bioorganic & Medicinal Chemistry*, 21, 15, 4502-4510.
- Tars, K.,Vullo, D., Kazaks, A., Leitans, J., Lends, A., Grandane, A., Zalubovskis, R., Scozzafava, A. ve Supuran, C.T., 2013. Sulfocoumarins (1,2-benzoxathiine-2,2-dioxides): a class of potentand isoform-selective inhibitors of tumor-associated carbonic anhydrases, *Journal of Medicinal Chemistry*, 56, 1, 293-300.
- Testa, B. ve Purcell, W.P., 1978. A QSAR study of sulfonamide binding to carbonic anhydrase as test of steric models, *European Journal of Medicinal Chemistry*, 13, 6, 509-514.
- Thakur, A., Thakur, M., Khadikar, P.V., Supuran, C.T. ve Sudele, P., 2004. QSAR study on benzene sulphonamide carbonic anhydrase inhibitors: topological approach using Balaban index, *Bioorganic & Medicinal Chemistry*, 12, 4, 789-793.
- Tuccinardi, T., Nuti, E., Ortore, G., Supuran, C.T., Rossello, A. ve Martinelli, A., 2007. Analysis of human carbonic anhydrase II: docking reliability and receptor-based 3D-QSAR study, *Journal of Chemical Information and Modelling*, 47, 2, 515-525.

- Tuccinardi, T., Ortore, G., Rossello, A., Supuran, C.T. ve Martinelli, A., 2007. Homology modeling and receptor-based 3D-QSAR study of carbonic anhydrase IX, *Journal of Chemical Information and Modelling*, 47, 6, 2253-2262.
- Wagner, J., Avvaru, B.S., Robbins, A.H., Scozzafava, A., Supuran, C.T., ve McKenna, R., 2010. Coumarinyl-substituted sulfonamides strongly inhibit several human carbonic anhydrase isoforms: solution and crystallographic investigations, *Bioorganic & Medicinal Chemistry*, 18, 14, 4873-4878.
- Weast, R.C., 1974. *Handbook of Chemistry and Physics*, Cleveland, Ohio, CRC Press, 109-112.
- Weber, A., Böhm, M., Supuran, C.T., Scozzafava, A., Sotriffer, C.A. ve Klebe, G., 2006. 3D QSAR selectivity analyses of carbonic anhydrase inhibitors: insights for the design of isozyme selective inhibitors, *Journal of Chemical Information and Modelling*, 46, 6, 2737-2760.
- Yarvet, Y.I.K., Olivson, A.I., Arguraia, R.K. ve Lippmaa, E.T., 1985. Study of molecular dynamics of bovine carbonic anhydrase B by ¹³C-NMR in two strongly polarizing magnetic fields. I. Intramolecular mobility of polypeptide chains of the native enzyme, *Journal of Molecular Biology*, 19, 5, 1358-1366.
- Zefirov, N.S., Kirpichenok, M.A., Izmailov, F.F. ve Trofimov, M.I., 1987. Scheme for the calculation of the electronegativities of atoms in a molecule in the framework of Sanderson's principle, *Doklady Akademii Nauk SSSR*, 296, 883-887.

ÖZGEÇMİŞ

Adı ve Soyadı : Zekiye Eda BOYLUĞ

Doğum Tarihi ve Yeri: 14.07.1989 / Gaziantep

E-posta adresi : zekiye_boyluğ@hotmail.com

EĞİTİM BİLGİLERİ (Kurum ve Yıl)

Lisans : Harran Üniversitesi Fen-Edebiyat Fakültesi Fizik Bölümü (2007-2012)

Yüksek Lisans : Aksaray Üniversitesi Fizik Ana Bilim Dalı (2013-2016)
Aksaray Üniversitesi İlköğretim Fen Bilgisi Eğitim Ana Bilim Dalı (2014-....)

Seminer ve Kongreler

11th International Symposium on Pharmaceutical Sciences (ISOPS-11) in Ankara, June 9-12, 2015

3rd International BAU Drug Design Congress in Istanbul, October 1-3, 2015

II. International Turkish Congress on Molecular Spectroscopy (TURCMOS 2015) in Antalya, September 13-18, 2015

Akdeniz Moleküler Modelleme ve Spektroskopi Derneği, Proje Hazırlama Eğitimi Semineri, 18 Eylül 2015