T.C
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRON KONFORMASYONEL-GENETİK ALGORİTMA QSAR METODU İLE PENİSİLİN TÜREVLERİNE AİT FARMAKOFOR GRUBUNUN BELİRLENMESİ VE BİYOAKTİVİTE HESABI

Tezi Hazırlayan
Ersin YANMAZ

Tezi Yöneten
Prof. Dr. Emin SARIPINAR

Kimya Anabilim Dalı
Yüksek Lisans Tezi

Temmuz 2010
KAYSERİ

T.C
ERCİYES ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

ELEKTRON KONFORMASYONEL-GENETİK ALGORİTMA QSAR METODU İLE PENİSİLİN TÜREVLERİNE AİT FARMAKOFOR GRUBUNUN BELİRLENMESİ VE BİYOAKTİVİTE HESABI

Tezi Hazırlayan

Ersin YANMAZ

Tezi Yöneten

Prof. Dr. Emin SARIPINAR

Kimya Anabilim Dalı

Yüksek Lisans Tezi

Bu çalışma TUBİTAK tarafından 107T385 nolu Proje ile desteklenmiştir

Temmuz 2010

KAYSERİ

Prof. Dr. Emin SARIPINAR danışmanlığında **Ersin YANMAZ** tarafından hazırlanan **"Elektron Konformasyonel-Genetik Algoritma Qsar Metodu ile Penisilin Türevlerine ait Farmakofor Grupların Belirlenmesi ve Biyoaktivite Hesabı "** adlı bu çalışma, jürimiz tarafından Erciyes Üniversitesi Fen Bilimleri Enstitüsü Kimya Anabilim Dalında **Yüksek Lisans** tezi olarak kabul edilmiştir.

16/07/2010

JÜRİ:

Başkan: Prof. Dr. Şaban PATAT

Üye : Prof. Dr. Emin SARIPINAR

Üye : Doç. Dr. İlhan Özer İLHAN

ONAY:

Bu tezin kabulü, Enstitü Yönetim Kurulunun 27/07/2010 tarih ve 2010/25-04 sayılı kararı ile onaylanmıştır.

27./07./2010

Prof. Dr. Nusret AYYILDIZ

Enstitü Müdürü

TEŞEKKÜR

Tez çalışmalarım süresince geniş bilgi ve deneyimleri ile her konuda yol gösterici olan, yardımlarını ve desteğini esirgemeyen sayın hocam Prof. Dr. Emin SARIPINAR' a sonsuz saygı ve teşekkürlerimi sunarım.

Yüksek lisans eğitimim süresince emeği geçen bütün hocalarıma sonsuz saygı ve teşekkürlerimi sunarım.

Ayrıca tez çalışmam boyunca yardımlarını esirgemeyen değerli çalışma arkadaşlarım Nazmiye Geçen, Kader Şahin, Mustafa Yıldırım, Fatih Çopur, Semiha Aydın ve hayatım boyunca yanımda olan ve her konuda benden maddi ve manevi desteklerini esirgemeyen sevgili aileme minnettarlıklarımı bildiririm.

ELEKTRON KONFORMASYONEL-GENETİK ALGORİTMA QSAR METODU İLE PENİSİLİN TÜREVLERİNE AİT FARMAKOFOR GRUBUNUN BELİRLENMESİ VE BİYOAKTİVİTE HESABI

Ersin YANMAZ
Erciyes Üniversitesi, Fen Bilimleri Enstitüsü
Yüksek Lisans Tezi, Temmuz 2010
Tez Danışmanı: Prof. Dr. Emin SARIPINAR

ÖZET

Bu çalışmada Penisilin türevi bileşiklerin antimikrobiyal aktivitesinden sorumlu farmakofor gruplarının belirlenmesi ve bileşik serilerinin nicel aktivite tahmini elektron konformasyonel genetik algoritma (EC-GA) metodu kullanılarak yapılmıştır. Bunun için genel amaçlı farmakofor grubunu, moleküler tanımlayıcılar (parametreler) ve aktivite hesaplamaları için kişisel bilgisayarlarda çalışan kapsamlı bir 4D-QSAR program paketi (EMRE, ECSP ve Aktivite kodları) geliştirilmiştir. Çalışmanın ilk kısmında, serilerdeki her bir bileşiğin konformasyonel analiz ve kuantum kimyasal hesaplamaları PM3 metodu ile yapılmıştır. Bu hesaplamalar sonucunda elde edilen verilerden yararlanılarak EMRE programı ile her bir konformer için üç boyutlu elektron konformasyonel uygunluk matrisleri (ECMC) hazırlanmıştır. Daha sonra ECSP programı ile bu matrisler belirli tolerans değerlerinde karşılaştırılarak aktiviteden sorumlu olan farmakofor grup yani elektron konformasyonel alt matris (ECSA) belirlenmiştir. Penisilin serisinde bileşiklerin her bir konformeri için 646 moleküler parametre EMRE programı ile hazırlanmıştır. PAB aktivitelerine etki eden en önemli parametre gruplarının belirlenmesi ve teorik aktivite değerlerini bulmak için genetik algoritma ve doğrusal olmayan en küçük kareler (lsqnonlin) yöntemi kullanılmıştır. Penisilin türevlerinde seriler eğitim ve test seti olmak üzere sınıflandırılarak optimizasyonu yapılmıştır. Bu çalışmada üç farklı set ile hesaplamalar yapılmıştır. Bu setler için regrasyon katsayısıları eğitim seti için sırasıyla 0.852, 0.897 ve 0.862; test seti için sırasıyla 0.822, 0.881 ve 0.886 bulunmuştur. Bu setler için çapraz doğrulama katsayıları ise sırasıyla 0.641, 0.801 ve 0.698 bulunmuştur. Bu çalışmada geliştirilen 4D-QSAR metodu ile oluşturulan modelin benzer yapılara sahip yeni bileşiklerin tasarlanmasında kullanılabileceği görülmüştür.

Anahtar Kelimeler: Penisilin, İlaç tasarımı, QSAR, Farmakofor, Elektron Konformasyonel Metot, Genetik Algoritma

PHARMACOPHORE IDENTIFICATION AND BIOACTIVITY PREDICTION FOR PENICILLIN DERIVATIVES BY ELECTRON CONFORMATIONAL-GENETIC ALGORITHM QSAR METHOD

Ersin YANMAZ
Erciyes University, Graduate School of Natural and Applied Sciences
M.Sc. Thesis, July 2010
Thesis Supervisor: Prof. Dr. Emin SARIPINAR

ABSTRACT

The electron-conformational genetic algorithm (EC-GA) method has been employed to reveal the pharmacophore (Pha) and to predict antimicrobial activity, studying in the class of penicillin derivatives. Hence we present a pharmacophore identification, molecular descriptor and activity calculation using a comprehensive 4D-QSAR program package (EMRE, ECSP and codes of activity) which runs on personal computers. In the first part, conformational analysis and electronic structure calculation of each of the molecules in the compound series are performed by PM3 method. The EC matrices of congruity (ECMCs) are constructed from data of conformational analysis and electronic structure calculation of each of the molecules in the compound series by EMRE programme. By compraring these matrices using ECSP programme for several acitve compounds of the training set a group of matrix elements is revealed that are common for these compounds within a minimum toerance, the EC submatix of acitivity (ECSA). 646 molecular descriptors are prepared for penicillin derivatives by EMRE. To select important variables for describing the inhibition activities and predict the theoretical activity, genetic algorithm and nonlinear least square regression methods (lsqnonlin) were applied for penicillin. Besides after compounds were divided into two subgroups of training and test set, optimization procedure was performed. In this study, calculations were made with three different sets. For this sets, regression coefficient values are 0.852, 0.897 and 0.862 for training; 0.822, 0.881 and 0.886 for test respectively. Cross validation coefficient values are 0.641, 0.801 and 0.698 respectively.

It was seen that the model constructed by 4D-QSAR method carried out in this study is applicable to design new congeneric compounds.

Keywords: Penicillin**,** Drug Design, QSAR, Pharmacophere, Electron Conformational Method, Genetic Algorith

İÇİNDEKİLER

KABUL VE ONAY……...i
TEŞEKKÜR i
ÖZET iii
ABSTRACT iv
TABLOLAR LİSTESİ..... vii
ŞEKİLLER LİSTESİ ix
KISALTMALAR VE SİMGELER xi

1. BÖLÜM 1
GİRİŞ 1
1.1. Giriş 1
1.2. İlaç 2
1.4. Hesaplamalı Kimya 5
1.5. Antibiyotikler 5
1.5.1. Antibiyotiklerin sınıflandırılması ve etki mekanizması 5
1.5.1.1. Penisilinler 6

2. BÖLÜM 8
QSAR METOTLARI 8
2.1. QSAR'a Genel Bakış 8
2.2. QSAR'ın Temel İlkeleri 9
2.3. QSAR Modelinin Geliştirilmesi 12
2.4. 2D-QSAR Analizi 15
2.5. 3D-QSAR Metodu 15
2.6. 4D, 5D ve 6D QSAR Metodları..... 16
2.7. Moleküler Modelleme ve QSAR 17
2.8. Biyoalıcı İle İlaç Molekülün Etkileşim Mekanizması 18

3. BÖLÜM 20
MATERYAL VE METOT..... 20
3.1. Materyal..... 20
3.2. Metot 20
3.3. Elektron Topolojik Metot 21
3.4. Elektron Konformasyonel Metot 22
3.4.1. ECMC Matrislerinin Oluşturulması..... 23
3.4.2. Farmakofor Tanımlama 23
3.4.3. APS ve AG Grupları..... 26
3.4.4. Aktivitenin Genel Formülü..... 27
3.5. Elektron Konformasyon-Genetik Algoritma (EC-GA) Metodu..... 30

3.5.1. EMRE Yazılım Sisteminin Geliştirilmesi 32
3.5.2. Adım 1 35
3.5.3. Adım 2 39
3.5.3.1. Boltzman Dağılımı 40
3.5.4. Adım 3 43
3.5.5. Adım 4 45
3.5.6. Adım 5 50
3.5.6.1. Moleküler Parametreler 51
3.5.7. Adım 6 55
3.5.7.1. Optimizasyon 56
3.5.7.2. Genetik Algoritma 58
3.5.7.2.1. Genetik Algoritmaların Temel Teoremi ve Basit Genetik Algoritma 59
3.5.7.2.2. Genetik Algoritmalarda Parametre Seçimi 63
3.5.7.2.3. Genetik Algoritmaların Uygulama Alanları 64
3.5.7.3. Genetik Algoritmanın Elektron Konformasyonel Yönteme Uygulanışı 64
3.5.7.4. İstatiksel Analiz 66
3.5.7.4.1. Çapraz Doğrulama (Cross Validation) 66
3.5.7.4.2. Dâhili doğrulama (İnternal Validation) 66
3.5.7.4.3. Harici Doğrulama (External Validation) 67

4. BÖLÜM 69

BULGULAR 69

4.1. ECMC Matrislerinin Hazırlanması ve Farmakofor Bulunması 69
4.2. Aktivite Hesaplamasında Kullanılan Parametreler 76
4.3. Penisilin Serinin Aktivite Hesaplaması 91
4.3.1. Penisilin Serisinin Eğitim ve Test Seti Olarak Hesaplanması 91
4.3.1.1. Penisilin Serisinde 60 Bileşik Eğitim ve 28 Bileşik Test Seti Olarak Aktivite Hesaplaması 92
4.3.1.2. Penisilin Serisinde 74 Bileşik Eğitim ve 13 Bileşik Test Seti Olarak Aktivite Hesaplaması 107

5. BÖLÜM 116

SONUÇ 116

5.1. EMRE, ECSP ve Matlab Yazılımları 116
5.2. Penisilin Türevlerinin Teorik Aktivite Hesaplama Sonuçları 117

KAYNAKLAR 134

ÖZGEÇMİŞ 142

TABLOLAR LİSTESİ

Tablo 2.1. QSAR teknikleri ve kullanılan parametreler. 17
Tablo 3.1. Penisilin Serisine Ait Temel İskelet, Sübstitüentler, Deneysel Aktivite Değerleri ve Konformer Sayıları……..……....36
Tablo 3.2. Penisilin serisinde 73. bileşiğin 15 konformeri için bağıl enerji ve boltzmann değerleri. 40
Tablo 4.1. a) 73. bileşiğin (referans) ilk konformerinde farmakofor gruba ait alt matris görülmektedir, b) 88 Bileşiğin en düşük enerjili konformerleri için minimum tolerans değerleri, c) 88 Bileşiğin 1222 konformeri için minimum tolerans değerleri…….75
Tablo 4.2. Penisilin serisi için hesaplamalarda kullanılan parametreler ve 73 bileşiğinin en düşük enerjili konformerine ait sayısal değerler..... 77
Tablo 4.3. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için elde edilen 1-7 κj değerlerine göre hesaplanan aktivite değerleri..... 93
Tablo 4.4. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için elde edilen 8-13κj değerlerine göre hesaplanan aktivite değerleri..... 95
Tablo 4.5. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için hesaplamalardan elde edilen parametreler ve κj değerleri..... 98
Tablo 4.6. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için leave one out hesaplama sonuçları..... 100
Tablo 4.7. Penisilin serisi 60 bileşik eğitim, 28 bileşik test için elde edilen 1–7 κj değerlerine göre hesaplanan logaritma cinsinden aktivite değerleri. 100
Tablo 4.8. Penisilin serisi 60 bileşik eğitim, 28 bileşik test için elde edilen 8–13 κj değerlerine göre hesaplanan logaritma cinsinden aktivite değerleri. 103
Tablo 4.9. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için aktivite değeri logaritmik olan hesaplamalardan elde edilen parametreler ve κj değerleri. 105
Tablo 4.10. Penislin serisi 60 bileşik eğitim, 28 bileşik test için aktivite değeri logaritma cinsinden olan hesaplamalarda leave one out hesaplama sonuçları. 107
Tablo 4.11. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için elde edilen 1–8 κj değerlerine göre hesaplanan aktivite değerleri..... 108
Tablo 4.12. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için elde edilen 9–14 κj değerlerine göre hesaplanan aktivite değerleri..... 110

Tablo 4.13. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için aktivite hesaplamalarında elde edilen parametre numaraları ve κj değerleri. 113
Tablo 4.14. Penisilin serisi 74 bileşik eğitim, 13 bileşik test seti için leave one out hesaplama sonuçları. 115
Tablo 5.1. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için 12 parametreye göre deneysel ve hesaplanan aktivite değerleri.……...119
Tablo 5.2. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için en iyi 12 moleküler parametre, 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri ve Kappa (kj) değerileri. 121
Tablo 5.3. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için aktivite değeri logaritma cinsinden olan hesaplamalarda 12 parametreye göre deneysel ve hesaplanan aktivite değerleri. 123
Tablo 5.4. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için en iyi 12 moleküler parametre, 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri ve Kappa (kj) değerileri. 126
Tablo 5.5. Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti için 14 parametreye göre deneysel ve hesaplanan aktivite değerleri. 128
Tablo 5.6. Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti için en iyi 14 moleküler parametre, 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri ve Kappa (kj) değerileri. 130
Tablo 5.7. Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti için hesaplama sonuçlarının makaledeki hesaplama sonuçları ile karşılaştırılması. 133

ŞEKİLLER LİSTESİ

Şekil 1.1. Bir bakteri hücresinde peptidoglikan (hücre duvarı) biyosentezinin ve çeşitli antibiyotiklerin etki yerlerinin şematik gösterimi 7
Şekil 3.1. İlaç tasarımında EMRE yazılım sisteminin basamakları..... 31
Şekil 3.2. EC-GA metodunda takip edilen aşamalar 34
Şekil 3.3. Penisilin (Beta laktam) molekülüne ait temel iskelet..... 35
Şekil 3.4. Penisilin serisinde 73 nolu bileşiğine ait 15 konformerin üst üste çakıştırılması 42
Şekil 3.5. Penisilin serisinde 1 nolu bileşiğine ait 14 konformerin üst üste çakıştırılması 42
Şekil 3.6. Üç boyutlu ECMC'nin gösterimi..... 43
Şekil 3.7. EMRE programına ait girdiler sekmesi..... 44
Şekil 3.8. EMRE programına ait matris hazırlama sekmesi 44
Şekil 3.9. Penisilin serisi için en basit bileşiğe göre diğer bileşiklerin fonksiyonel grupları ve atomlarının düzenlenmesi yani moleküllerde birbirine karşılık gelen atomlar..... 48
Şekil 3.10. ECSP programına ait arayüz..... 48
Şekil 3.11. Emre programına ait parametrik sekmesi..... 50
Şekil 3.12. Optimum tasarım prosesi..... 56
Şekil 3.13. Genetik algoritmanın aşamaları 66
Şekil 4.1. Penisilin serisinde 73. bileşiğinin 1. konformerine (73_1) ait ECMC……70
Şekil 4.2. Penisilin serisinde 73. bileşiğinin 15. konformerine (73_15) ait ECMC 71
Şekil 4.3. 73. bileşik için üst gruba ait matris ve bileşiğin en düşük enerjili konformerinin şekli 74
Şekil 4.4. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği.....99
Şekil 4.5. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği.....106
Şekil 4.6. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği.....114
Şekil 5.1. Penisilin serisi için Referans bileşiğin ECSA, farmakofor grubu temsil eden atomlar, 1222 Konformer için tolerans değeri 118

Şekil 5.2. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin eğitim seti için çizilmiş R^2 grafiği 122
Şekil 5.3. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin test seti için çizilmiş R^2 grafiği 123
Şekil 5.4. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin eğitim seti için çizilmiş R2 grafiği 127
Şekil 5.5. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin test seti için çizilmiş R2 grafiği 127
Şekil 5.6. Penisilin serisinde 14 parametre ile hesaplanan teorik aktivitenin eğitim seti için çizilmiş R^2 grafiği 131
Şekil 5.7. Penisilin serisinde 14 parametre ile hesaplanan teorik aktivitenin test seti için çizilmiş R^2 grafiği 132

KISALTMALAR VE SİMGELER

QSAR : Kantitatif Yapı Aktivite İlişkisi

SAR :Yapı Aktivite İlişkisi

3D-QSAR :Üç Boyutlu Kantitatif Yapı Aktivite İlişkisi

2D-QSAR :İki Boyutlu Kantitatif Yapı Aktivite İlişkisi

HQSAR :Hologram Kantitatif Yapı Aktivite İlişkisi

Pha :Farmakofor

APS :Aktiviteyi perdeleyici grup

AG :Aktiviteyi artırıcı yardımcı grup

CADD :Bilgisayar destekli ilaç tasarımı

HTS :Çok Sayıdaki Bileşiği Ayırma Yöntemi (High Throughput Screening)

CoMFA :Karşılaştırmalı Moleküler Alan Analizi

MSA :Moleküler Şekil Analizi

CoMSIA :Karşılaştırmalı Moleküler Şekil Analizi

CoMMA :Karşılaştırmalı Moleküler Moment Analizi

ECMC :Elektron Konformasyonel Uygunluk Matris

ECSA :Elektron Konformasyonel Alt Matris

PAB : İlaçın Yüzde Bağlanması

$\mathbf{PAB_{deneneysel}}$: İlaçın Yüzde Bağlanmasının Deneysel Değeri

$\mathbf{PAB_{hesaplanan}}$: İlaçın Yüzde Bağlanmasının Hesaplanan Değeri

Psa : Parametrenin sayısal değeri

E : E istatistik değeri

$\mathbf{R^2_{eğitim}}$: Eğitim seti için regrasyon katsayı değeri

$\mathbf{R^2_{test}}$: Test seti için regrasyon katsayı değeri

$\mathbf{q^2}$: Dahili çapraz doğrulama değeri

$\mathbf{q^2_{ext1}}$: Harici çapraz doğrulama değeri (Eğitim ortalama)

$\mathbf{q^2_{ext2}}$: Harici çapraz doğrulama değeri (Test ortalama)

1. BÖLÜM

GİRİŞ

1.1. Giriş

Bakteriyel hastalıkların kontrolünde antibakteriyel ilaçların geliştirilmesi çok önemli adımlardan biridir. Bakterilerin çoğalmasını engelleyen ya da bakterileri öldüren biyolojik kaynaklı ve sentetik olarak elde edilen maddelere antibiyotikler denir. Antibiyotikler denince ilk akla gelenler hiç kuşkusuz Beta laktam grubu antibiyotiklerdir. Penisilinler Beta laktam grubu antibiyotikler içerisinde yer alır.
Penisilin türevi antibiyotikler günümüzde en sık kullanılan antibiyotik türevlerinin başında gelir. Penisilin türevi antibiyotiklerin incelenmesi [1] ve antibakteriyel ilaçların geliştirilmesi öncelikli hedefimizdir.

İdeal ilaç geliştirmedeki amaç daha etkili, daha az toksik ve daha az yan etkili olan ilaçlar elde etmektir. Yeni ilaç geliştirilmesi, uzun, zahmetli ve oldukça yüksek harcamaları gerektiren bir süreçtir. Son zamanlarda yapılan bir hesaplamaya göre, yeni ilaç geliştirilmesi yaklaşık 12–15 yılı ve 1 milyar USD harcamayı gerektirmektedir. Sürecin bu kadar uzun olması, yeni ilaç geliştirilmesinin zorluğu hakkında ipucu vermektedir.

Deneysel çalışmalar yapmadan bilgisayar hesaplamalarından yararlanarak ilaçların yapıları hakkında önbilgiye sahip olunabilmektedir. Elde edilen bu ön bilgiler kullanılarak ilaçta olması gereken özellikler belirlenerek uygun bileşiklerin sentezi gerçekleştirilebilir. Bilgisayar-Destekli ilaç tasarımı, (Computer-Assisted Drug Design, CADD) ilaç geliştirme çalışmaları için ortaya çıkan bir tekniktir [2] ve ilaç olarak kullanılan, biyolojik olarak aktif bileşiklerin optimizasyonu, tasarımı ve araştırılması aşamalarını içeren bilgisayar destekli tekniklerin, QSAR da dâhil, çok geniş bir

uygulama alanına sahiptir. Bilgisayar-Destekli ilaç tasarım yöntemleri ve programları, deneysel yöntemler kullanılarak bir ilacın keşfinden ticari olarak satılmasına kadar geçen sürenin ve maliyetinin düşürülmesi amacıyla, son yıllarda hem ilaç firmaları hem de üniversitelerde yaygın bir şekilde kullanılmaktadır. Bu nedenle pek çok uygulama alanı bulmuştur.

Başlangıçta QSAR metodu bileşiklerin fizikokimyasal özelliklerine göre aktiviteyi açıklamıştır. Fakat son zamanlarda biyomoleküller üç boyutlu olarak ele alınıp elektronik özellikler ile aktivite değerleri bulunmuş ve bu çalışmalar 3D-QSAR olarak adlandırılmıştır. QSAR yaklaşımlarındaki tüm gelişmeler hızlı bilgisayarların kullanımı ile olmuştur.

Hesaplamalı metotlar ile yeni moleküllerin dizaynı; ilaçların biyolojik aktivitelerinin QSAR teknikleri ile incelenmesi, veri analizlerinin matematiksel ve istatistiksel yöntemleri; teorik ve hesaplamalı bilgisayar metotlarının uygulamaları yapılır.

Türkiye'de QSAR çalışmaları yapılmasına rağmen, henüz bu konuda program geliştirilmemiş olması, bu konuda çalışma yapan araştırma gruplarını paket programları yurtdışından yüksek ücretlerle temin etmek zorunda bırakmaktadır (örneğin piyasada çok kullanılan bir program olan sybyl'ın bir yıllık kira bedeli yaklaşık 25.000 Euro). Algoritmasını ortaya koyduğumuz ve büyük oranda yazılımlarını başardığımız QSAR ile ilgili programların (EMRE ve ECSP) piyasada bulunan ve yüksek maliyetli ticari programlara alternatif olacağını düşünmekteyiz. Hazırlanan yazılımlarla; deneysel aktivitesi bilinen bileşiklerin teorik aktivitelerinin öngörülmesi, aktivitesi deneysel olarak bilinmeyenlerin aktivitelerinin öngörülmesi ve bilinmeyen bileşiklerin tasarlanması amaçlanmaktadır.

1.2. İlaç

İlaç, canlı hücre üzerinde meydana getirdiği tesir ile bir hastalığın teşhisini, iyileştirilmesi veya semptomlarının azaltılması amacıyla tedavisini veya bu hastalıktan korunmayı mümkün kılan, canlılara değişik uygulama yöntemleri ile verilen doğal, yarı sentetik veya sentetik kimyasal preparatlardır. İlaçlar, ağrıları ya da başka rahatsız edici durumları dindirmek, zihin ve bedenin normal dışı durumlarını düzeltmek ve denetim altında tutmak için kullanılırlar. İlaçlar vücudun belirli bölgelerine etki ederek biyolojik süreci etkileyen moleküllerdir [3]. İlaçlar çok çeşitlidir ve vücudu değişik yollardan

etkiler. Örneğin vitamin türünden ilaçlar vücudun iyi çalışması için gerekli kimyasal maddeleri sağlar.

İlaçlar genellikle aşağıda verilen üç aşamada çalışır:

- Bakteri, parazit ya da virüslerin biyolojik fonksiyonlarına müdahale eder. (Antibiyotikler ve HIV ilaçları)
- Bakteri, parazit ya da virüslerin sistemlerimiz ile etkileşimini engellemek
- Biyolojik fonksiyonlarımızın artmasına müdahale eder. (Aşılar, insülin, kanser ilaçları)

Bir ilacı bağlayıp farmakolojik etkilerine aracılık eden, hücre yüzeyi veya içinde bulunan hedef makromoleküle reseptör denir. Reseptör kavramının temelleri Ehrlich ve Langley'in 19. yy'nin sonları ve 20. yy'nin başlarında yaptığı deneysel çalışmalara dayanır. Sayıca değerlendirildiğinde en önemli ilaç reseptör sınıfı proteinlerden oluşur. Hormon reseptörleri, büyüme faktörlerinin reseptörleri, metabolik veya düzenleyici yolaklarda rol oynayan enzimler (dihidrofolat redüktaz, asetilkolinesteraz gibi), transport işlevi olan proteinler (Na+K+-ATPaz) veya yapısal görevi olan proteinler (tübülin gibi) önemli reseptör proteinlerdir. Nükleik asitler özellikle kanser kemoterapisinde önemli ilaç reseptörleridir. Birçok ilaç fizyolojik önemi olan bu tip reseptörler üzerinde etki gösterir. Agonistler organizmadaki fizyolojik reseptörlere bağlanarak endojen düzenleyici bileşiklerin etkilerini taklit ederler. Antagonistler ise ilgili reseptöre bağlanmakla birlikte intrinsik (iç) düzenleyici aktiviteleri olmadığı için sadece agonistin bağlanmasını engelleyerek etki gösterirler [3].

1.3. Bilgisayar Destekli İlaç Tasarımı

- 60 –70ler… QSAR
- 80ler…Moleküler Modelleme
- 90 larınbaşı… 3D (3 Boyutlu) Tarama, Hedef Tabanlı ilaç Tasarımı (de Novo tasarımı)
- 90larınsonu… Kombinatoryal Kimya, HTS (HighThrougputScreening)
- 00ler… Tüm Yöntemlerin bir arada kullanımı

Bilgisayar-Destekli ilaç tasarımı, (Computer-Assisted Drug Design, CADD) ilaç geliştirme çalışmaları için ortaya çıkan bir tekniktir [2]. İlaç olarak kullanılan biyolojik olarak aktif bileşiklerin optimizasyonu, tasarımı ve araştırılması aşamalarını içeren

bilgisayar destekli tekniklerin tümünü kapsar. Bilgisayar-Destekli ilaç tasarım yöntemleri, deneysel yöntemler kullanılarak bir ilacın keşfinden ticari olarak satılmasına kadar geçen sürenin ve maliyetinin düşürülmesi amacıyla son yıllarda hem ilaç firmaları hem de üniversitelerde yaygın bir şekilde kullanılmaktadır. Bu nedenle pek çok uygulama alanı bulmuştur.

Bilgisayarlar, ilaç tasarımı yöntemlerinin önemli bir kısmını oluşturmaktadır. Yapı analizleri, yapı karşılaştırmaları, öncü bileşik tasarımı, aktif konformasyon ve farmakofor (Pharmacophore, Pha) belirlenmesi, protein ve bağlanma alan yapısı, ligand bağlama ve QSAR (Quantitative Structure Activity Relationship)'ıda içeren çok geniş bir uygulama alanına sahiptir. Son 20 yıl boyunca bu alandaki çalışmalar "yapı-aktivite ilişkisi" (Structure-Activity Relationship, SAR) ve "nicel yapı-aktivite ilişkisi" (Quantitative SAR, QSAR) başlığı altında yayınlanmıştır. Başlangıçta QSAR metodu aktiviteyi bileşiklerin fizikokimyasal özelliklerine göre açıklamıştır. Fakat son zamanlarda biyomoleküller üç boyutlu olarak ele alınıp elektronik özellikleriyle aktivite değerleri bulunmuş ve bu çalışmalar 3D-QSAR olarak adlandırılmıştır. QSAR yaklaşımlarındaki tüm gelişmeler hızlı bilgisayarların kullanımı ile olmuştur [4–6].

İlk bilgisayar destekli, yapı temelli ilaç tasarım metotları 1980'lerin başında ortaya çıkmıştır. Bunun için ilk olarak Hammet eşitliği kullanılmış olup zaman içerisinde gelişim göstermiştir. Günümüzde ise bu gelişim halen devam etmektedir [7].

Bilgisayar-destekli ilaç tasarımında, ligand-temelli (ligand-based etki gösteren moleküllerin yapısından reseptör yapısının yorumlanması) ve yapı-temelli (structure-based bilinen reseptör yapısından hareketle etki gösterebilecek moleküllerin tasarlanması) olmak üzere 2 temel yöntem bulunmaktadır. Eğer ilacın vücut içinde etkileştiği hedef protein yapısı bilinmiyorsa, QSAR (quantitative-structure activity relationship) gibi yöntemler kullanılırken, "Hedef protein" yapısı biliniyorsa Moleküler dinamik(MD), simulasyon (zamana karşı konformasyon analizi) ve Docking (protein-ligand kilitlenmesi) yöntemleri kullanılmaktadır.

Moleküler Modelleme, Kalıp Tanıma (pattern recognition), Kümeleme analizi (Cluster Analysis), Reseptör Uyum veya Reseptör Tanımlama Yöntemleri bilgisayar destekli ilaç tasarımı ile ilaç özelliği gösteren yeni moleküllere ulaşmak üzere kullanılan önemli yöntemlerdir [8].

1.4. Hesaplamalı Kimya

Hesaplamalı kimya teorik kimyacılar tarafından geliştirilmiş matematiksel yöntemleri uygular ve elde edilen sonuçları yorumlar, böylece deneysel kimya ile teorik kimya arasında bir köprü kurar. Hesaplamalı kimya ile sadece kararlı molekülleri değil, aynı zamanda kısa-ömürlü, kararsız araürünler ve geçiş hallerini de çalışmak mümkün olur. Bu şekilde, gözlem yolu ile elde edilmesi mümkün olmayan moleküller ve tepkimeler hakkında bilgi sahibi olunabilir. Bu hesaplamalar ile elde edilen nitel veya nicel sonuçlar, kimyacıların çok faydalı öngörülerde bulunmasını sağlar.

Deneysel çalışmaları desteklemek ya da deneysel çalışma yapmadan elde edilecek sonuçları önceden tahmin edebilmek amacıyla hesaplamalı yöntemleri kullanacak olan araştırmacılar için üç farklı seçenek vardır. Moleküler mekanik yöntemi (MM), yarı-denel (semiempirik) yöntem ve ab initio yöntemidir [9].

1.5. Antibiyotikler

Antibiyotikler, bakteri veya mantarlar tarafından sentez edilen ve diğer mikroorganizmaların gelişimini engelleyen ve sonuçta onların ölümüne yol açan maddelerdir. Ancak genel kullanımda antibiyotik teriminin içine mikroorganizma ürünü olmayan (sülfonamidler, kinolonlor gibi) sentetik antimikrobik ajanlar da eklenmiştir. Bu güne kadar yüzlerce antibiyotik tanımlanmış olup bunların önemli bir bölümü geliştirilerek enfeksiyon hastalıklarının tedavisinde kullanılmaktadır. Bakteriyel, fungal ve viral replikasyon mekanizmalarının moleküler düzeyde aydınlatılması ile bu mikroorganizmaların yaşam döngülerini bozabilecek yeni bileşiklerin gelişimi hızlanmıştır [10].

1.5.1. Antibiyotiklerin sınıflandırılması ve etki mekanizması

Antibiyotikler kendi aralarında antibakteriyel etki spektrumu, etki mekanizması, fiziksel ve kimyasal özellikler bakımından farklılıklar gösterirler. Antimikrobik ilaçların sınıflandırması için çeşitli öneriler olmakla birlikte, aralarında farklı veya benzer özelliklerin bulunması nedeniyle başarılı bir sınıflandırma yapılamamaktadır. Geleneksel ve en yaygın kullanılan sınıflandırma kimyasal yapı ve önerilen etki mekanizmasına göre yapılmıştır. Antibiyotikler etki mekanizmalarına göre Sitoplâzma

membran geçirgenliğini engelleyenler, Protein sentezini engelleyenler, Nükleik asitlere etki edenler, Antimetabolik etki gösterenler, **Hücre duvarı sentezini engelleyenler** şekilde özetlenebilir [10].

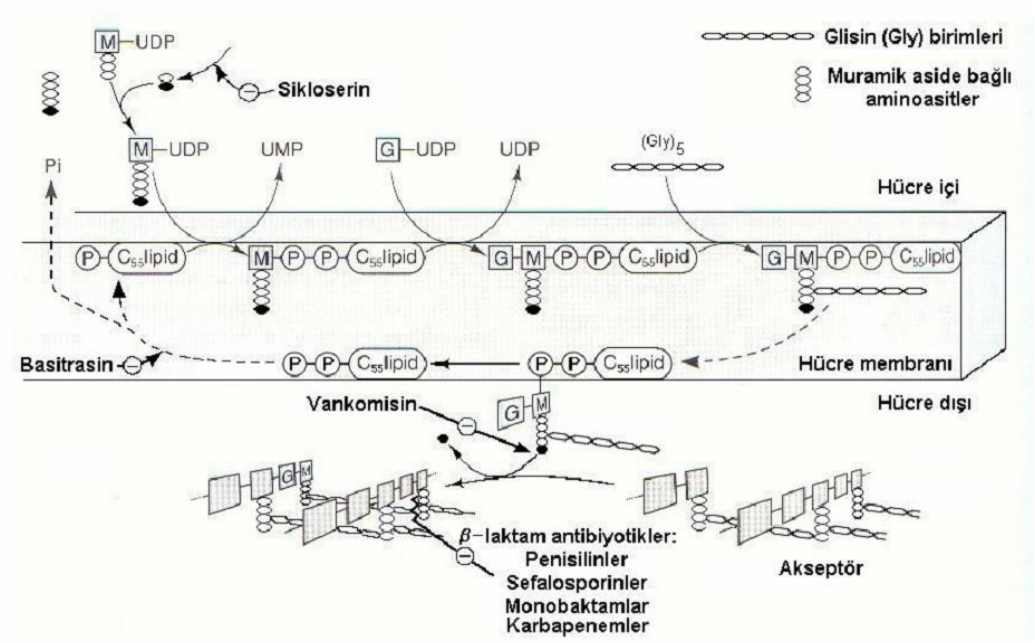
Hücre duvarı sentezini engelleyen antibiyotikler, kimyasal yapıları içinde bir beta laktam halkası içerdiklerinden "beta laktamlar" olarak adlandırılırlar. Beta laktam halkası biri azot, üçü karbon olan 4 üyeli doymuş bir halkadır. Hücre duvarındaki belli proteinlere bağlanarak, hücre duvarının devamlılığı için süregelen sentezi durdururlar. Bunun sonucunda bakteri şeklini kaybederek ölür. Bu yüzden beta laktamlar, hücre duvarı olmayan bakterilere etkisizdir. Beta laktam antibiyotikler; penisilinler, sefalosporinler, karbapenemler, monobaktamlar, beta laktamaz inhibitörleri (klavamlar) olmak üzere 5 gruba ayrılırlar.

1.5.1.1. Penisilinler

Penisilinler beta laktam grubu antibiyotikler içerisinde yer alır. İlk kez 1928'de Fleming tarafından bir mantar olan penicillium notatum'un salgıladığı antibakteriyal bir maddenin Stafilokok kültürlerinde lizise (organik yapıların parçalanması, yok olması.) neden olduğu fark edilmiş ve bu antibakteriyel maddeye penisilin demiştir. Daha sonra 1940'lı yıllarda Florey ve arkadaşları tarafından yeterli düzeyde saflaştırılmış ve klinik kullanımıyla ilgili çalışmalar başlatılmıştır. O tarihten bugüne kadar birçok penisilin türevli bileşikler geliştirilerek bakteriyal patojenlerin neden olduğu durumlarda sıklıkla kullanılabilir duruma gelinmiştir. Bütün penisilinler temel yapıyı 6-Aminopenisillonikasit yani "APA" çekirdeği oluşturur. 6-APA bir beta laktam halkası ile buna bağlı bir tiazolidin halkasından ibarettir. 6-APA çekirdeğine değişik yan zincirlerin eklenmesiyle önemli farmokolojik ve antibakteriyal farklılık gösteren birçok penisilin türevi elde edilmiştir. Penisilinlerin etki spektrumlarına göre temel penisilinler (benzil penicillinler), penisilinoza dirençli penisillinler, geniş spektrumlu nonantipseudomonal penisilinler, antipseudomonal penisilinler olarak 4 gruba ayrılır.

Penisilinler bakteriyel hücre duvarı sentezinin son basamağını (transpeptidasyon ya da çapraz-bağlanma) inhibe ederek ozmotik olarak dayanıksız bir membranın oluşumuna neden olurlar. Membranın ozmotik bakımdan dayanıksız olması hücrenin parçalanmasına ve dolayısı ile bakterinin ölümüne neden olur. Bu nedenle hücre duvarı sentez inhibitörleri bakterisid etkilidir. Penisilinlerin antimikrobiyal etkinlikleri

moleküler büyüklük, elektriksel yük ve hidrofobisiteleri ile ilişkilidir. Penisilinler sadece peptidoglikan hücre duvarı sentezleyen ve hızlı büyüme gösteren mikroorganizmalara karşı etkilidirler. Buna bağlı olarak hücre duvarı olmayan mikobakteriler, protozoalar, funguslar ve virüslere karşı etkisizdirler. Penisilinlerin bakteri hücre duvarı sentezini inhibe etme mekanizmaları Penisilin bağlayıcı proteinler (PBP), Transpeptidazın inhibisyonu ve Otolizinler (mürein hidrolazlar) [10] olmak üzere üç grupta incelenirler. Aşağıda şekil 1.1'de Bir bakteri hücresinde peptidoglikan (hücre duvarı) biyosentezinin ve çeşitli antibiyotiklerin etki yerlerinin şematik gösterimi verilmiştir.



Şekil 1.1. Bir bakteri hücresinde peptidoglikan (hücre duvarı) biyosentezinin ve çeşitli antibiyotiklerin etki yerlerinin şematik gösterimi.

Bir bakteri hücresinde peptidoglikan (hücre duvarı) biyosentezinin ve çeşitli antibiyotiklerin etki yerlerinin şematik gösterimi: Hidrofilik özellikteki disakkarit-pentapeptit pirofosfat köprüleri (-P-P-) ile büyük bir lipide (C55 lipid) bağlı olarak hücre membranından transfer edilir. Hücre dışında büyümekte olan peptidoglikan tabakasına (akseptör) enzimatik olarak bağlanır. Son mekanizmayı oluşturan transpeptidasyon reaksiyonunda Gly5'in serbest ucu akseptörde bulunan bir M'nin peptid yan zincirine bağlanırken terminal amino asit (alanin) kopar. Tekrar taşıyıcı olarak işlev görmeden önce bir fosfat grubunun (Pi) kopması ile lipid oluşturulur. (M= N-asetil muramik asit; G= N-asetil glukozamin)

2. BÖLÜM

QSAR METOTLARI

2.1. QSAR'a Genel Bakış

İlaç tasarımının temel özelliği; sadece istenilen etkiyi meydana getirme gücü ve yararlılık açısından değil aynı zamanda farmakokinetik özellikler ve yan etkiler açısından da sürekli olarak daha iyi analogları elde etmek amacıyla öncü bileşiklerin analoglarını sentez etmek ve onların biyolojik aktivitesini test etmektir. Bu konuda kabul edilen temel varsayım şöyledir: kimyasal yapıdaki herhangi bir değişiklik biyoaktivitede pozitif ya da negatif bir değişiklik meydana getirir. Böyle bir sistematik sebep-sonuç ilişkisine "yapı aktivite ilişkisi, structure-activity relationship (SAR)" denir. SAR'ın amacı; değişen ilaç yapısının kimyasal sonuçlarını belirlemek ve daha sonra kimyasal yapıdaki ve özelliklerdeki hangi değişikliklerin daha iyi biyolojik aktivite sağlayacağını saptamaktır. QSAR modelleri ile biyolojik aktiviteleri bilinen benzer bileşiklerin kimyasal yapılarından yola çıkarak yeni ya da test edilmemiş kimyasalların biyolojik aktiviteleri (ya da özellik, reaktivite gibi) belirlenebilir.

Araştırmacılar SAR çalışmalarına 19. yy'da başlamışlardır. Moleküllerin fizikokimyasal özelliklerini biyolojik aktivite ile nicel olarak ilişkilendiren ve "nicel yapı-aktivite ilişkisi, quantitative structure-activity relationship (QSAR)" olarak bilinen bu kavram ilk olarak 1960'ların başında Corwin Hansch [11] tarafından ortaya konuldu. O zamandan beri gelişen birçok yeni metot, ilaç araştırma alanında büyük değişiklikler meydana getirdi. Bu alandaki farklı bakış açılarını ortaya koyan birçok çalışma yayınlanmıştır [12–18].

Bu kavram zaman içerisinde gelişerek 2D-QSAR'dan 3D-QSAR'a ve son zamanlarda da 4D-QSAR'a [19] kadar gelişmiştir.

2.2. QSAR'ın Temel İlkeleri

Nicel yapı aktivite ilişkisinin (Quantitative Structure Activity Relationship) ilk genel formülü, Crum-Brown ve Fraser tarafından 1868'de ortaya koyulmuştur. Farklı alkoloidlerle ilgili çalışmalarında temel azot atomlarının alkilasyonunun dikkate değer şekilde biyolojik etkinlik meydana getirdiğini gördüler. Böylece biyolojik aktivitenin (Φ), kimyasal yapının ("constitution") bir fonksiyonu olduğunu varsaydılar:

$$\Phi = f(C) \tag{2.1}$$

Richet [20] bileşiklerin toksisitesinin sudaki çözünürlüklerinin tersi bir yol takip ettiğini keşfetti. Böyle bir ilişki eşitlik 2.1'e karşılık gelmektedir. Burada ΔΦ, biyolojik aktiviteler arasındaki farktır. Bu fark kimyasal ve özellikle de fizikokimyasal özelliklerdeki değişikliklerden ΔC kaynaklanmaktadır. ΔC, kimyasal yapıdaki değişiklikleri ifade etmektedir.

$$\Delta\Phi = f(\Delta C) \tag{2.1}$$

Bazı QSAR eşitlikleri eşitlik 2.1'e uyar. Çünkü biyolojik aktiviteler arasındaki fark, sadece lipofilitedeki ve/veya diğer fizikokimyasal özelliklerdeki değişiklikler ile ilişkilendirilebilir.

Meyer [21] ve Overton [22] birbirinden bağımsız olarak yaptıkları çalışmalarında yağ-su dağılma katsayısı olarak ifade edilen lipofilite ile narkotik aktiviteler arasında doğrusal ilişki olduğunu gözlemlediler. Fühner ve Neubauer [23] homolog bir seri içerisinde biyolojik aktivitenin geometrik olarak arttığını (yani $1:3:3^2:3^3$ gibi) fark ettiler. Biyolojik aktiviteyi arttırıcı grup katkısının bir kanıtı niteliğindeki bu sonuç daha sonraları diğer araştırmacılar tarafından da doğrulanmıştır.

QSAR metodu 1950'lerin ortalarından itibaren hızla gelişmeye başlamıştır: Bruice, Kharasch, and Winzler [24] ilk Free-Wilson tipi analizi olarak ele alınabilecek çalışmalarında tiroid hormonunun bir serisi için grup katkılarının biyolojik aktivite değerlerine etkisini formülüze ettiler.

Zahradnik [25–27] organik bileşiklerin reaktivitesini nicel olarak tanımlamak için kullanılan Hammett eşitliği (eşitlik 2.3) [28] kavramını biyolojik aktivite verilerine eşitlik 2.4'de uygulamaya çalıştı.

$$\log k_{R\text{-}X} - \log k_{R\text{-}H} = Q\sigma \tag{2.3}$$

$$\log \tau_i - \log \tau_{Et} = \alpha\beta \qquad (2.4)$$

Bu biyolojik Hammett eşitliğindeki τ_i terimi bir serideki i. üyenin aktivite değerini gösterir. τ_{Et} ise aynı serideki etil bileşiğinin biyolojik aktivite değeridir. β, sübtitüent sabiti (Hammett eşitliğindeki reaksiyon sabitine Q karşılık gelir). α terimi de biyolojik sistemi tanımlayan bir sabit olup Hammett eşitliğindeki σ elektronik parametresine karşılık gelir. Ne yazık ki eşitlik 2.4 sadece belirli olmayan biyolojik aktiviteler için geçerlidir (Çoğunlukla homolog serilerde ve belirli bir lipofilite değer aralığında).

1962'de Hansen, sübstitüe benzoik asitlerin toksisitesi ve sübstitüentlerin σ elektronik sabitleri arasında ilk gerçek Hammett tipi ilişkiyi türetti. Aynı yıl Corwin Hansch tarafından ilk QSAR çalışması "The correlation of the biological activity of phenoxyacetic acids with Hammett substituent constants and partition coefficients" [29] yayınlandı.

Modern QSAR teknolojisi 1964'de Hansch – Fujita [30] ve Free – Wilson'nın [31] birbirinden bağımsız olarak ortaya çıkardıkları yayınları ile başlamıştır. Bu iki çalışma iki yeni QSAR metodunun gelişimine yol açtı. Bu çalışmalar daha sonra sırasıyla Hansch ve Free – Wilson analizi olarak anılmıştır.

Hansch analizi olarak adlandırılan ve yaygın olarak kullanılan QSAR metodu aynı zamanda "ekstra termodinamik ya da lineer serbest-enerji ilişkisi yaklaşımı" olarak da bilinir. Biyolojik etkinin kimyasal bileşiğin fizikokimyasal özelliklerinin bir fonksiyonu olduğunun tanımlanması, nicel yapı-aktivite ilişkisi (QSAR) analizlerinde Corvin Hansch'ın ekstratermodinamik yaklaşımının temelini oluşturmaktadır. Böylece homolog bir dizin içerisinde moleküllerin, elektronik, sterik, hidrofobik ve yapısal gibi değişik özelliklerinin biyolojik etki üzerine gösterdikleri olası katkılarının kantitatif değerleri bulunabilmektedir. Hansch analizi lineer, multilineer ya da bilineer regresyon analizi yöntemleri kullanarak biyolojik aktivite değerlerini fizikokimyasal özelliklerle ilişkilendirir. Buna göre Hansch analizi gerçekten bir özellik-özellik ilişkisi modelidir. 1962'de Hansch ve arkadaşları bitki gelişimini düzenleyicilerinin yapı aktitivite ilişkisi ve bunların Hammett sabiti ve hidrofobisiteye bağımlılığına dair çalışmalarını yayınladılar [11].

İkinci çalışma, Hansch analizini destekleyen ve belirli yapısal değişiklikler için oldukça kullanışlı olan Free – Wilson analizi olarak bilinmektedir. Free – Wilson analizi,

QSAR'ın ilk yıllarında geliştirilen ve indikatör değişkenlerin kullanımı ile ilgili bir tekniktir. "Matematiksel model", "katkı modeli" ya da "de novo yaklaşımı" Free – Wilson metoduyla eş anlamlı olarak kullanılır. Rasgele seçilen bir referans bileşikten ayrılan her yapısal özellik için bir indikatör değişken oluşturulur [32, 33].

Her iki çalışma da matematiksel metot (eşitlik 2.1'deki f) olarak çoklu regresyon analizini kullanır fakat kimyasal özelliklerin tanımlanması açısından farklılık gösterirler. Free – Wilson analizi bileşiklerin iki boyutlu yapısından doğrudan türetilen kimyasal fragmentlere dayanırken, Hansch analizinde sübstitüent sabitleri ve diğer fizikokimyasal tanımlayıcılar kullanılır.

Günümüzde eşitlik 2.1'deki f terimini açıklamak için çok sayıda matematiksel metot mevcuttur. Bunlardan en sık kullanılanları: çoklu regresyon analizi (multiple regression analysis), temel bileşen ve faktör analizi (principal component and factor analysis), temel bileşen regresyon analizi (principal component regression analysis), kısmi en küçük kareler yöntemi (partial least square, PLS), diskriminant analizi ve diğer sınıflandırma metotları ve yapay sinir ağları. Kimyasal yapıyı tanımlayan çok sayıda kimyasal tanımlayıcı ile çeşitli matematiksel metotlar birleştirilmiştir. Bu tanımlayıcılar için ansiklopedik yol gösterici niteliğindeki bir çalışma Tedeschini ve Consonni tarafından Handbook of Molecular Descriptors [34] adı altında sunulmuştur. Bu tanımlayıcıların tamamının kullanışlı olduğu kanıtlanmış değildir. Genel olarak bahsedilirse, deneysel nicelikler olarak sınıflandırılabilirler: log P, pKa (bu niceliklerde hesaplanabilirler), spektroskopik veriler; substituent sabitleri (elektronik, hidrofobik ve sterik); moleküler modelleme ve kuantum kimyasal hesaplamalardan elde edilen parametreler gibi.

Tipik biyolojik aktivite ölçümleri (eşitlik 2.1'deki C terimi), belirli bir etki üreten bir bileşiğin, bağlanma, birleşme, inhibisyon sabitleri ya da oran sabitlerine karşı çizilen doz-cevap eğrisinden (örn. ED_{50} ya da ID_{50}) elde edilen molar konsantrasyon (C) değerleridir. Daha aktif bileşikler için daha geniş aralıkta değerler elde etmek amacıyla, ayrışma sabitleri ve molar konsantrasyon temelli niceliklerin genellikle çarpmaya göre tersi alınır. Termodinamik ya da kinetik sebeplere dayanarak, bu tür parametreler logaritmik dönüşüm yapılarak serbest-enerji ile ilgili niceliklere dönüştürülebilir. Böylece, eşitlik 2.1'deki Φ terimi için ifadeler $pC = -\log C = \log 1/C$ (örn. pED_{50} ya da pIC_{50}), log K (K; bağlanma, inhibisyon veya oran sabiti) ve $\log 1 / K_d$ (K_d; ayrışma

sabiti) şeklini alır. Genel olarak biyolojik ölçümün logaritmik dönüşümü sadece Hansch analizinde (veya diğer lineer serbest ilişkiye dayanan yöntemler) değil aynı zamanda nicel biyolojik ölçümlere uygulanan tüm QSAR yaklaşımlarında da kullanılır. Logaritmik dönüşümün kullanılma sebeplerinden birisi sonuçların daha iyi karşılaştırılabilmesidir. Bazen de sonuçlar sadece tek bir dozda ölçülen %etki olarak ifade edilir. Aslında bu tür veriler Hansch tipi ve ilgili diğer QSAR yaklaşımları için uygun değildir. Deneyimler gösteriyor ki yine de bu tür veriler birkaç yüzdeden %100'e kadar tüm yüzde aralığı mevcut olmak şartıyla logaritmik dönüşümden sonra anlamlı nicel yapı aktivite ilişkisi vermektedir. Böyle değerler için iyi bir alternatif logaritmik dönüşümüdür:

$$\Phi = \log (\%etki(100 - \%etki)) \qquad (2.5)$$

Başka bir alternatif ise %etki değerlerini, sınıflandırma metotlarıyla daha sonra analiz etmek üzere bir sınıflandırma şemasına dönüştürmektir. Biyolojik ölçümler sadece biyolojik etkinliğin değerlendirilmesine izin veriyorsa bu yöntemler de gereklidir.

2.3. QSAR Modelinin Geliştirilmesi

Qsar modelinin geliştirilmesi için genel olarak üç bileşen gereklidir:

i. Bir bileşik grubu için deneysel biyolojik aktivite ölçümlerini içeren veri seti,
ii. Bu bileşik grubu için moleküler yapı ya da özelliklere ait veriler (değişkenler, tanımlayıcılar gibi),
iii. Bu iki veri seti arasındaki ilişkiyi kurmak için kullanılan istatistiksel metotlar.

i. QSAR modelinin geliştirilmesindeki sınırlayıcı faktör, yüksek kaliteli deneysel verilerin bulunmasıdır. QSAR analizinde deneysel verilerin hem doğru hem de kesin olması anlamlı bir model geliştirmek için şarttır. QSAR değerlendirmelerinde kullanılan veriler ya literatürden doğrudan alınır ya da QSAR-tipi analizler için özel olarak oluşturulur. Bu veriler benzer kimyasal bileşik serilerinden oluşur ya da kimyasal bir sınıf içerisinde bile yapısal farklılıkları sağlar. Bu farklılık daha zor QSAR modellerinin genelleştirilerek daha yaygın bir şekilde uygulanabilmesini sağlar.

İyi bir QSAR:

- Makul bir istatistiksel analiz için yeterli sayıda molekül incelenmeli
- Regresyon modeli için, nicel olarak ölçülmüş ve geniş bir aralıkta değişen biyolojik etkinlik değerleri olmalı (yani farklı büyüklüklerde değerler olmalı) ya da sınıflandırma modeli için, moleküllerin her sınıftaki dağılımı uygun olmalı
- Yeni kimyasalların güvenilir bir şekilde tahmini için uygulanabilir olmalı (değerlendirme ve uygulanabilirlik alanı)
- Modellenmiş sonuçtan teknik bilgi elde edilmesine imkân vermeli.

ii. Kimyasal tanımlayıcılar deneysel, kuantum kimyasal ve deneysel olmayan parametreler içerir. Deneysel parametreler ölçülebilir ya da tahmin edilebilir ve fizikokimyasal parametreleri içerir. Deneysel olmayan parametreler; atomlara, sübstitüntlere ya da tüm moleküle ait olabilir. Bunlar tipik yapısal parametrelerdir. İki boyutlu yapısal bilgilerden geliştirilebilir veya molekülün üç boyutlu konformasyonlarından hesaplanabilir.

QSAR modellerinde kullanılan çok sayıda parametre mevcuttur. Bunlar fizikokimyasal, kuantum kimyasal ve bağlanma özellikleridir. Fizikokimyasal parametreler bir molekülün çözünebilirlik ve iyonlaşma sabitleri yanında hidrofobik, elektronik ve sterik özellikleriyle ilgilidir. Kuantum kimyasal özellikler yük ve enerji değerlerini, bağlanma özellikleri de biyolojik makromolekülleri içerir.

Modern QSAR yaklaşımlarında teorik moleküler parametre setlerinin kapsamlı bir şekilde kullanılması gittikçe yaygınlaşmaktadır. Böylece bir kimyasalın bütün yapısal özellikleri alınarak moleküler yapı sayısal değerlere dönüştürülebilir. Teorik parametreleri kapsamlı olarak hesaplayan çok sayıda yazılım mevcuttur. Bunlardan bazıları şunlardır: Adapt [35], Oasis [36] Codessa [37], MolConnZ [38] ve Dragon'dur [39]. Teorik parametrelerin avantajı, belirli bir yazılım kullanılarak sentezlenmemişlerde dâhil olmak üzere tüm kimyasallar için hesaplanabiliyor olmasıdır. Bu ticari programlardan Codessa ve Dragon'un her ikisi de topolojik, geometrik, yapısal, termodinamik, elektrostatik ve kuantum mekaniksel özellikleri hazırlamaktadır. Codessa programı bu değişkenleri elde edebilmek için Gaussian ve Ampac output dosyalarını kullanırken, Dragon programı Sybyl, Molfiles, Multiple Sd, HyperChem, Smiles ve MacroModel output dosyalarını kullanır. Bu programlar ile bir bileşiğin sadece en düşük enerjili bir konformerini dikkate alarak parametreler hesaplanmaktadır.

Hâlbuki bir bileşiğin yüzlerce konformeri olabilmektedir. Gaussian, Ampac, Sybyl, Molfiles, Multiple Sd, Hyperchem, Smiles gibi programlar konformer hesabını yapamamaktadır. Bu nedenle Codessa ve Dragon programları sadece bir bileşiğin bir konformeri ile ilgili parametreleri hazırlamaktadır. Spartan [40] paket programı ise Monte Carlo yöntemine göre bileşiklerin konformer hesaplarını yapabilmekte relatif olarak enerjilerini hesaplayabilmektedir. Çalışmalarımızda bileşiklerin bütün konformerlerini dikkate aldığımız için Spartan programı çalışmalarımız açısından en iyi programlardan biridir.

Bileşiklerin bütün konformerlerini hesaplamaya dâhil ettiğimiz için CODESSA ve DRAGON programlarını moleküler parametreleri hazırlamak için kullanamamaktayız. Bu nedenle hem moleküler parametreleri hazılamak hem de ülkemizde kendi yazılımlarımızın geliştirilmesine katkıda bulunmak için çalışma grubumuz tarafından EMRE adı verilen bir program geliştirilmiştir. Geliştirilen bu program Spartan 02, Spartan 04, Spartan 06 ve Spartan 08 programlarının birbirinden farklı output dosyalarını kullanmaktadır. EMRE programının diğer iki programa göre üstünlüğü ise hem matris hem de farmakofor grup belirlendikten sonra farmakofor düzlemine moleküldeki en uzak atomun dik uzaklığı, farmakofor düzlemi ile iki atom arasındaki açı, atomlar arası açı, torsiyon açısı ve van der waals yarıçapı gibi çeşitli özellikleri de hazırlamasıdır. Maliyeti ise verilen diğer iki programdan daha düşüktür.

iii. Yapı-aktivite ilişkisinin uygun korelâsyonunu geliştirmek için seçilen QSAR metodu oldukça önemlidir. Buna rağmen moleküler değişkenler ile biyolojik aktivite arasındaki ilişki lineer olabilir ya da olmayabilir. Günümüzde QSAR modelinde en küçük kareler yöntemi (PLS) ile çoklu lineer regresyon (MLR) kullanılır. Biyolojik aktivite hesaplamalarında karşılaşılan en yaygın problem değişkenlerin sayısının bileşiklerin sayısını aşmasıdır. Hesaplama sonunda kalan parametre sayısının bileşik sayısını aşmaması istatistiksel hesaplamalarda önemlidir [41, 42].

Başlangıçta QSAR olarak ortaya çıkan bu metot zamanla gelişim göstererek yeni QSAR metotları şeklinde uygulama alanı bulmuştur. Bunlardan bazılarına aşağıda değinilmiştir.

2.4. 2D-QSAR Analizi

2D-QSAR analizi kantitatif ilişkileri elde etmek için lineer regresyon tekniklerini kullanır. 2D-QSAR da kullanılan moleküler parametreler genellikle üç grupta toplanır.

- Moleküler bağlanma gibi iki boyutlu parametreler
- Moleküler yüzey alanı gibi üç boyutlu parametreler
- LogP gibi fizikokimyasal özellikler

2D-QSAR parametreleri yapısal, topolojik, geometrik, elektrostatik, kuantum kimyasal ve termodinamik parametreler olarak sınıflandırabiliriz. En basit tanımlayıcı, geometrik veya elektronik yapıya bakmadan moleküler yapıyı yansıtan yapısal parametredir ki bu parametre atom sayısı, moleküler ağırlık, bağ sayısı ve H bağı sayısı ile ilgilenir. Topolojik parametreler, moleküldeki iki atom çifti arasındaki bağ bilgilerini kapsar. Geometrik parametre, yapının üç boyutlu koordinatlarını gerektiren, molekül yüzey alanı, moleküler yoğunluk ve hacim gibi özellikleri içine alır. Kuantum kimyasal parametreler ise moleküllerin elektronik özellikleri hakkında bilgi sağlayan HOMO ve LUMO gibi tanımlayıcılardır. Termodinamik parametreler, ilk olarak Hansch analizinde kullanılan molar refraktivite, oluşum ısısı, hidrofobik özellik ve oktanol ve sudaki serbest solvatasyon enerjisidir [41, 43].

2.5. 3D-QSAR Metodu

Biyolojik aktivite ile kimyasal bir serinin üç boyutlu özellikleri arasındaki bağıntıyı kullanan 3D-QSAR, son yıllarda QSAR metotlarının en dikkat çekeni olmuştur [44]. Hansch ve free-Wilson analizleri Üç boyutlu (3D) nicel yapı-aktivite ilişkisinin (3D-QSAR) temelini oluşturur [30, 45]. 3D-QSAR metodlarından günümüzde en çok kullanılan karşılaştırmalı moleküler alan analizi (comparative molecular field analysis, CoMFA) [46] ve Karşılaştırmalı Moleküler Şekil Analizi (comparative molecular similarity indices analysis, CoMSIA) [47] dir.

3D-QSAR da genel kurallar aşağıda maddeler halinde gösterilmiştir.

- Konformasyonel analiz tekniklerini kullanarak minimum enerjili konformer seçilir.
- İki atom arasındaki mesafe, halkalı sistemlerin optimizasyonu, dönor-akseptör hidrojen bağı ya da yüklü fonksiyonel gruplar tanımlanır.

- Aktif Küme (Perform Cluster Analysis) sınıflandırma metodları ile bağlayıcı konformer seçilir. "Aktif Küme" aktif bileşiklerin yüksek konsantrasyonu ve inaktif bileşiklerin düşük konsantrasyonuna sahiptir. Aktif kümede her bir yapının en düşük enerjili konformeri reseptör ile etkileşen konformer olarak alınır.
- Farmakoforun, dipol momenti, polaritesi, nokta yükleri, moleküler elektrostatik potansiyeli, HOMO ve LUMO orbitalleri, şekil ve hacim gibi özellikleri belirlenir.

Önemli farmakofor özellikleri seçilir [19].

2.6. 4D, 5D ve 6D QSAR Metodları

Son yıllarda 2D-QSAR ve 3D-QSAR metotlarına ek olarak yeni QSAR metotları geliştirilmiştir. 4D-QSAR, 5D-QSAR, 6D-QSAR, Hologram QSAR (HQSAR), Inverse QSAR, Binary QSAR bunlardan bazılardır. Aşağıda 4D-QSAR, 5D-QSAR, 6D-QSAR ayrıntılı olarak açıklanmıştır.

Hopfinger 1997 yılında 3D-QSAR modeline dördüncü boyutu kazandırdı ve 4D-QSAR analizi terimini ortaya çıkardı. Bu metodun CoMFA'dan başlıca farkı, konformasyonları eşleştirme işlemi yaparak, biyolojik aktivite bilgileri ile moleküler yapıların bir setinden geliştirilen modelde hem moleküler konformerleri hemde moleküllerin üst üste çakışması incelemelerini birleştirmesidir. 4D-QSAR analizinin dördüncü boyutu eşleştirme ortalamasından kaynaklanır. Bu metoda göre, çoğu 3D-QSAR metodlarından farklı olarak, aktif konformasyon minimum enerjili konformer değildir. Sistematik konformasyonel araştırma yerine, konformasyonel eşleştirme özelliklerini belirlemek için MDS (multi-temperature molecular dynamics) yi kullanır. Minimum enerjili konformerin 2 kcal/mol içinde olan tüm konformerleri çalışılan bileşik serisindeki her bir bileşiğin en aktif konformeri olarak tanımlanır. Düşük enerjili konformerlerin bu serisi en iyi sonucu veren 3D-QSAR modellerinde ayrı ayrı hesaplanır. Moleküler çakışma problemi benzer örnekleme ve hesaplama tekniği ile çözülmüştür. Her bileşiğin konformerleri, CoMFA metodunda olduğu gibi, en iyi çakışmayı bulmak için daha önceden belirlenmiş dikdörtgen şekilli bir yüzeye yerleştirilir ve 3D parametreleri bilgisayarda hesaplanır, daha sonra PLS teknikleri kullanılarak biyolojik aktivite ile korelâsyonu yapılır. 3D-QSAR modelleri PLS tekniği ile ortaya çıkarılmıştır. Bu metot en iyi çakışmayı sonuçlandırana kadar tekrar edilir. Aktif konformer, aktivite değerini

en iyi veren, uygun 3D-QSAR modelinden seçilir. Bu metot 3D-QSAR modellerini geliştirmek için kullanılmıştır [48].

Vedani son yıllarda QUASAR olarak adlandırılan paket programda 4D-QSAR yaklaşımını ortaya çıkarmıştır. Vedani ve grubuna göre bu program, çok yönlü konformasyonların etkilerinin birleştirilmesi ile 4D-QSAR kavramlarını kullanır. Bu metot özellikle, ligandın reseptöre bağlanma serbest enerjisi tahmin edilirken, reseptör yapısı bilinmediği zamanlarda kullanışlıdır [49].

Vedani tarafından geliştirilen 5D-QSAR yöntemi ile QSAR tekniklerine daha güçlü, yeni bir boyut kazandırıldı. 5D-QSAR, birçok ligandın proteine bağlanması sürecinde meydana gelen değişikliklerin de incelendiği bir yöntemdir [50].

6D-QSAR tekniği yine Vedani tarafından geliştirilen bir modeldir. 6D-QSAR tekniği diğer QSAR tekniklerine ek olarak çözücü molekülleri ve çözünen madde arasında zayıf kovalent bağların oluştuğu çeşitli çözünme modellerini de inceler [49]. Tablo 2.1'de bütün QSAR tekniklerinde kullanılan parametreler özet olarak verilmiştir.

Tablo 2.1. QSAR teknikleri ve kullanılan parametreler.

QSAR Tekniği	**Kullanılan Parametreler**
1D-QSAR	pKa, logP gibifizikokimyasal parametreler
2D-QSAR	1D-QSAR+Yapısal, geometrik, elektrostatik, termodinamik parametreler
3D-QSAR	2D-QSAR+Elektrostatik, sterik, hidrofobik parametreler
4D-QSAR	3D-QSAR+Konformasyonlar, protonlanma ve stereoizomerlerle ilgili parametreler
5D-QSAR	4D-QSAR+Ligand-protein bağlanmasında oluşan konformasyonel değişiklikler ile ilgili parametreler
6D-QSAR	5D-QSAR+Çözünme ile ilgili parametreler

2.7. Moleküler Modelleme ve QSAR

Moleküler modelleme, bilgisayarlı kimya ve grafiksel tasarım teknikleri kullanılarak moleküler yapı ve özellikleri araştırmak için kullanılan bir tekniktir [51]. Moleküler modelleme yaklaşımları, bilgisayarlı analiz ve interaktif tasarım olmak üzere iki ana

gruba ayrılır. Modern moleküler modelleme yaklaşımlarından en önemli üçü, yapı analizi, homolog modelleme ve etkileşimdir. Aslında en iyi modelleme teknikleri, moleküler dinamik, moleküler mekanik ve kuantum mekanik olmak üzere üç temel sınıfa ayrılır [52].

Çoğu moleküler modelleme çalışmaları üç basamak içerir. İlk adımda moleküler içi ve moleküller arası etkileşimleri açıklamak için bir model seçilir. Moleküler modellemede kullanılan en genel iki model QM (kuantum mekanik) ve MM (moleküler mekanik) hesaplamaları ile elde edilir. Bu modeller, hesaplanan sistemde moleküllerin ve atomların enerji düzenlemesine imkân tanır ve bu modeller atom ve moleküllerin pozisyonundaki değişimlerin sistemin enerjisini nasıl değiştirdiğini belirlemeyi sağlar. Moleküler modelleme çalışmalarının ikinci basamağında konformasyonel araştırma, Monte Carlo Simülasyonu veya moleküler dinamik hesaplaması yapılır. Son olarak hesaplama sonuçları analiz edilir. Sadece özellikler hesaplanmaz aynı zamanda uygun bir biçimde gerçekleştirilmesini yapar.

Kısaca moleküler modellemeye başlamak için üç ana metod kullanılır [53]:

- İlk olarak, özellikle bağ açısı ve bağ uzunlukları gibi geometrik özellikler yapılandırılır.
- Molekül optimize edilerek uygun geometriye sahip olan fragmentler ortaya çıkarılır.
- X-ışınları kristalografisi, nötron difraksiyonu veya NMR gibi fiziksel deney sonuçlarından elde edilen bilgiler kullanılarak yapılandırma işlemi yapılır.

2.8. Biyoalıcı İle İlaç Molekülün Etkileşim Mekanizması

Biyolojik aktivite, incelenen molekül ve biyoalıcının etkileşimi ile doğrudan ilişkilidir. Bu etkileşimin mekanizması oldukça karmaşık olduğu gibi biyokimyanın en zor problemlerinden birisidir. Biyoalıcılar hakkındaki bilgiler genellikle sınırlıdır ve ilaç molekülleri ile etkileşimleri, metabolizma gibi harici etkilerden dolayı oldukça karmaşıktır. Bu problem, ilaç-biyoalıcı etkileşiminin her bir kademesinde değişik biyolojik yöntemlere ve fizikokimyasal metotlara başvurmakla çözülebilir. Fakat genel olarak bu yol çok uzun ve çözülmesi zordur. Diğer taraftan bilinen moleküler yapı ve maddelerin biyolojik aktiviteleri üzerine oldukça fazla miktarda deneysel bilgiler mevcuttur. İlaç-biyoalıcı etkileşiminin biyomekanizmasını dikkate almadan, bu probleme aktif ve inaktif bileşikleri deneysel metotlarla tespit ederek aralarındaki

benzerlikleri ve farklılıkları açıklayan çeşitli çalışmalar bulunmaktadır. Bu çalışmalardaki bilgilere dayanarak, aktiviteden sorumlu bileşiklerin temel özelliklerine karar verilebilir ve biyolojik aktivite için bazı tahmini kurallar formüllendirilebilir.

Kullanılan QSAR metotlarında problemin en zor kısmı, biyoalıcı ile ilaç molekülünün etkileşim mekanizmasını net olarak ortaya koyamamasıdır. Bu nedenle, ilaç-biyoalıcı etkileşim mekanizması bilinmeyen bir kapalı kutu gibi düşünülebilir. Sadece bu kapalı kutudaki girdiler ve çıktılar analiz edilebilir ve ayrıntılı olarak karşılaştırılabilir. Bu işlemde, çıktılar, biyolojik aktivite olarak tanımlanan deneysel olarak belirlenmiş kantitatif biyoalıcı-ilaç molekülü etkileşimidir. Bu kapalı kutu için girdinin kesinliği daha da azdır. Hakikaten molekülün hangi tür özelliklerinin biyolojik aktivite formülünde kullanılacağı belirli değildir.

Molekülün hangi tür özelliklerinin biyolojik aktivite formülünde kullanılacağı belirli olmadığından incelenecek girdilerin kesinliği azdır. Birçok QSAR yaklaşımında, moleküler modelleme metodu bilgisayar hesaplamalarında kullanıldı (topolojik indisler, moleküler grafikler, kartezyen koordinatlar, atomik yükler, bağ mertebesi, bağ uzunlukları gibi). Molekülü tanımlayan bu metotlar genellikle biyolojik aktiviteyi doğru şekilde tahmin edecek kadar etkili değildirler. Tanımlama biçiminin parametreleri genellikle fiziksel ve kimyasal değişkenlere gereksinim duyarlar. Bu parametreleri kimyacılar geçmiş deneyimlerine dayanarak doğrudan kullanamazlar ve direkt olarak yeni bileşiklere uygulayamazlar. Bilgisayar hesaplamalarıyla QSAR' da çözüm üretmek bu yüzden anlamlı ve temel gereksinimdir. QSAR metotlarının ana problemi doğru ve yeterli parametreleri bularak genel ve hassas biyolojik aktivite tahmininde bulunmasındaki zorluğudur.

Bizim çalışmalarımızda ligat temelli 4D-QSAR Metodlarından olan ve çalışma grubumuz tarafından geliştirilen Elektron Konformasyonel-Genetik Algoritma (EC-GA, EMRE) metodu kullanılmıştır. Daha önceki çalışmalarımızda bu metod başarılı bir şekilde uygulanmış olup tartışma ve sonuçlarla ilgili bilgiler makalelerde verilmiştir [54, 55]. Bu metodun ayrıntıları 3. bölümde sunulmuştur.

3. BÖLÜM

MATERYAL VE METOT

3.1. Materyal

Bu çalışmada:

1. Spartan 02 ve Spartan 06 paket programlarında konformasyon ve kuantum kimyasal hesaplamalar yapılmıştır. Bu çalışmada; semiempirik hesaplamalar ile konformasyon ve kuantum kimyasal hesaplamalar yer almaktadır.
2. Her bir bileşiğin her bir konformerinin yapısını üç boyutlu olarak belirleyen topolojik matrisleri (Electron Conformational Matrix of Congruity, ECMC) hazırlayan ve aktivite hesaplamasında kullanılmak üzere bileşiğin özelliklerini gösteren sayısal parametreleri hazırlayan EMRE programı,
3. ECMC matrislerini atomlara göre karşılaştırarak elektron konformasyonel alt matrisi (Electron Conformational Sub Matrix, ECSA) belirleyen ECSP programı,
4. Matlab 7.0 paket programı ve bu program ile yazılarak biyolojik aktiviteyi hesaplayan program kodları,
5. Yazılımların geliştirilmesi için Delphi 7.0 ve C programları,
6. HP Workstation (iş istasyonu) xw 8200, Fujitsu-siemens celcius iş istasyonu,
7. Core 2 Quad işlemcili 5 adet pc, Dual 2 core işlemcili 1 adet pc, HP Notebook,
8. Veri depolama aygıtı,
9. 10 kVA'lık online güç kaynağı kullanılmıştır**.**

3.2. Metot

Çalışmalarımızda kuantum kimyasal hesaplamaları yapmak üzere spartan 02 ve 06 programı kullanılmıştır. Kuantum kimyasal hesaplamalar içerisinde semiemprik, HF, DFT, Moller Plesset (MP2) olmak üzere çeşitli hesaplama yöntemleri bulunmaktadır.

Bu metotlar arasında en hızlı olan semiemprik metodudur. Biyoalıcı ile molekülün etkileşimde bileşiğin geometrik yapısı çok önemlidir. Semiempirik yöntemlerden molekülün geometrik yapısını en iyi veren metot PM3 metodu olduğu için bizim çalışmalarımızda bu metot kullanılmıştır. Teorik aktivite hesaplamaları Matlab 7.0'da yazılan kodlar yardımıyla yapılmaktadır. Hesaplamalar hibrid metot olan Elektron Konformasyon-Genetik Algoritma (EC-GA) metodu kullanılarak yapılmıştır.

3.3. Elektron Topolojik Metot

A.S. Dimoglo ve çalışma arkadaşları tarafından geliştirilen elektron topolojik metodun [56] genel özellikleri aktiviteyi bazı alt yapısal özelliklere bağlaması, moleküllerin yapısını elektronik faktörlerle tanımlayan bir matris kullanması ve aktivite özelliklerinin seçimi için olasılık tahmini yapılmasıdır. Elektron Topolojik Metot (ETM) moleküler mekanik yöntemle bileşiklerin Z matrisini çıkararak eski programlardan olan CNDO yöntemiyle elektronik ve geometrik hesaplamalar yapmakta daha sonra bileşiklerin sadece tek konformeri için üç boyutlu matris oluşturarak pha'yı bulmaktadır. Elektron-Topolojik Metodu ile her bir molekülün Elektron-Topolojik Uygunluk Matrisi (ETMC) ya da üç boyutlu ETMC (TDETMC) matrisi hazırlanır. Molekülün elektronik özelliklerini tanımlayan atomik yükler valens aktiviteleri polarize edilebilirlik ve HOMO-LUMO enerjileri gibi atomik parametreler köşegen elemanları olarak seçilir. Köşegen dışı elemanlar; polarize edilebilirlik, bağ mertebesi (Wiberg indeks) ve bağ enerjisi gibi bağın elektronik parametrelerinden birisi olabilir. ETM, SAR yöntemleri gibi analiz edilen çok sayıda veri oluşturur ve başlangıç olarak bileşikleri aktif ya da inaktif olarak sınıflandırır. Ele alınan bileşikler benzer temel yapıya sahip değil ise klasik QSAR teknikleri ile ETM verilerini analiz etmek ve aktif bileşikleri tahmin etmeyi imkânsız kılmasa da zorlaştırır. ETM verileri üç boyutlu kimyasal yapı ile biyolojik aktiviteyi yarı nicel (aktif ya da inaktif) olarak ilişkilendiren önemli bir veri tabanı sağlamaktadır.

ET metodunda aktivite kantitatif olarak biliniyorsa incelenen seri belli bir aktivite değerinden itibaren ikiye bölünür. Bu değerin üstündekiler aktif altındakiler inaktif olarak kabul edilir. Her grup için alt matris bulunur. Bulunan alt matrislerin elemanlarının karşılaştırılması ile aktivite alt matris parametrelerinin bir fonksiyonu olarak ifade edilir. Oysa inaktif olarak nitelendirilen bileşikler aktif olarak

sınıflandırılan bileşiklere göre düşük olsada aktivite göstermektedirler ve aktivitelerindeki bu düşüş moleküle bağlı olan sübstitüentlerin farklı olmasından kaynaklanmaktadır. Çoğu zaman aktiflikleri arasında büyük bir fark olmadığı için keyfi olarak yapılan bu bölme hatalara götürmektedir. Bu yönüyle sübstitüentlerin etkisini dikkate almamaktadır. Bu da yöntemin en önemli eksikliklerinden biridir. Bunu elimine etmek için zaman zaman birden fazla farmakofor hesaplama yoluna gidilmiştir. ETM'de Pha sadece yüksek aktiviteli bileşiklerde aranmakta, ECM'de ise incelenen bütün bileşik serisi içinde aranmaktadır.

ETM de bileşiklerin bütün konformerleri değil de tek bir konformeri hesaba katılır. Biyoalıcının aktif durumdaki üç boyutlu yapısı bilinemediği için biyoalıcının ilgili molekülün konformerlerinden (bir bileşiğin yüzlerce konformer yapısı bulunmaktadır.) hangisi ile etkileştiği bilinmemektedir. Hâlbuki moleküllerin düşük enerjili olan birden çok konformeri mevcuttur ve bunların o bileşiğin aktivitesine katkısı ihmal edilemez. Her bir konformerin aktiviteye katkısı Boltzmann dağılımına göre dikkate alınması gerekir. ETM' de göz ardı edilen bu etken Elektron Konformasyonel Metot (ECM) adı verilen bir metotta ise sadece farmakofor grubu bulmak için dikkate alınmıştır. Aktivite hesaplamalarında ise konformerler hesaba katıldığında işlemin çok karışık olduğundan aktiviteyi bu şekilde hesaplayamamışlardır ve bu nedenle konformerleri dikkate almamışlardır. Bizim hesaplamalarımızda hem ECMC ler karşılaştırılırken hem de aktivite hesaplamaları yapılırken bütün konformerler dikkate alınmıştır.

3.4. Elektron Konformasyonel Metot

EC metotta [57–60] ise ETM'den farklı olarak bileşiklerin konformerleri de hesaba katılarak ve üç boyutlu matrislerini kullanarak farmakofor grubu belirlenmiştir. Bunun için bütün bileşiklerin tüm konformerleri referans bileşik ile verilen tolerans aralığında karşılaştırılır. Tüm konformerler de ortak olan atom grubu belirlenir. Metot biyolojik aktivitesi bilinen bir bileşik serisinde belirli biyolojik aktiviteden sorumlu tek bir Pha olduğunu varsayar ve prensipte bir yapı aktivite ilişkisidir. Ancak ECM'de Pha'nın varlığı aktivite için gereklidir fakat yeterli şart değildir. Bunun için hem sterik engeller hem de aktiviteyi azaltan (APS) veya aktiviteyi artıran (AG) gruplar da hesaba katılır [61]. Ayrıca Bersuker farmakoforlara bağlı olarak aktivitenin hesaplanmasında

kullanılmak üzere eşitlik 3.3'de verilen aktivite denklemini geliştirerek eğitim ve test bileşiklerinin aktivitelerini hesaplayabilmektedir.

3.4.1. ECMC Matrislerinin Oluşturulması

ECMC, molekülün hem elektronik hem de geometrik yapı özelliklerini içerir. Daha önceleri ETM de, her bir molekülü temsil eden ETMC için sadece en düşük enerjili olan temel hal (ground-state) konformeri dikkate alınırdı. Ancak EC metodunda, relatif enerji farkı 1.5 kcal/mol den büyük olan konformerler silindikten sonra, geriye kalan düşük enerjili her bir konformer için ECMC matrisleri oluşturulur [58, 59].

EC metodunda kullanılan ECMC nin diyagonal elemanları, ilgili atomun yüklerini gösterir nondiyagonal elemanlar birbirine bağlı atomlar ise bağ derecesi, bağlı değilse atomlar arasındaki mesafeyi gösterir. ETM metodunda atom yükleri yaygın olarak kullanılmıştır [62]. I.B. Bersuker, EC metodu ile gerçekleştirdiği çalışmalarında atom yükleri yerine etkileşim indeks değerini kullanmıştır. Bu çalışmada ise mulliken yükleri kullanılmıştır.

3.4.2. Farmakofor Tanımlama

İlaç tasarımında (Pha) terimi ana kavramlardan biri olup 1900'lerin başında Ehrlich [63] tarafından ortaya çıkarılmış ve Pha, araştırmacılara bağlı olarak çeşitli şekillerde tanımlanmıştır. Ancak Pha çoğunlukla, kesin bir biyolojik aktivite sağlamak zorunda olan bileşiğin, kimyasal, yapısal ve fiziksel özelliklerinin kombinasyonlarını kapsayan bir grup özelliği olarak ifade edilir [60]. İncelenen molekül serisinde ilaç etkisinden sorumlu olan Pha grubudur. Deneme setindeki aktif moleküllerde aktiviteden sorumlu belli bir geometrik düzende üç boyutlu yapısı olan atomlar veya atom gruplarına farmakofor denir. Bu tanımlama, belli bir moleküler sistemde Pha'nın varlığında veya yokluğunda aktivitenin kalitatif olarak (aktivite var yada yok) belirlenmesinde kullanılabilir. Fakat Pha'nın var yada yok şeklinde tanımlanması, aktifliğin kantitatif ölçümü için yetersizdir. Örneğin Pha var fakat aktiflik küçük ise, bileşik deneysel olarak inaktif şeklinde sınıflandırılabilir [61].

Reseptörün üç boyutlu yapısı bilinmediği durumlarda farmakofor tanımlaması, yeni farmakolojik ilaçların tasarımında bir anahtardır ve oldukça önemlidir. Pha modelleri, reseptör yapıları deneysel olarak açıklanamadığı zaman veya açıklanması zor olduğunda

reseptör-ligand etkileşimini açıklamada kimyacılara yardım eder. Farmakofor kavramı Bilgisayar Destekli İlaç Tasarımında reseptör veya enzim ile etkileşen çeşitli moleküllerdeki genel üç boyutlu özelliklerin araştırılmasında kullanılır [64]. Pha çoğunlukla, biyolojik aktivite sağlamak zorunda olan bir bileşiğin, fizikokimyasal (logP, ovality, hacim, yüzey alanı vb. gibi) geometrik (uzaklık, açı, torsiyon açısı, atomların farmakofor düzlemine uzaklığı ve açısı vb. gibi) ve elektronik (HOMO, LUMO, yükler, kimyasal sertlik, dipol moment, elektronegatiflik vb gibi) özelliklerinin kombinasyonlarını kapsayan bir grup özelliği olarak ifade edilir [60, 65–69].

I.B.Bersuker EC metodunda, Pha nın tanımını aşağıdaki üç önemli aşamada göstermiştir.

1. Pha sadece atomlar ile değil moleküldeki atomların elektronik özellikleri ile de ilgilidir. Bilindiği gibi farklı bileşiklerde, aynı atomlar için farklı ve farklı atomlar için aynı olabilen elektronik özellikler olabilir.
2. Pha nın geometrik parametreleri ve elektronik özelliklerinin her ikisi de aktif bir bileşikten diğerine belirli sınır değerler içerisinde değişebilir ve aktivite bu değişikliklerin (Pha esnekliği) bir fonksiyonu olabilir.
3. Pha aktivitenin tanımlanmasında gereklidir fakat tek başına yeterli değildir. Pha nın olmasına rağmen aktivite Pha dışındaki gruplar tarafından azaltılabilir (APS) ya da artırılabilir (AG). Bu gruplar, Pha nın biyoalıcı ile uygun etkileşimini engelleyen APS ve moleküldeki hidrofobisite gibi özellikleri sağlayan AG gruplarıdır [61].

Pha'nın belirlenmesi için burada anlatılan bilgiler temel olarak alınır. İlaç moleküllerinde, konformerleri, geometrisi ve elektronik özellikleri önemlidir. EC metodunda Pha belirlenmesi, kuantum kimya hesaplamalarından elde edilen moleküler özelliklerin bilgisayarla matris haline getirilmesi ile kantitatif biyolojik aktivite tahmini ise regresyon analizi ve molekülün özelliklerinin değerlendirilmesi ile yapılır [61]. Bunun sonucu olarak Pha, farklı bileşiklerin bir serisinde gözlemlenmiş biyolojik aktivitenin nedenini anlamak ve daha aktif bileşiklerin tasarımına yardımcı olmak için kullanılır.

EC metodunda bileşiklerin ECMC'lerinde matrisin diyagonal ve non-diyagonal elemanlarını verilen belirli tolerans aralıklarında, karşılaştırarak bütün aktif moleküller için ortak olan, inaktif moleküllerde ise bulunmayan elektron topolojik özellikler tespit edilir. Matrisleri karşılaştırmak için bileşik serisindeki istenilen yapı referans olarak

seçilebilir. Çalışmalarda hangi bileşiğin hangi konformerinin referans olarak seçildiği belirtilmelidir. En basit yapılı bileşiğinde aktivitesi olduğuna göre farmakofor grubu bu bileşikte de bulunmak zorundadır. Bu nedenle referans alınmasında bir mahsur bulunmamaktadır. ETM de ise aktifliği en fazla olan molekül referans olarak seçilir ve bu yöntemde bileşikler aktif ve inaktif olarak ikiye bölündüğü için matrislerin karşılaştırılmasında aktivitelerde dikkate alınır.

ECMC'ler hazırlandıktan sonra referans bileşiğe göre karşılaştırma işlemi yapılarak bütün aktif bileşiklerin bulundurduğu ortak özellik Pha olarak bulunur ve bu özelliği temsil eden ve matris elemanları sayısı daha az olan, elektron konformasyonal alt matrisi (ECSA) oluşturulur.

QSAR çalışmalarında bilinen yöntemlerin çoğu bileşiklerin enantiyomerlerinden birini diğerinden ayırmak için yetersiz kaldığı halde, ECMC de enantiyomerlerin farkı ortaya konulabilir. İki enantiyomerin matris elemanlarının mutlak değerleri aynı olmalarına rağmen, iki matrisdeki karşılıklı pozisyonları farklıdır (açı ve torsiyon açısı). Eğer ECSA (Pha) enantiyomerlerden birinde var ise, diğer enantiyomerde olmasını gerektirmez; eğer Pha her ikisinde de varsa, anti pharmacophore shielding (APS) ve auxiliary group (AG) gruplarının pozisyonları farklı olacağından farklı aktiviteleri ortaya çıkarırlar. Moleküler sistemde atomları numaralandırma yolu (saat yönü veya tersi) önemlidir. Çünkü bu işlem ECMC de matris elemanlarının düzenini belirler. Bir tanesinin seçilmesiyle bütün konformerler ve enantiyomerler için korunur.

ECSA sonuçları, inceleme altındaki aktif Pha yı gösterir. Matris elemanları elektronik ve konformasyonal özelliklere karşılık geldiği için, ECSA prensip olarak, molekülün aktif konformeri, topoloji ve yük dağılımı ile açıklanan, aktif bölgesi (Pha) hakkında bilgi verir. Tolerans değerlerinin verilmesi şartı ile, elektronik ve konformasyonel parametreler kullanılarak aktifliğin sayısal değeri, aktiflik kalitatif (aktiflik var yada yok) olarak bilindiği zaman bile kantitatif olarak elde edilebilir. Aktiflik kantitatif olarak bilindiği zaman bu işlem aktiflik değerinin her bir değeri için tekrarlanabilir ve daha sonra elektronik ve konformasyonel parametreler aktifliğin bir fonksiyonu olarak dikkate alınır. ECSA tarafından bulunan Pha tanımı, aktivite için yeni bileşiklerin taranmasında kullanılabilir.

Bir bileşik serisinde aktivitenin azalmasına Pha nın olmaması ya da aktivite azaltıcı grupların varlığı sebep olabilir. Matrislerin minimum aralığını (tolerans değerleri)

karşılaştırmak için bileşik serisindeki aktif ve düşük aktiviteli bileşikler şeklinde sınıflandırmak gerekir. Pha deneme setinden elde edilen ECSA, diğer aktif bileşiklerin ECSA ları ile karşılaştırılr.

ECMC'nin ETMC ye göre diğer üstünlüğü de AG ve APS grupları [60, 70] tanımlayarak hesaba katmasıdır. Bu grupların aktiviteye katkıları, Pha ile birlikte hesaplanmasıdır. Aslında Pha'i bulunduran bir konformerdeki her bir atomun aktiviteye sayısal bir değer yansıtması söz konusudur ancak bu atomların incelen seri içinde molekül içinde bulunan en etkili olan parametreler göz önüne alınabilmektedir.

3.4.3. APS ve AG Grupları

Pha aktivite için çok önemli şartlardan biri olmasına rağmen tek başına yeterli değildir. Aynı Pha ya sahip farklı bileşiklerin neden farklı aktivite gösterdiğini açıklayamaz. Pha aktivite için gerekli bir grup atomu içerir ancak molekülde Pha dışında aktiviteye etki eden diğer gruplarda (Out-of-Pha groups, OOP) bulunmaktadır. Bu gruplar atomlar olabildiği gibi atom grupları da olabilir ve aktiviteyi artırıcı ya da azaltıcı etki gösterebilirler. Bunlar sterik sınırlamalara neden olan APS gruplar ve aktiviteye hidrofobisite gibi etkilerle katkıda bulunan AG gruplardır ki bu gruplar hidrojence zengin olan ve aktiviteyi artırıcı özellik gösteren gruplardır. Bu atomların hidrofob özellikleri aktiviteyi artırıcı özellik gösterirler. APS gruplar ise biyoalıcı ile etkileşim esnasında sterik engel veya perdeleme yaparak aktiviteyi azaltıcı yönde etki ederler.

Pha dışındaki bazı gruplar biyoalıcı ile etkileşime katılmaz. Pha ile bu inert gruplar molekülün temel iskeletini oluştururlar. Pha ile temel iskelet arasındaki fark, yaklaşık olarak aynı aktiviteye sahip olan bileşikler için ECMC elemanlarının farklı olması veya bu elemanların olmamasıdır. Bu gruplar, eğer varsa, Pha nın dışında kalmış fazla etkisi olmayan bölgelerini veya biyoalıcının boşluklarındaki bazı boş yerleri işgal edecekleri için hiçbir katkıda bulunmazlar ve bunlar ligand-biyoalıcı etkileşimine katılmazlar. Temel iskelet kabaca, yaklaşık olarak aynı aktiviteye sahip farklı bileşiklerin üst üste çakışmaları ile ortaya çıkarılabilir.

AG ve APS grupların etkisi onların pozisyonuna ve reaktifliğe bağlıdır. Pha atomlarından farklı olarak APS ve AG grupların parametreleri ve bunların sayısal değerleri bir bileşikten diğerine değiştiği gibi konformerden konformere de değişmektedir. Bu grupların her birinin aktiviteye katkısını açıklamak için ECMC den

elde edilen elektronik ve yapısal parametreler kullanılır. Bileşikten bileşiğe APS ve AG grupları ve bu grupların pozisyonları değiştiği için bu parametrelerin (sayısı ve çeşitliliği) seçimi kısmen keyfidir. Bunların seçimi için en iyi model regresyon analizidir. APS ve AG'ye karar vermek için aktif bileşiklerin üst üste çakıştırılmış yapılarını incelememiz gerekir. Biyoalıcıdaki konformasyonel değişiklikler bilinirse AG doğrudan hesaplanabilir ancak çalışmalarda biyoreseptör bilinmediği için AG yi tanımlamak daha zordur [61].

3.4.4. Aktivitenin Genel Formülü

APS grupları aktiviteyi kısmen veya tamamen azaltan ve sterik engele neden olan atom grupları olarak da tanımlanır. Eğer biyoalıcı yapısı biliniyorsa hem sterik engeli hemde aktiviteyi azaltan yükleri doğrudan hesaplanabilir. Biyoalıcının yapısı bilinmiyorsa bu işlem zorlaşır. Bu problemi çözmek için, Pha da olduğu gibi bileşiklerin relatif aktivite değerleri ile yapısal ve elektronik parametrelerin istatistiksel karşılaştırılması yapılabilir. Bunun için EC metotta hem sterik engelleri hemde aktiviteyi azaltan yükleri ortaya çıkarmak için APS değerlerini kantitatif olarak (sayısal olarak) karakterize eden S ile gösterilen bir parametre dikkate alınır. Bileşikler ile biyoalıcılar arasındaki etkileşim diğer moleküller arası etkileşim (en azından, sterik ve diğer enerji bariyer faktörleri göz önüne alınabilir) ile benzediğinden dolayı aktivite (A) eksponansiyel olarak (e üzeri bir fonksiyon olarak) S ye bağlıdır. Yani AG ve APS gruplarının etkisini hesaba katmak için aktiviteyi artıran ya da azaltan S fonksiyonunu kullanırız. S fonksiyonu AG ve APS'yi temsil eden bir grup parametreye bağlıdır.

$$A \sim e^{-S} \tag{3.1}$$

Molekül-biyoalıcı etkileşiminde AG ve APS grupları exp(-E'n/kT) nin bir faktörü olarak E' miktarı kadar aktiviteyi artırır ya da azaltırlar. APS için E'>0 ve aktivite azalırken, AG grupları için E'<0 ve aktivite artar. $S_{ni}=E'_{ni}/kT$ olarak gösterilmiştir ve genel olarak S fonksiyonu aşağıdaki eşitlik ile ifade edilir.

$$S_{ni} = \sum_{j=1}^{N} \kappa_j a_{ni}^{(j)} \tag{3.2}$$

Burada $a_{ni}^{(j)}$, (ECMC de verilen a_{ii} parametrelerden farklıdır.) n'inci bileşiğin i'inci konformerindeki APS veya AG nin j'inci tür özelliğini temsil eden (bağ açısı, torsiyon

açısı, yük, atomik etkileşim indeksi (*II)* ve atomlar arası uzaklık) bir parametre olup her bir problem için spesifik olduğu söylenilebilir. N seçilen AG parametrelerinin sayısı ve κ_j toplu analiz sonucundan elde edilen değişken bir katsayıdır.

$a_{ni}^{(j)}$ parametresinin seçimi önemlidir. Pha için $a_{ni}^{(j)}$ parametresi atomik etkileşim indeksi, değerleri ve ECSA'dan elde edilen Pha atomları arasındaki mesafeler olarak alınabilir. APS gruplar için $a_{ni}^{(j)}$ parametresinin seçimi daha fazla kimyasal bilgileri gerektirir. APS parametreleri ise Pha nın ve temel iskeletin dışında kalan pozisyonlarına göre belirlenen sterik faktörler olarak alınır. κ_j değerleri ilgili parametrenin birimine bağlı olarak farklı birimlerde değerler alır ve o parametre ile κ_j nin çarpımı birimsiz bir ifadedir. κ_j nin daha büyük değeri, hesaplanan biyoaktiviteye karşılık gelen parametrelerde daha etkilidir. Yani $\kappa_j>0$ ve $a_{np}^{(j)}< a_{lp}^{(j)}$ için aktivite bu bileşiklerde artar ve $a_{np}^{(j)}> a_{lp}^{(j)}$ ise azalır. $\kappa_j<0$ ise tam tersi doğrudur. Kısacası κ_j katsayısının mutlak değerinin büyüklüğü, aktivite üzerindeki etkiyi gösterir ve APS ve AG gruplarını üç boyutlu yapı içinde tanımlamak Pha'dan sonra ikinci derecede öneme sahiptir [61].

Konformasyonel analiz, bilgisayar destekli ilaç tasarımında genellikle oldukça karışık olarak görülen bir problemdir ve çoğunlukla organik yapılara sahip ilaç moleküllerinin birçok konformeri vardır. Bu nedenle bu problemi basitleştirmek için ilaç-biyoalıcı bağlanmasındaki enerji farklılıkları için, eşitlik 3.2 ile birlikte sıcaklığın ve enerjinin bir fonksiyonu olarak her bir konformerin Boltzman dağılımını hesaba katarak n. bileşik için aktivite genel formülünü aşağıdaki eşitlik halinde elde ederiz:

$$A_n = A_l \frac{\sum_{i=1}^{m_l} e^{-E_{li}/kT} \sum_{i=1}^{m_n} \delta_{ni}[Pha] e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_n} e^{-E_{ni}/kT} \sum_{i=1}^{m_l} \delta_{li}[Pha] e^{-S_{li}} e^{-E_{li}/kT}} \qquad (3.3)$$

A_l aşağıda açıklanacağı üzere sabittir ve m_n n. molekülün konformer sayısıdır. Bu formülde δ, Dirac δ fonksiyonunun bir çeşididir.

$$\delta_{ni}[\text{Pha}] = \begin{cases} 0, \text{ Pha bulunmadığı zaman} \\ 1, \text{ Pha bulunduğu zaman} \end{cases}$$

Bu formülde sadece Pha'i bulunduran konformerler aktiviteye katkıda bulunur ve bu katkılar onların APS ve AG gruplarının etkilerine göre ağırlıklı olarak alınır ve enerji aktif konformasyonlardaki moleküllerin relatif sayısına bağlıdır. Bu sayılar E_{ni}

konformerin enerjisiyle hızla azalır (E_{ni} de konformasyonun enerji değeri yaklaşık 2.8 kcal/mol ise tüm konformerler göz önüne alındığı zaman bu değerin katkısı 0.01 den daha azdır.) [61]. Bu formülde A_l sabittir ve bu değeri belirlemek için, çalışılan molekül setinden (eğitim seti) aktivitesi bilinen bir referans molekül (l) seçeriz ve bu molekül için aktiviteyi (A_n) eşitlik 3.3 ile hesaplarız.

APS/AG grupların parametrelerinin tahmini, onları ECMC den ayıracak bir algoritma bulunmadığı için zaman alıcıdır ve böylece her bir konformer için ayrı ayrı hesap yapılması gerekir. Özellikle reseptör yapısı bilinemediği durumlarda, reseptörle etkileşim sırasında molekülün yapısında meydana gelen konformasyonel değişim ihtimalinden dolayı belirsizlik ortaya çıkar.

Bu problem molekül-biyoalıcı etkileşimlerinin bazı genel özelliklerini dikkate alarak çözülmüştür. İlk olarak, enzim katalizli reaksiyonlarda, reaksiyonun daha önceden dengede olduğu hesaba katılarak Boltzman dağılımı kullanılır. Önemli bir nokta da, önemli yoğunluktaki konformerlerin arasında 2–3 kcal/mol olan enerji farkından daha büyük olan substrat-enzim bağlanma enerjisidir (10–20 kcal/mol). Bu durumda ligandın en az bir konformeri Pha ya sahiptir ve diğer tüm konformerler enerji kazanmak ve resptöre bağlanmak için denge dönüşümüne uğrarlar [61].

$$A_n = A_l e^{-\left(E_{np} - E_{lp}\right)/kT} e^{-\left(S_{np} - S_{lp}\right)} \tag{3.4}$$

Bersuker eşitlik 3.4'de gösterilen formül kullanarak bir bileşiğin tüm konformerlerini hesaba katarak aktivite hesaplaması için algoritma geliştiremediğinden dolayı sadeleştirilmiş formül olan ve en düşük enerjili bir konformeri hesaba katan eşitlik 3. 8'yi kullanmıştır. Diğer konformerlerin inaktif pozisyonda olduğunu kabul etmiştir. Bu çalışmanın en zor kısmını kj değerlerini hesaplama kısmı oluşturmaktadır. Bersuker çalışmalarında cross-validation yöntemlerinden olan leave one out metodunu kullanarak en uygun kappa sayısını ve değerini hesaplama yoluna gitmiştir [61].

Molekül setindeki bileşiklerin bilinen deneysel aktiviteleri kullanılarak yukarıdaki formül çözülür ve hesaplanan aktivite değeri A_n^{hes}, deneysel aktivite değeri A_n^{den}, ve N molekül sayısı olmak üzere, $\sum_{n=1}^{N} \left| A_n^{den} - A_n^{hes} \right|^2$ fonksiyonu üzerinden en küçük kareler yöntemi kullanılarak değişken κ_j katsayıları elde edilir. Bu yolla elde edilen κ_j ile

aktiviteleri bilinmeyen aynı tür başka bileşikler için aktivite tahmini yapılabilir. Δ, Pha'dan dolayı SE bağlanmasının enerji farkıdır ve E'_{np} ise APS ve AG gruplarından dolayı bağlanma enerjisindeki azalmadır. Bağlanma ve dönüşümden sonra yani 300 K de E_0 yerine $-\Delta + E_{np}+E'_{np}$ eşitliği kullanılır [61].

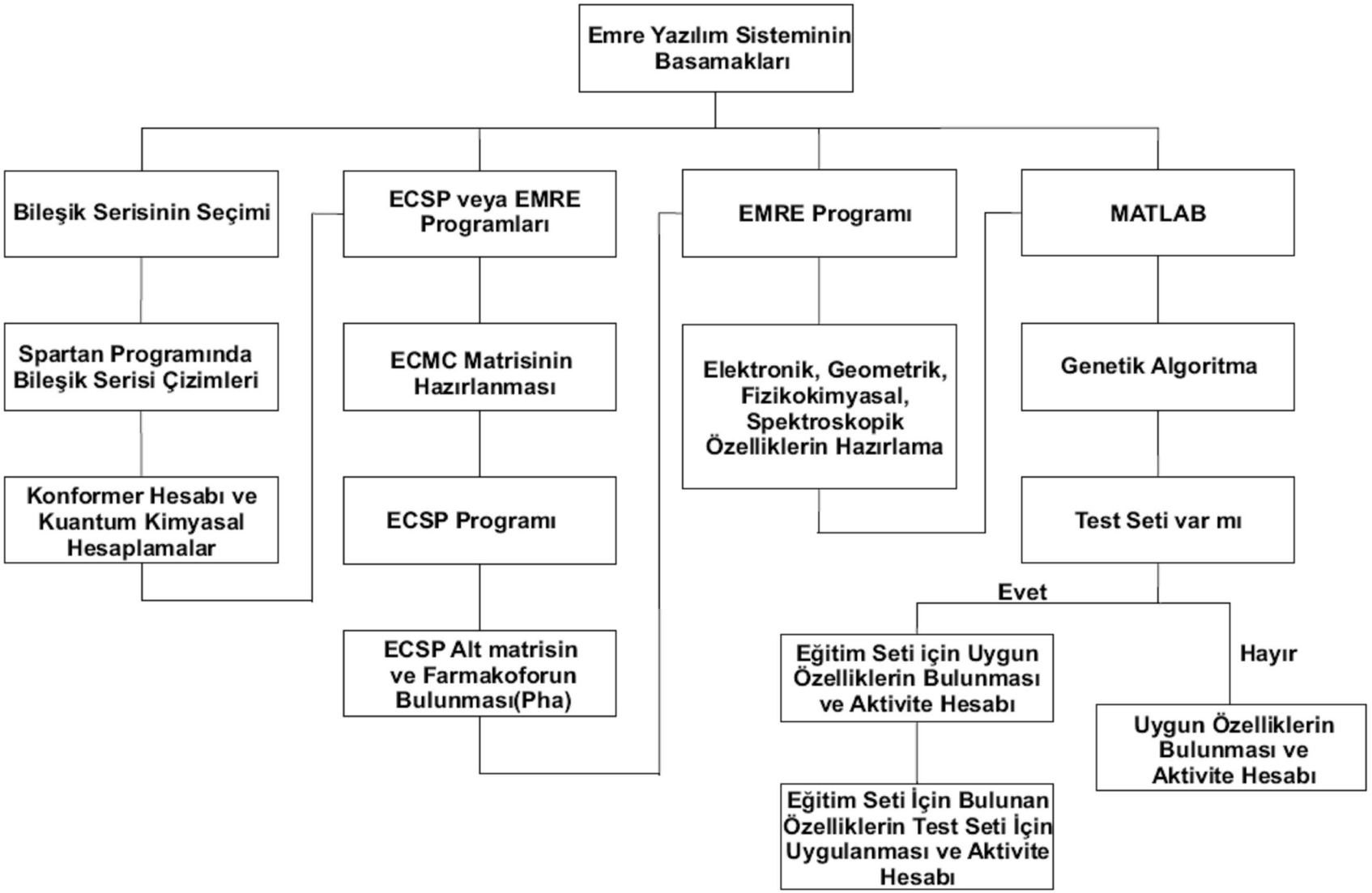
EC Metodu bizim çalışmalarımızın temelini oluşturmaktadır. EC metodunu esas alarak geliştirilen EC-GA, Elektron Konformasyon-Genetik Algoritma, yöntemi biyolojik aktivitenin hesaplamaları gerçekleştirilmiştir.

3.5. Elektron Konformasyon-Genetik Algoritma (EC-GA) Metodu

EC metodunda birini dışarıda bırakarak çapraz doğrulama (leave-one-out cross validation) yöntemi ile sadece en düşük enerjili tek bir konformer kullanılarak teorik aktivite hesaplamaları yapılmaktadır. Farmakofor belirlenmesinde ise serideki bileşiklerin bütün konformerlerini dikkate almaktadır [58–60, 62, 71]. Bu çalışmada kullanılan ve bizim tarafımızdan geliştirilen paket programlarda ise (EMRE, ECSP programlarının ve Matlab 7.0 için yazılan kodlar) bütün konformerler aktivite hesaplamalarına da dâhil edilmiştir.

EC-GA Metodu, Pha grubu, aktiviteyi artıran (AG) ve aktiviteyi azaltan (APS) grupları tanımlamak ve biyolojik aktiviteyi hesaplamak için kullanılan bir metottur. EC-GA Metodu, EC Metottan farklı olarak aktiviteye etki eden parametrelerin belirlenmesi ve aktivite hesaplanmasında genetik algoritma optimizasyon tekniğini kullanmaktadır.
Bu metotta molekülü ve konformerlerini tanımlayacak yüzlerce parametre hazırlamak ve bunların içinden hem eğitim hem de test seti için en uygun olan parametreleri seçmek mümkündür. Çünkü bazı parametreler eğitim seti için uygun olurken test seti için kötü olabilmektedir.

İleride ayrıntılı olarak açıklanacağı gibi geliştirilen bu yazılımların tamamını şekil 3.1'de gösterilen ilaç tasarımında EMRE yazılım sistemi şeklinde isimlendirilmiştir. Şekilden görüldüğü gibi EC-GA metodu içerisinde Spartan paket programı, çalışma grubumuz tarafından geliştirilen EMRE, ECSP programı ve son olarakta yine çalışma grubumuz tarafından geliştirilen, aktiviteyi ve kj değerlerini hesaplamak için Matlab 7.0 paket programında genetik algoritma optimizasyon tekniğini kullanarak hesaplamaları gerçekleştiren yeni bir yazılım olmak üzere üç farklı program kullanılmaktadır.



Şekil 3.1. İlaç tasarımında EMRE yazılım sisteminin basamakları.

3.5.1. EMRE Yazılım Sisteminin Geliştirilmesi

Sistem geliştirme aşamalarından programlama, test ve dokümantasyon adımları yazılım geliştirme olarak adlandırılır. Bu çalışmada geliştirilen yazılımlar TÜBİTAK tarafından desteklenen proje ile hizmet alımı şeklinde yapılmış olup yazılım mühendisleri ve programcılar tarafından bizim tarafımızdan belirlenen algoritma mantığı çerçevesinde yazılmıştır. Bu konudaki çalışmaların uzun yıllar devam edeceğini düşünmekteyiz. Çünkü yeni program geliştirme uzun bir süreci kapsamaktadır. Bu yazılımları geliştirmekteki amacımız QSAR çalışmalarında kullanılabilecek yeni programları ülkemize kazandırabilmekdir. Ayrıca hesaplamalarda çok sayıda veri kullanıldığından ve bunların her birinin formatı farklı olduğundan hesaplamaları basit olarak el ile yapmak mümkün değildir. Örneğin sadece ECMC matrislerinin hazırlanması mümkün olmamaktadır. Bu nedenle matris hazırlanması için yeni yazılımlar geliştirilmesi yoluna gidilerek çalışma grubumuz tarafından Delphi 7.0 programlama dilini kullanan EMRE programı geliştirilmiştir.

Çalışma grubumuz tarafından bu amaçla, ileride ayrıntıları açıklanacak programların geliştirilmesi için aşağıda verilmiş olan genel kurallar uygulanmıştır. Yazılım geliştirme (veya programlama) için izlenmesi gereken aşamalar şunlardır:

1. Program analizi;

   A Program amaçlarının saptanması: Çözülmek istenen problemin açık tanımı yapılmalıdır. Yazılması gereken programın hedefleri saptanmalıdır.

   B Girdilerin saptanması: Programa veri olarak neler girileceği belirlenmelidir.

   C Çıktıların saptanması: Program çıktılarının neler olacağı belirlenmelidir. Raporlar, grafikler ekranda görüntülenebilir veya yazıcıdan bastırılabilir.

   D İşlemlerin tanımlanması: Programa girilen verinin çıktıya dönüştürülürken yapılması gereken işlemler tanımlanmalıdır.

   E Program gerçekleştirme yapılabilirliğinin araştırılması: Programın gerçekleştirilmesi için bütçenin durumu, süresi, eleman yeterliliği gözden geçirilmelidir.

   F Analizin belgelenmesi: Programın hedefleri, girdisi, çıktısı ve gereken işlemler, daha sonra başvuru amacıyla belgelenmelidir.

2. Program tasarımı; Program tasarımı, iki aşamadan oluşur.

A ilk aşamada, program mantığı belirlenerek sıradüzen diyagramı (hierachy chart) hazırlanır.

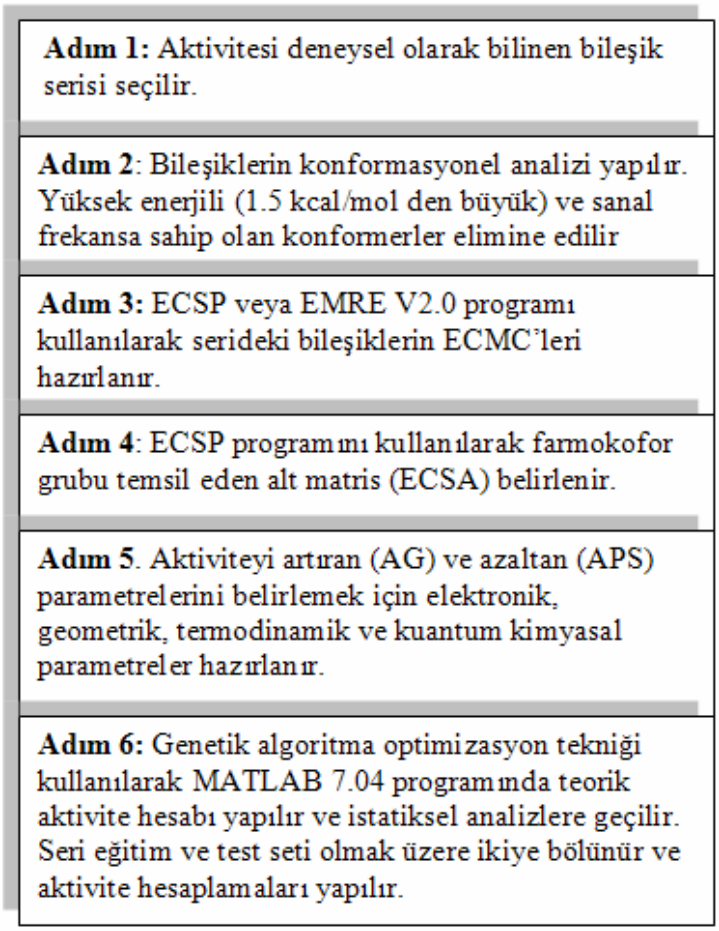
B Daha sonra, ayrıntılı tasarım yapılarak, program, sözdekod (pseudocode) veya akış diyagramı (flowchart) yollarından biriyle açıklanır. Ayrıntılı tasarımda, sırasal, koşul, döngü gibi çeşitli kontrol yapıları kullanılır.

3. Program kodlanması Çoğu kişi, kodlamayı programlama olarak sanırlar. Ancak kodlama, programlama aşamalarından yalnız biridir. Şu aşamalardan oluşur: A Programlama dili seçimi: Yüzlerce programlama dili vardır. Bütün diller her amaca uygun değildir. Bazı diller matematiksel işlemler için, bazıları veritabanı yönetimi için daha uygundur. Bu dillerden, programın amacına uygun olanı veya kurumda yaygın olarak kullanılanı, programlama yapmak için seçilir. B Programın yazılması: Program bilgisayar kullanılarak seçilen programlama dilinin kurallarına göre yazılır. Program, daha sonra üzerinde başka programcıların değişiklik yapmasına olanak tanıyacak şekilde kolay anlaşılır ve açıklayıcı olmalıdır.
4. Program testi Programı yazanlar kodlama sırasında hata yapabilirler. İki tür hata vardır: sözdizimi (syntax) ve mantık (semantic veya logic). A Sözdizimi hatası, seçilen programlama dilinin kurallarına uyulmamasıdır. B Mantık hatası, programın istenen işlevi gerçekleştirmemesi veya yanlış olarak gerçekleştirmesidir. Program testi sırasında program, sözdizimi ve mantık hatalarından temizlenir. Hataların temizlenmesi (debug) için çeşitli yöntemlerden bir ya da bir kaç tanesi kullanılabilir: Masabaşı inceleme: Programcı, program kodlarını kontrol ederek sözdizimi veya mantık hatalarını düzeltmeye çalışır. Ayrıca mantık hatalarını bulmak için, örnek verinin program komutları tarafından nasıl işlendiği el veya hesap makinesi ile takip ederek doğru sonuçlar elde edip etmediğini bulmaya çalışır. Derleyici programı: programlar, Derleyici program tarafından makine diline çevrilip çalıştırılır. Program çalıştırılmadan önce, programın sözdizimi hatasının olmaması gerekir. Derleyici program, sözdizimi hatalarının bulunmasına yardımcı olur. Gerçek veriyle çalıştırma (alfa testi): Programın bütün sözdizimi hataları temizlendikten sonra, mantık hatası olup olmadığı araştırılır. Bunun için çıktısı bilinen girdiler verilerek programın doğru sonuçlara ulaşıp ulaşmadığına bakılır ve mantık hataları düzeltilir.

Beta testi: En sonunda, program bir grup potansiyel kullanıcıya deneme için bir süre kullandırılır. Kullanıcılar, karşılaştıkları sorunları programcılara bildirirler.

5. Program belgeleme ve bakım Belgeleme, programı ve nasıl kullanılacağını anlatan yazılı açıklamalardır. Aslında belgeleme, yazılım bittikten sonra değil, yazılım geliştirmenin bütün aşamalarında yapılır. Bu aşamada, daha önce hazırlanan belgeler gözden geçirilerek son haline getirilir. Belgeleme, programın içinde, programın nasıl çalıştığını ve ne yaptığını anlatan yorum satırları ve kullanım kılavuzu olarak yapılır. Belgeleme, programda daha sonra yapılacak güncellemeler açısından önemlidir.

Bu mantık çerçevesi içinde bu yazlımlar ETM, ECM ve EC-GA metotlarını da kapsamaktadır. Bu yazılımlar kullanılarak bütün bileşik serileri için ECMC matrisleri ve geometrik, elektronik ve fizikokimyasal parametreler hazırlanabilir ve teorik aktivite hesaplamaları yapılabilir. Şekil 3.2'de ise EC-GA metodunda takip edilen aşamalar gösterilmektedir [54, 55].

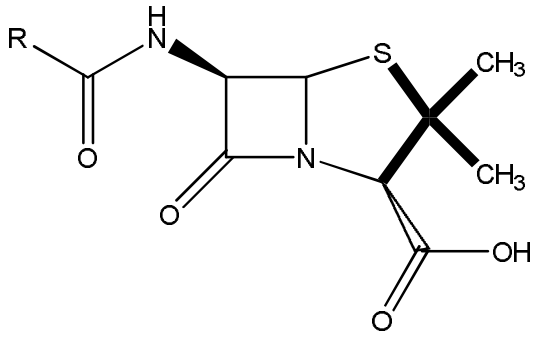


Şekil 3.2. EC-GA metodunda takip edilen aşamalar.

3.5.2. Adım 1

QSAR modelinin geliştirilmesindeki en önemli etken deneysel verilerin doğru ve güvenilir olmasıdır. Deneysel veriler doğru ve kesin olmaması durumunda anlamlı bir model elde etmek ve geliştirmek imkânsızdır. Uygun bir QSAR modelinin geliştirilmesi için bu verilerin benzer kimyasal bileşik serilerinden oluşması gerekmektedir. Bu nedenle bu çalışmada doğruluğuna ve sonuçların güvenilir olduğuna inandığımız penisilin [1] serisi literatürden alınmıştır.

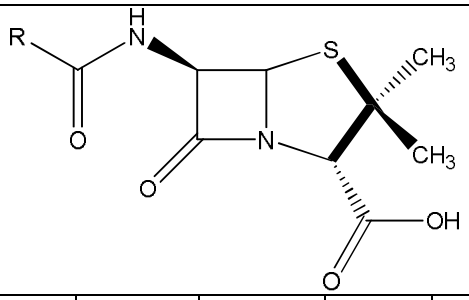
Bu çalışmada ele alınan deneysel aktivitesi bilinen bileşik serisi aşağıda şekil 3.3 ve tablo 3.1 ’de verilmiştir:

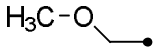
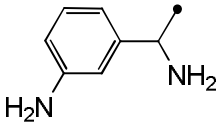
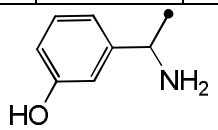
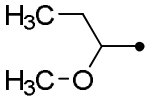
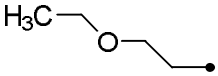
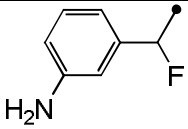
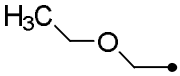
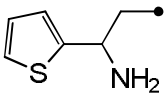
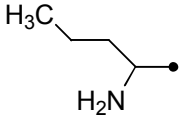
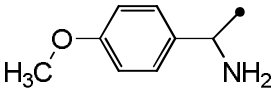
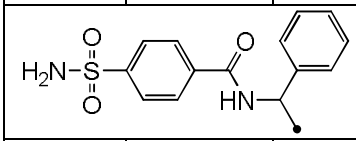
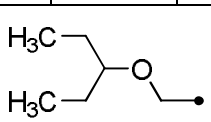
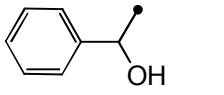
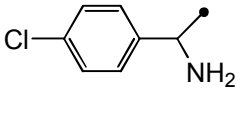
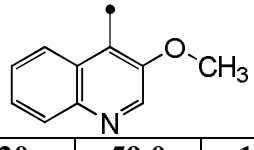
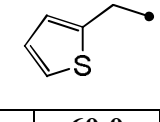
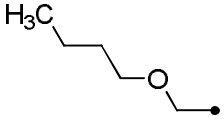
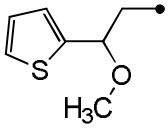
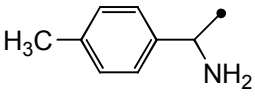


Şekil 3.3. Penisilin (Beta laktam) molekülüne ait temel iskelet.

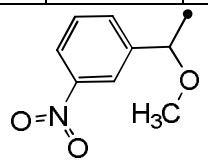
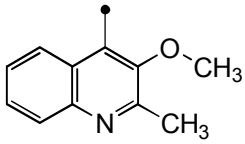
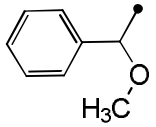
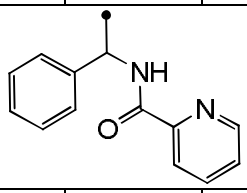
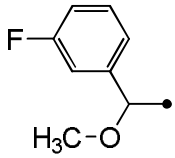
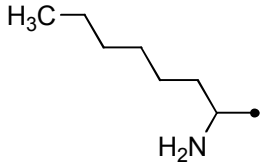
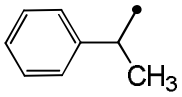
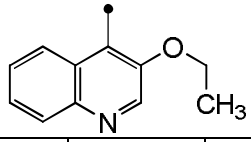
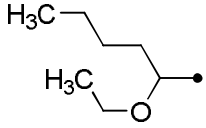
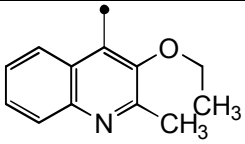
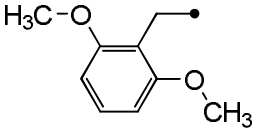
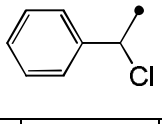
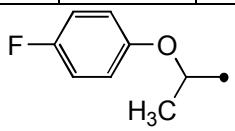
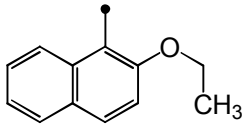
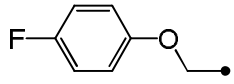
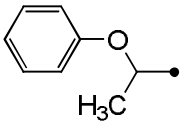
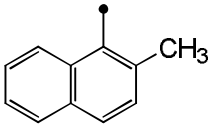
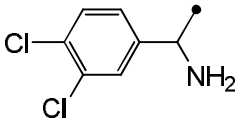
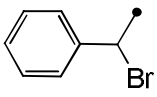
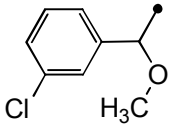
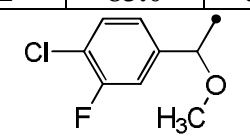
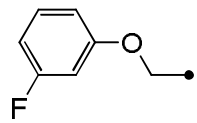
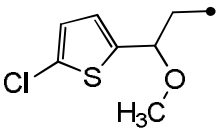
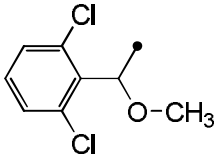
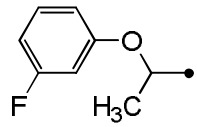
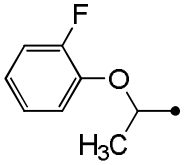
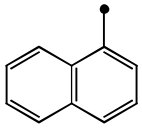
Şekil 3.3’de ele alınan bileşik serisine ait temel iskelet görülmektedir. Tablo 3.1’de Penisilin (Beta laktam) seri için temel iskelete bağlı olan sübstitüentler, hesaplamalarımız sonucu belirlenen konformer sayıları ve deneysel aktivite değerleri yer almaktadır.

Tablo 3.1. Penisilin Serisine Ait Temel İskelet, Sübstitüentler, Deneysel Aktivite Değerleri ve Konformer Sayıları.

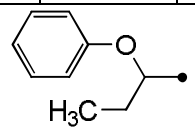
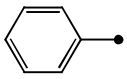
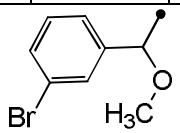
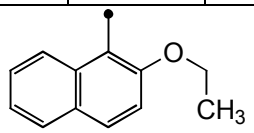
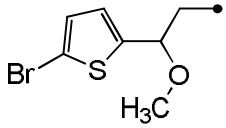
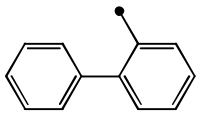
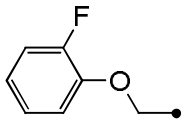
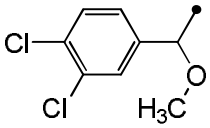
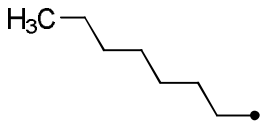
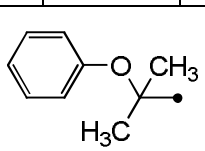
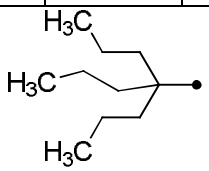
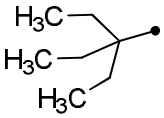
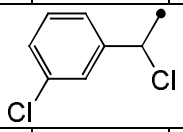
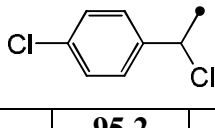
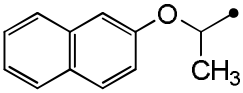
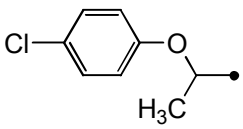
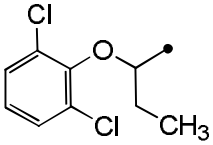
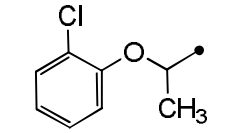
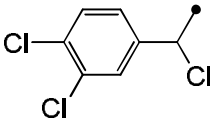
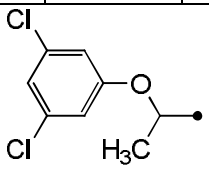
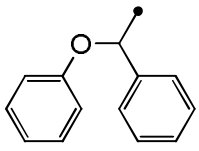
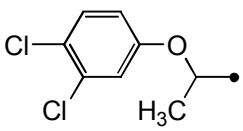
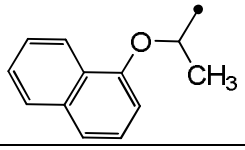
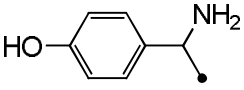


Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.
1	**7.2**	**14**	**2**	**12.0**	**13**	**3**	**15.0**	**8**
								
4	**16.8**	**20**	**5**	**18.0**	**7**	**6**	**20.0**	**13**
			**R = 0**					
7	**25.0**	**15**	**8**	**26.0**	**22**	**9**	**28.0**	**9**
								
10	**32.0**	**12**	**11**	**33.0**	**11**	**12**	**38.0**	**18**
								
13	**42.0**	**18**	**14**	**47.0**	**15**	**15**	**53.2**	**10**
								
16	**55.0**	**14**	**17**	**57.0**	**24**	**18**	**58.0**	**25**
								
19	**58.5**	**19**	**20**	**59.0**	**16**	**21**	**60.0**	**11**
								

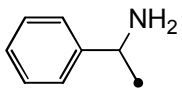
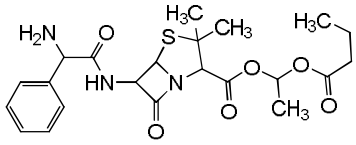
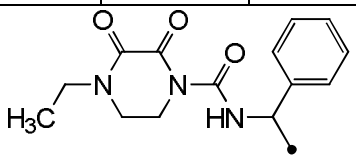
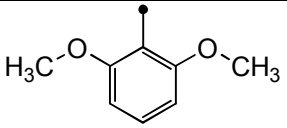
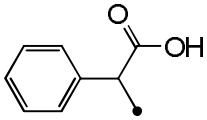
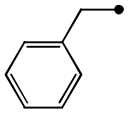
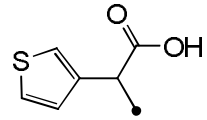
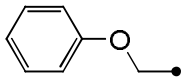
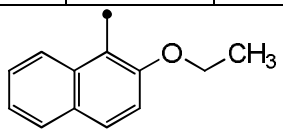
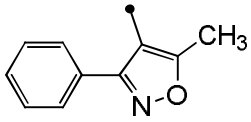
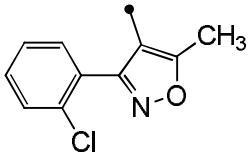
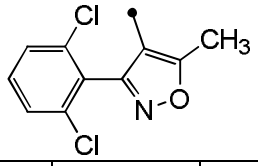
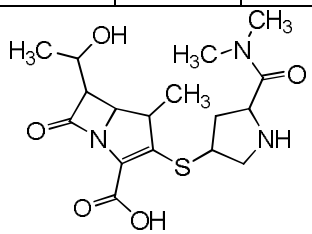
Tablo 3.1’ in devamı

Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.
22	**60.0**	**17**	**23**	**61.7**	**31**	**24**	**62.0**	**8**
								
25	**63.0**	**14**	**26**	**65.0**	**13**	**27**	**66.2**	**9**
								
28	**68.0**	**17**	**29**	**69.7**	**21**	**30**	**74.0**	**17**
								
31	**74.5**	**14**	**32**	**77.0**	**13**	**33**	**78.0**	**10**
								
34	**80.0**	**10**	**35**	**80.0**	**26**	**36**	**81.5**	**16**
								
37	**81.5**	**13**	**38**	**82.1**	**11**	**39**	**82.2**	**17**
								
40	**82.5**	**11**	**41**	**83.0**	**25**	**42**	**83.0**	**32**
								
43	**83.5**	**24**	**44**	**83.6**	**12**	**45**	**84.0**	**12**
								
46	**84.0**	**14**	**47**	**86.0**	**26**	**48**	**86.0**	**9**
								

Tablo 3.1’ in devamı

Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.
49	**86.1**	**13**	**50**	**87.0**	**8**	**51**	**88.0**	**9**
								
52	**89.3**	**8**	**53**	**89.7**	**11**	**54**	**91.0**	**9**
								
55	**91.5**	**10**	**56**	**92.0**	**10**	**57**	**92.4**	**13**
								
58	**92.5**	**12**	**59**	**93.3**	**15**	**60**	**93.6**	**20**
								
61	**94.0**	**24**	**62**	**94.0**	**19**	**63**	**94.7**	**9**
								
64	**94.8**	**8**	**65**	**95.2**	**12**	**66**	**95.6**	**13**
								
67	**96.0**	**11**	**68**	**96.0**	**12**	**69**	**96.5**	**14**
								
70	**97.0**	**21**	**71**	**97.0**	**13**	**72**	**97.2**	**9**
								
73	**97.4**	**15**	**74**	**97.4**	**8**	**75**	**18.0**	**14**
								

Tablo 3.1' in devamı

Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.	Bil. No	PAB_{den}	Konf.
76	18.0	15	77	20.0	8	78	30.0	14
								
79	39.0	18	80	50.0	3	81	60.0	13
								
82	65.0	14	83	80.0	11	84	89.0	4
								
85	92.0	9	86	95.0	15	87	96.0	12
								
88*	2.0	12						
								

* ile simgelenen 88'nolu bileşik Beta laktam antibiyotiklerin Karbapenem sınıfı antibiyotiklerinden olan Meropenem'dir.

3.5.3. Adım 2

Çalışmanın bu basamağında Spartan paket programında molekül iskeletlerinin çizimi yapılıp**,** iskelet dinamizmine göre en düşük enerji halleri minimize edilerek molekül yapıları çıkartılmıştır. Semi emprik PM3 metoduyla konformerler belirlenerek kuantum kimya hesaplamaları yapılmıştır. Konformasyonel analizi ve kuantum kimyasal hesaplamaları yapılarak yüksek enerjili (1.5 kcal/mol den büyük olan) ve sanal frekansa sahip konformerler elimine edilmiştir. Çünkü sanal frekansın bulunması molekülün geçiş haline karşılık gelmektedir ve molekülün tam optimize edilemediği anlamına gelir. Daha sonra moleküller üst üste çakıştırılarak molekülün aynı yapıya sahip yüksek enerjili konformerleri silinir. Boltzmann dağılımına göre oda sıcaklığında düşük enerjili

konformerler seçilir. Düşük enerjili konformerlerin oda sıcaklığında bulunma ihtimalleri daha yüksektir. Düşük enerjili konformerler ile yüksek enerjili konformerler arasında bileşiklere bağlı olarak 1/1000 gibi dağılım farkı ortaya çıkmaktadır. Spartan paket programının kullanıldığı bu aşamadan sonra, çalışma grubumuz tarafından yazılan ve geliştirilmesi devam eden programlar devreye girecektir ve bundan sonraki aşamalarda bu programlar kullanılacaktır.

3.5.3.1. Boltzman Dağılımı

Boltzman dağılımı; farklı enerjili hallerin popülâsyonlarını hesaplamak için ortaya atılan bir formüldür. Boltzmann dağılımının bilinmesi gereken en önemli özelliklerinden biri popülâsyon dağılımlarının enerji ile sıcaklığın üstel bir fonksiyonu olduğudur. Bu formül E_i ve E_j hallerdeki her bir bileşiğin konformer sayılarının oranı olan N_i/N_j büyüklüğünü

$$N_i/N_j=e^{-(E_i-E_j)/kT} \tag{3.9}$$

şeklinde tanımlar. Boltzmann dağılımının yorumlanmasına ilişkin önemli bir nokta, bu ifadenin seviyelerin değil enerji hallerinin popülâsyonlarına işaret ettiğidir [72]. Boltzmann dağılımı dikkate alınarak düşük enerjili konformerler seçilmiş ve diğer konformerler elimine edilmiştir. Ancak relatif enerji değerleri ile boltzman dağılımları karşılaştırılırsa, relatif enerji değeri azaldıkça boltzman dağılımı değerlerinin arttığı (boltzman dağılımının yüksek olması oda sıcaklığında düşük enerjili yapıların bulunma ihtimalinin daha fazla olduğunu gösterir) görülmüştür. Tablo 3.2'de Penisilin serisi için konformerler, her bir konformerin relatif enerjisi ve bu enerjilere karşılık gelen Boltzmann değerleri verilmiştir. Tabloda görüldüğü gibi konformerlerin bağıl enerjileri arttıkça Boltzmann dağılımı değerleri azalmaktadır.

Tablo 3.2. Penisilin serisinde 73. bileşiğin 15 konformeri için bağıl enerji ve boltzmann değerleri.

Konformer	rel. E (kcal/mol)	Boltzmann Dağılımı
73_1	0.0000	0.1889
73_2	0.0416	0.1761
73_3	0.2210	0.1301
73_4	0.3169	0.1106
73_5	0.4281	0.0917
73_6	0.4994	0.0813
73_7	1.0130	0.0342
73_8	1.0537	0.0319

Tablo 3.2'nin devamı

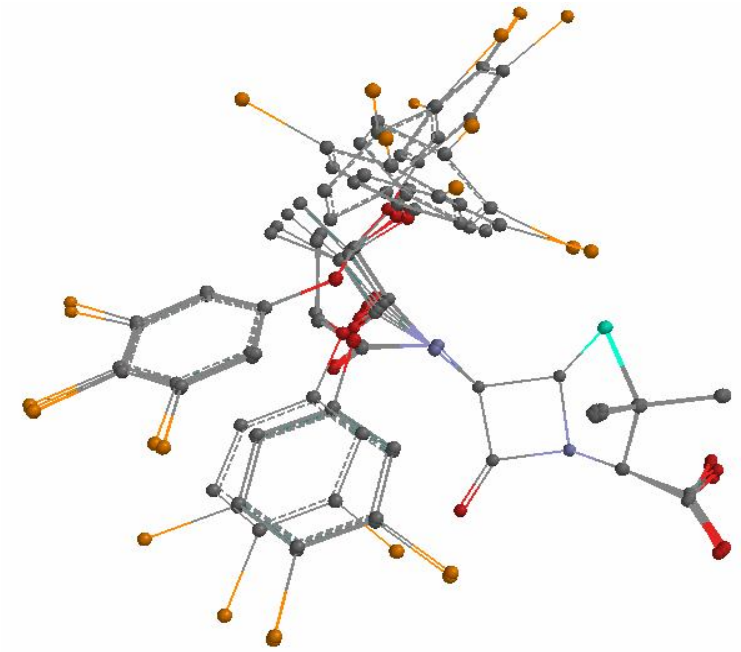
Konformer	rel. E (kcal/mol)	Boltzmann Dağılımı
73_9	1.1026	0.0294
73_10	1.1926	0.0252
73_11	1.3015	0.0210
73_12	1.3025	0.0210
73_13	1.3271	0.0201
73_14	1.3469	0.0195
73_15	1.3578	0.0191

Relatif enerjisi 1.5 kcal/mol den büyük olanlarda bu dağılım oldukça küçük (yaklaşık 10^{-4}-10^{-5}) olmaktadır. Literatürde verilen bilgilere göre [61] yaklaşık olarak 1.5 kcal/mol den büyük olan konformerler ihmal edilmiştir. Kalan konformerler için kendi aralarında bazı atomlar üzerinden üst üste çakıştırma işlemi (align) yapılmış ve birbirinin aynı ya da yakın enerjili olan birden fazla konformer varsa yüksek enerjili konformerler silinmiştir.

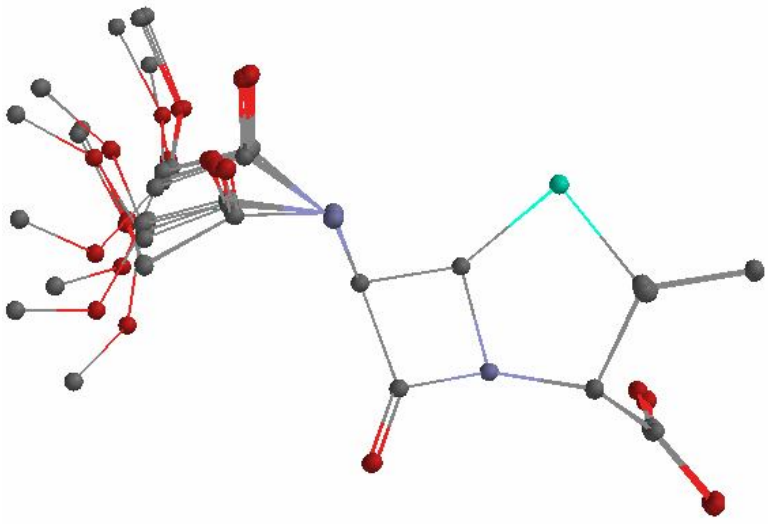
Bu çalışmada *.txt dosyalarındaki bilgiler içerisinden atomların yükleri ve kartezyen koordinatları, moleküle ait enerjiler elektron konformasyonel matrislere dönüştürüldü. Bu matris değerleri, moleküllerin karşılaştırma ile aktif fragmentinin (ya da Pha) bulunması için kullanıldı.

Moleküler özelliklerin (sterik durum, elektrostatik potansiyel ve hidrofobisite gibi) üç boyutta göz önüne alınması, aktif ve inaktif bileşikler arasındaki benzerlikler ve farklılıkları belirlemek için ilk adımdır. Bir bileşik serisi için genel olan üç boyutlu yapısal özellikleri ortaya çıkarmak için en kolay yol üst üste çakıştırma (süperimpozisyon) işleminin uygulanmasıdır. Topolojik analoglar ele alındığında üst üste çakıştırma işlemi bellidir. Fakat kimyasal olarak ilişkili olmayan yapılar karşılaştırıldığında bu açık değildir. Üç boyutta ortak stereokimyasal özellikleri paylaşan farklı moleküllerin nasıl benzer biyolojik aktivite ortaya koyduklarını gösteren çeşitli terapötik alanlarda kapsamlı analizler yapılmıştır. Bu ise yeni kılavuz bileşiklerin rasyonel tasarımında üç boyutlu farmakoforik modellerin gelişmesine yol açmıştır. Molekül yapıları genellikle atomik pozisyonlarının temel alınmasıyla düzene sokulmuştur. Küçük relatif enerjili olan ve üst üste çakışmayan bileşiklerin konformerlerinin Spartan çıktıları txt dosyaları halinde kaydedilmiştir. 88 bileşikli Penisilin serisinde her bir bileşik için konformer hesaplamaları yapılmış yüksek enerjili olanlar ve imajiner frekansa sahip olanlar silinmiştir. Konformer hesaplamalarında bazı

bileşiklerin konformerleri tam olarak optimize olmadığı için bu konformerlere ait titreşim frekansları negatif çıkmıştır. Negatif frekans değerleri ise geçiş halinde söz konusudur. Bu nedenle sanal frekansa sahip olan konformerler de silinmiştir. Bütün bileşikler için çok fazla sayıda konformer elde edilmiştir. Hepsinin gösterimi oldukça uzun olacağından Penisilin (Beta laktam) serisine ait 73. ve 1. bileşikleri için sırasıyla şekil 3.4 ve 3.5'de, konformerlerinin üst üste çakıştırılma işlemi gösterilmiştir.



Şekil 3.4. Penisilin serisinde 73 nolu bileşiğine ait 15 konformerin üst üste çakıştırılması.

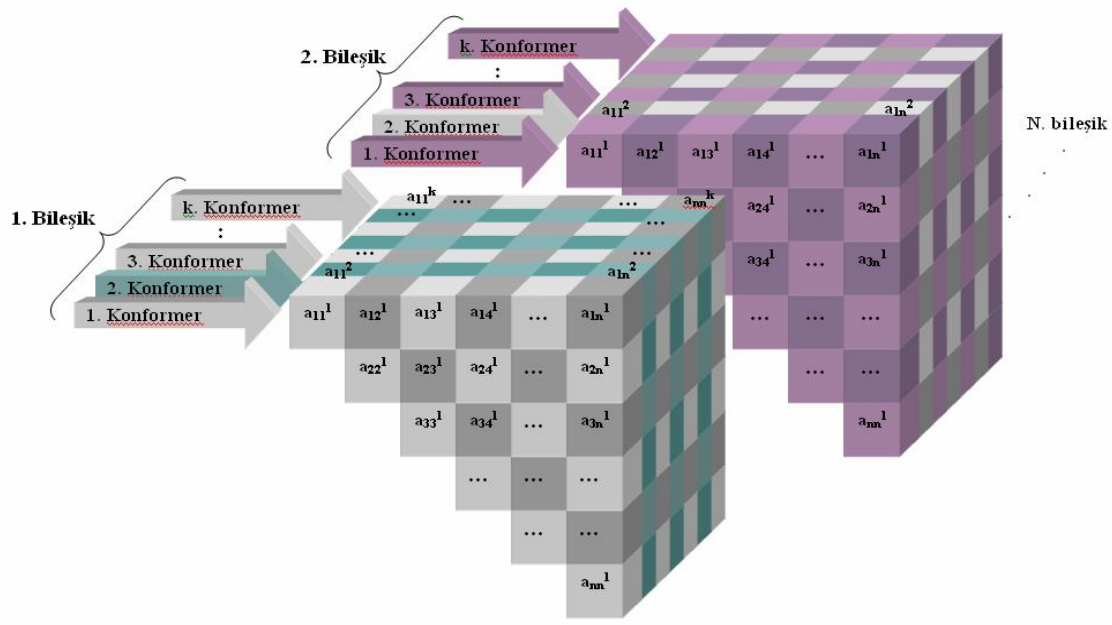


Şekil 3.5. Penisilin serisinde 1 nolu bileşiğine ait 14 konformerin üst üste çakıştırılması.

3.5.4. Adım 3

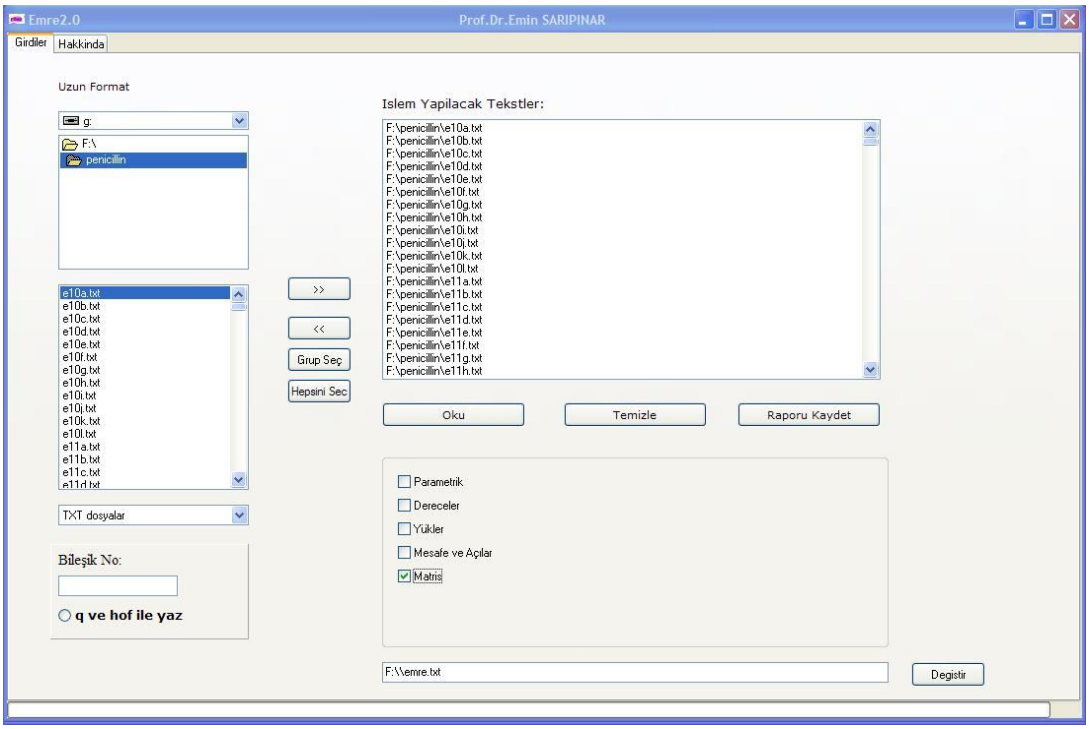
Referans bileşiğin (aktivitesi en yüksek olan bileşiğin en düşük enerjili konformeri referans olarak seçilir) fonksiyonel atomlarına göre serideki bileşiklerin her bir konformeri için ECMC matrisleri hazırlanmıştır. Diyagonal elemanlarına göre simetrik olan bu kare matrisin sadece üst kısmını içeren üçgen matris kullanılır. Matrisin diyagonal elemanları atomların elektronik özelliklerini tanımlayan atomik yüklerden oluşur. Non-diyagonal elemanlar (a_{ij}) ise iki çeşittir. Birincisinde, i ve j ile ifade edilen atomlar birbirine kimyasal olarak bağlı ise a_{ij} elemanı bağ derecesini gösterir. İkincisinde ise, i ve j bağ yapmamış atomlar ise bu eleman atomlar arası mesafeyi verir. Böylece her bir konformer bütün özelliklerinin elektronik ve geometrik olarak temsil edildiği kabul edilir. Bunları hazırlamak için EMRE programı kullanılmıştır.

Şekil 3.6'da ECMC üç boyutlu olarak gösterilmiştir. Her bir matris küçük alt birimlerden oluşmuştur ve bu alt birimler moleküldeki atomların atomik özelliklerini ya da atomlar arasındaki özellikleri temsil etmektedir. Matrislerdeki sütun/satır sayıları moleküldeki atom sayısına (n) karşılık gelmektedir. Matrislerin dikey tabakalardan oluştuğu düşünülürse her bir tabaka bileşiğin farklı bir konformerine karşılık gelmektedir. Böylece her bir bileşik için o bileşiğin konformer sayısı (k) kadar tabaka olacak ve her bir tabakada moleküldeki atomlara ait özellikler yer alacaktır. Bu şekilde serideki her bir bileşik için benzer bir matris oluşturulmuştur.



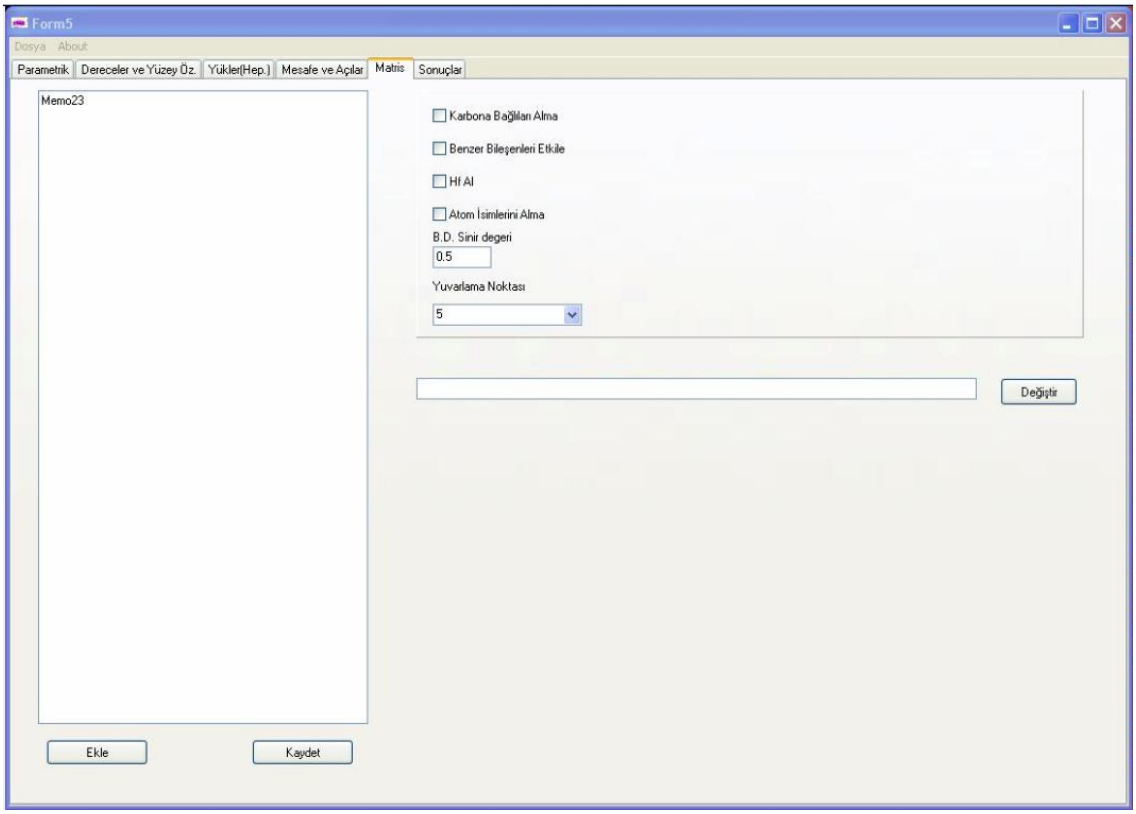
Şekil 3.6. Üç boyutlu ECMC'nin gösterimi.

EMRE programına ait girdiler sekmesi şekil 3.7'de gösterilmektedir.



Şekil 3.7. EMRE programına ait girdiler sekmesi.

EMRE programına ait matris sekmesi şekil 3.8'de gösterilmektedir.



Şekil 3.8. EMRE programına ait matris hazırlama sekmesi.

3.5.5. Adım 4

Farmakofor (Pha) grubun bulunması, çoğunlukla, kesin bir biyolojik aktivite sağlamak zorunda olan bileşiğin, kimyasal, yapısal ve fiziksel özelliklerinin kombinasyonlarını kapsayan bir grup özelliği olarak ifade edilir. Pha tanımlayan grup ECSA alt-matrisi ile gösterilir. Adım 3'te olduğu gibi; hazırlanan bu matrisler kullanılarak ECSP adını verdiğimiz program ile Pha'yı temsil eden alt matris belirlenmiştir. ECSP programı ile alt matris belirleme işleminde süreyi en aza indirmek ve boyutu çok büyük olan matrislerle hesaplama yapabilmek için program paralel hale getirilmiştir. Literatür çalışmaları incelendiğinde bileşiklerin üç boyutlu matris ile belirtilmesi çalışmaları ve farmakofor grubun bulunması ile ilgili çalışmalar 1991 yılında aynı çalışma grubunda olan A.S. Dimoglu ve I.B. Bersuker tarafından gerçekleştirilmiştir [57]. Daha sonraları bu kişiler çalışmalarına ayrı devam etmişlerdir. Elektron Topolojik Metot (ETM) ile çalışan A.S. Dimoglu ve grubu moleküler mekanik yöntemle bileşiklerin Z matrisini çıkararak eski programlardan olan CNDO yöntemiyle elektronik ve geometrik hesaplamalar yapmakta daha sonra bileşiklerin sadece tek konformeri için üç boyutlu matris oluşturarak pha'yı bulmaktadır. Bunun için tolerans değerlerini tek tek girilerek aylarca süren hesaplamalar yapılmaktadır. ETM yöntemi ile hesaplama yapılan programların dos ortamında ve PC 286 makinelerde çalıştığı bilinmektedir. ETM yönteminde sadece mulliken yükleri kullanılırken bizim programımızda istenildiği takdirde mulliken, elektrostatik ve atomik valens yüklerine göre matrisler oluşturulup hesaplamalar yapılabilmektedir. I.B. Bersuker ise bileşiklerin konformerlerinide hesaba katarak üç boyutlu matrislerini kullanarak farmakofor grubunu belirlemiştir. Bunun için referans bileşiğine göre diğer tüm konformerler % bağıl sapma olarak hesaplanmıştır. Çalışmalarında sonuçları tolerans değerleri olarak vermiştir. Bu konuda yaptığı çalışmalardan 2008 tarihli yayınında ELCOME isimli bir programdan söz etmekte fakat ayrıntılı olarak bilgi vermemektedir. Bu nedenle ayrıntıları literatürde verildiği kadarıyla bilmekteyiz. Grubumuz tarafından geliştirilen ECSP programı ise bileşiklerin en düşük enerjili tek konformerini hesaba kattığı gibi bütün konformerlerinide dikkate alarak hesaplama yapabilmektedir. Ayrıca hesaplamaları belirli tolerans değerleri arasında tarama yaptığı gibi % bağıl sapma olarak veya her ikisinin karışımını dikkate alarak hesaplama yapabilmektedir. Bizim program istenilen bileşik serisi için çok rahatlıkla uygulanabilmektedir. ECSP programında alt matris belirleme işlemi

yapılırken her bileşik için spartan atomlara farklı numara vermektedir ve her bileşik için atomların numaralandırmasına karşılık gelen atm dosyası oluşturulmuştur. Bu dosya en basit yapılı bileşiğe göre oluşturulmuştur. Şekil 3.9'da bileşiklerdeki her bir atomun karşılaştırılarak yazılmasıyla oluşan ve atm adı verilen dosyadan örnek gösterilmiştir. Görüldüğü gibi serideki her bir bileşiğin temel iskeletindeki atomlar bu dosyada üst üste çakıştırılarak yazılmıştır. Bu dosya alt matris hesaplaması yapılırken atomların karşılaştırılmasında programın çalışma hızını artırmakta ve çalışma grubumuza oldukça büyük bir kolaylık sağlamaktadır.

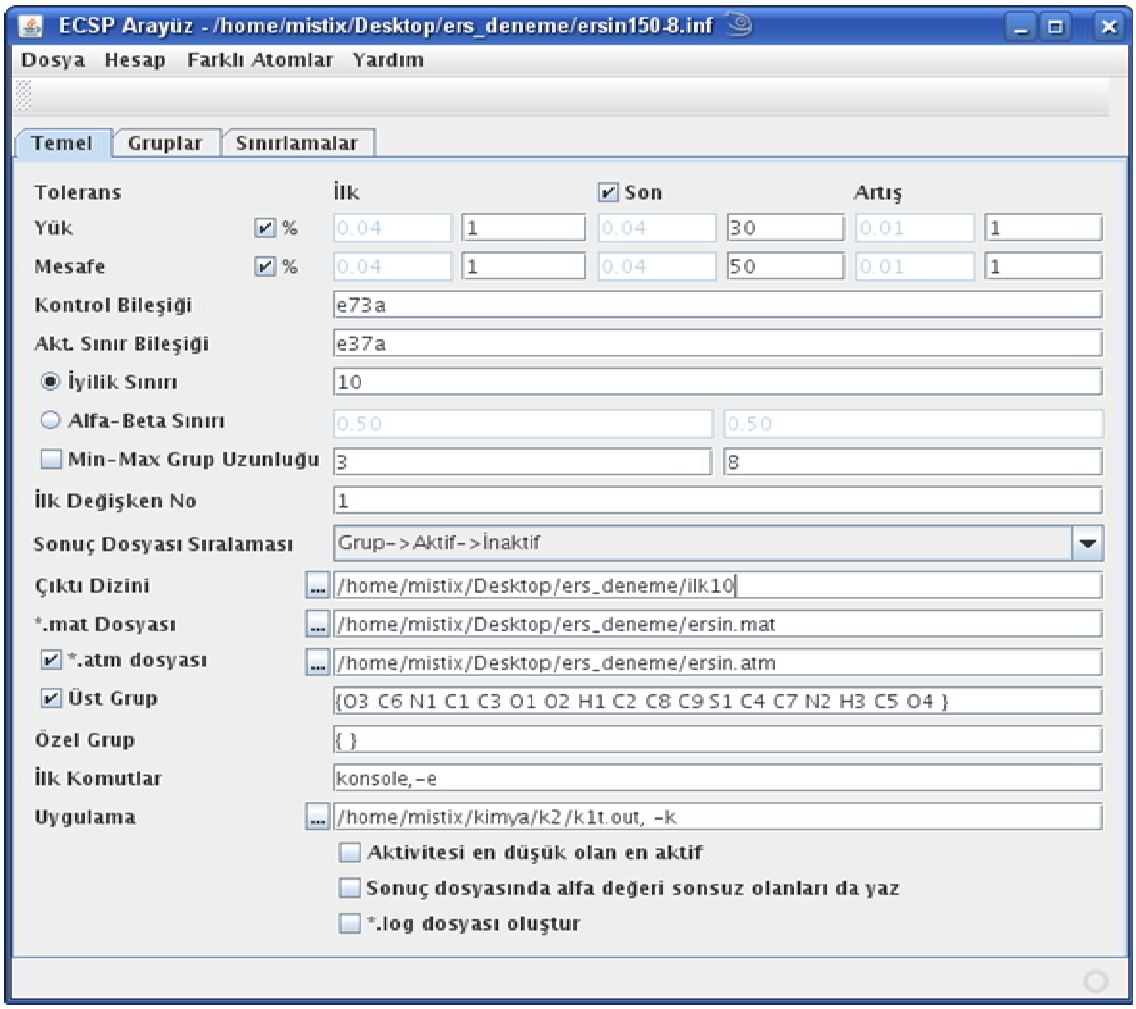
1 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
2 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
3 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
4 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
5 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4
6 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
7 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
8 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
9 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
10 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
11 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
12 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
13 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
14 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
15 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
16 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
17 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C11
18 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
19 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
20 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
21 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
22 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
23 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C13
24 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
25 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
26 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
27 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
28 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
29 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C19
30 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
31 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C19
32 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
33 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
34 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
35 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C11
36 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10

37 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
38 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C11
39 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
40 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
41 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
42 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
43 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
44 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
45 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
46 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
47 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
48 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C11
49 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
50 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
51 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
52 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C11
53 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
54 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
55 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
56 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
57 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
58 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
59 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
60 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
61 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
62 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
63 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
64 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
65 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
66 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
67 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
68 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
69 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
70 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
71 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
72 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
73 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
74 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
75 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
76 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
77 O3 C6 N1 C1 C3 O1 O2 C17 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
78 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
79 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
80 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
81 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
82 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
83 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
84 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C16
85 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10

86 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
87 O3 C6 N1 C1 C3 O1 O2 H1 C2 C8 C9 S1 C4 C7 N2 H3 C5 O4 C10
88 O3 C6 N1 C1 C3 O2 O1 H1 C2 S2 - C13 C4 C7 C5 C17 O4 H24 -

Şekil 3.9. Penisilin serisi için en basit bileşiğe göre diğer bileşiklerin fonksiyonel grupları ve atomlarının düzenlenmesi yani moleküllerde birbirine karşılık gelen atomlar.

Alt matris belirleme işlemi için grubumuz tarafından geliştirilen ECSP programına ait ara yüzü şekil 3.10'da verilmektedir.



Şekil 3.10. ECSP programına ait arayüz.

Programın arayüzünde bulunan sekmelerin açıklaması:

- **İlk / Son:** Yük ve mesafe için başlangıç ve bitiş tolerans değerleri
- **Artış:** Tolerans değerlerinde ne kadarlık artışlarla tarama yapılacağı
- ***.mat:** Konformerlere ait matrislerin bulunduğu dosya
- **Çıktı Dizini:** Çıktıların kaydedileceği dizin

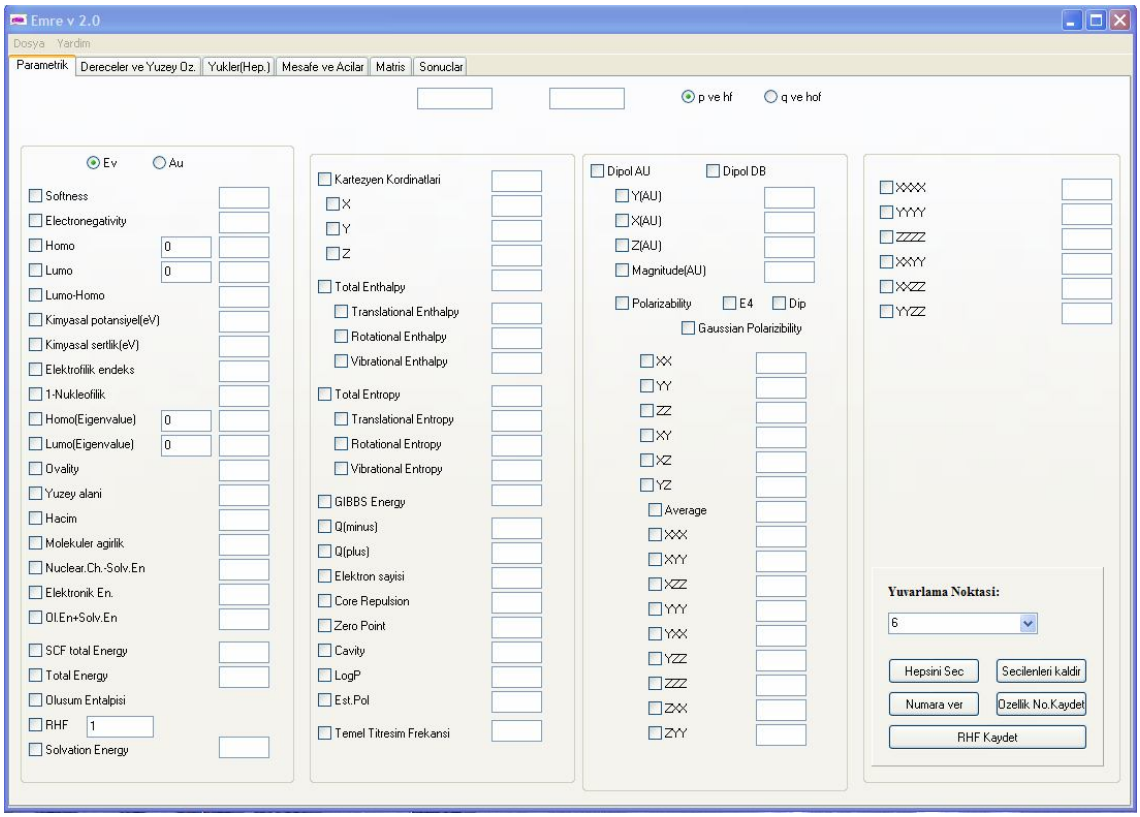
- **Kontrol Bileşiği**: En yüksek aktiviteye sahip bileşiğin en düşük enerjili konformeri
- **Akt. Sınır bileşiği**: En düşük aktiviteye sahip bileşiğin en yüksek enerjili konformeri
- **İyilik Sınırı:** Serideki bütün konformerler içerisinde Pha'yı içeren minimum konformer sayısı
- **Üst grup:** Bütün bileşiklerde ortak olan yapı
- Programın sol üst kısmında yer alan "gruplar" bölümü ise, Pha olduğunu düşündüğümüz grupları analiz ermek istediğimiz zaman oldukça kullanışlıdır. Verilen tolerans değerleri aralığında tarama yaptıktan sonra hesaplama sonuçlarını analiz ederek belirtilen gruplar için istatistiksel değerleri bu aralıklar dâhilinde verir. Buna göre en uygun Pha belirlenir.
- İstenilen gruplara ait alt matrisi, bu grupların hangi konformerlerde bulunduğunu ve konformerlerin kaç tanesinde bu grubun bulunduğu hakkında bilgi edinebiliriz.

Bersuker ve arkadaşları tarafından yapılan çalışmada [73] izlenen adımların bir kısmı bizim çalışmamızda da izlenmiştir. Bu işlem için önce aktivitesi en yüksek olan ilk 10 bileşik seçilir. İlk 10 bileşik için en düşük enerjili konformerler esas alınarak belirli tolerans değerleri içerisinde karşılaştırma yapılmıştır. Burada bulunan ortak en büyük matris atomları üst grup olarak belirlenir. Daha sonra moleküller belirli tolerans değerleri içinde veya en küçük bağıl sapmaya göre köşegen matris elemanları (yükler) karşılaştırılır. Bu karşılaştırmada iki matrisin belirli tolerans değerlerinde birbirlerine benzeyen alt matrisleri aranmaktadır. Bunun için her iki matrisin 3x3 lük bütün alt matrisleri, 4x4 lük bütün alt matrisleri, ... ,nxn lik bütün alt matrisleri karşılaştırılmalıdır. Yani 21x21 lik bir referans matris ile 20x20' lik bir matrisin benzer alt matrisleri aranırken, en az $P(21,3)xP(20,3)+P(21,4)xP(20,4)+...+P(21,20)xP(20,20)=1.98x10^{38}$ adet karşılaştırma yapılmalıdır. Ancak gereksiz eleman karşılaştırmasının önüne geçebilecek bir yapı ve algoritma tasarlanmıştır.

Referansın her bir atomunun yükü, molekülün her bir atomunun yükü ile karşılaştırılır. Yükler farkı tolerans değerinden büyük olan atomlar elenir. Bir sonraki aşamada sadece yük testinden geçen atomların mesafeleri karşılaştırılacağından önemli bir eleme gerçekleştirilmiş olur. Bu işlemden sonra diğer moleküllerin en düşük enerjili konformerlerinin ECMC'lerinden bir alt matris (Pha) belirlemek için referans bileşiğin en düşük enerjili konformerine ait matris ile karşılaştırmalar yapılır.

3.5.6. Adım 5

Alt matrisin belirlenmesinden sonra moleküllerde aktiviteye etki eden gruplar, serideki bileşiklerin temel iskeletine bağlı olan sübstituentlerin etkisine göre araştırılmıştır. Aktiviteyi azaltan veya artıran elektronik, geometrik ve fizikokimyasal gibi özellikleri belirlemek için APS ve AG değerlerini kantitatif olarak karakterize eden S fonksiyonu eşitlik 3.2'de formülize edilmiştir. Serideki tüm bileşikler için farklı $a_{ni}^{(j)}$ değişkenleri belirlendikten sonra, txt dosyalarından aktivite hesaplama formülündeki formata uygun hale getirilebilmesi için, Delphi 7.0 programlama dili kullanılarak yazılan yeni bir program (EMRE) ile aktivite üzerinde etkisi olduğu tahmin edilen bu değişkenler düzenlenir. Şekil 3.7'de EMRE programına ait girdiler sekmesi gösterilmiştir. Şekil 3.11'de EMRE programına ait parametrik sekmesi gösterilmiştir. İlgili bileşik serisinin hesaplanmasında Parametrik sekmesinde hangi özellikler kullanılacaksa ilgili özellikler buradan seçilmektedir.



Şekil 3.11. Emre programına ait parametrik sekmesi.

3.5.6.1. Moleküler Parametreler

Bu çalışmada kullanılan elektronik, geometrik ve fizikokimyasal parametreler kısaca açıklanmıştır.

HOMO (highest occupied molecular orbital) bir moleküldeki en yüksek enerjili dolu orbitali, LUMO (lowest unoccupied molecular orbital) ise bir moleküldeki en düşük enerjili boş orbitali belirtir. ε_{HOMO} ve ε_{LUMO} şeklinde gösterilir [74].

Hardness: Hardness (η) moleküllerin dayanıklılığıdır ve lumo ile homo enerjileri arasındaki farkın yarısına eşittir. Sonuçlar au (atomic unite) cinsinden elde edilir. Şu şekilde gösterilir.

$$\eta \approx \frac{1}{2}(I - A) \approx \frac{1}{2}(\varepsilon_{\mathrm{LUMO}} - \varepsilon_{\mathrm{HOMO}}) \tag{3.5}$$

$$I \approx -\varepsilon_{\mathrm{HOMO}} \quad \text{ve} \quad A \approx -\varepsilon_{\mathrm{LUMO}} \tag{3.6}$$

Formuldeki *I* iyonizasyon potansiyeli, *A* ise elektron afinitesini göstermektedir.

Sistemden bir elektron çıkarıldığı zamanki enerji değişimine iyonizasyon potansiyeli (I), bir elektron eklendiği zamanki enerji varyasyonuna ise elektron afinitesi (E) denir [74].

Elektoron afinitesi: Bir ligantın bir donör atomdan tam olarak bir elektron alma yatkınlığıdır. eigenvalue LUMO nun ters işaretlisidir.

İyonizasyon potansiyeli: eigenvalue HOMO nun ters işaretlisidir.

Kimyasal sertlik: Hardness ile aynı formulle hesaplanır fakat sonuç eV (electron volt) cinsinden bulunur.

Softness: Softness kimyasal reaktivitenin derecesini ölçen moleküllerin bir özelliğidir**.** Ayrıca Hardness in tersidir. S şeklinde gösterilir**.**

$$S = \frac{1}{\eta} \tag{3.7}$$

Kimyasal potansiyel: μ şeklinde gösterilir ve aşağıdaki formülle hesaplanır [74].

$$\mu \approx -\frac{1}{2}(I + A) \approx \frac{1}{2}(\varepsilon_{\mathrm{HOMO}} + \varepsilon_{\mathrm{LUMO}}) \tag{3.8}$$

Elektronegatiflik: Bir molekülde ki atomun elektronları kendine çekme kapasitesidir. Kimyasal potansiyelin zıt işaretlisidir. χ şeklinde gösterilir.

$$\chi = -\mu \qquad \chi = \frac{I + A}{2} \tag{3.9}$$

Elektrofilik indeks: Donör ve akseptör arasındaki elektron alışverişi sırasında oluşan enerji azalmasının bir ölçüsüdür. (ω^+)şeklinde gösterilir ve aşağıdaki formülle hesaplanır [74].

$$\omega^+ = \frac{\mu^2}{2\mu} \tag{3.10}$$

Nükleofilik indeks: Elektrofilik indeksin tersidir. $\omega^- = 1/\omega^+$ ile formülize edilir.

Ovality: Bir molekülün nasıl küre yada silindirik şekilde olacağının bir ölçüsüdür.

Hacim ve alan oranına bakılarak aşağıdaki şekilde tanımlanır.

$$O = A / (4*Pi* ((3*V) / 4*Pi) \wedge (2/3)) \tag{3.11}$$

A: alan, V: hacim, O: ovality i ifade etmektedir.

Bağ dipol moment: Bir moleküldeki kimyasal bağın polarlığının bir ölçüsüdür. Bağ dipolu (μ) şu şekilde verilir:

$$\mu = \delta\, d \tag{3.12}$$

$+\delta - \delta-$ şeklinde gösterilen bağ dipolunda ki d kısmi pozitif ve kısmi negatif yükler arasındaki uzaklığı verir. $+\delta - \delta-$ ise bağ ile paralel olan ve pozitiften negatife giden bir vektör demektir.

Nükleer yüklerin bir setteki kuantum mekanik dalga fonksiyonunun dağılması 3 yolla mümkündür. Bunlar mulliken yükler, natural yükler ve elektrostatic yükler dir.

1. Mulliken yük standart mulliken analizini kullanarak iki atom merkezinin yoğunluk/dalga fonksiyonunu alır ikiye böler, böylece herbir bileşen atomundaki elektron bulutunun yarısı alınmış olur.

2. Natural yükler mulliken yüklere benzerdir fakat büyük setler için matematiksel hesaplamalarla daha iyi sonuçlar elde edilir. Natural ve mulliken yükler herbir atomun yükünün kimyasal reaktivitesini saptamak için kullanılan en iyi yöntemlerdir.

3. Elektrostatik yük bütün dalga boylarının meydana getirdiği elektrostatik alandan yüklerin oluşturulduğu sayısal bir metottur. Molekül ya da atomun elektrostatik potansiyelinin yarısı söz konusu olduğunda elektrostatik yükü kullanmak daha uygundur.

Polarizability: Polarizability yük dağılımının bağıl bir ölçüsüdür. Yani bir atom yada moleküldeki elektron bulutudur ki bu en dış elektrik alanla molekül yada atomun biçiminin bozularak iyon yada dipole yakın bir görünüme sahip olmasıyla gerçekleşir.

α ile gösterilen elektronik polarizability bir atomun indüklenmiş dipol momentinin (P), bu dipol momenti oluşturan elektrik alana (E) oranıdır [75].

$$P = \alpha E \tag{3.13}$$

Bağ açısı: Bir moleküldeki üç atom arasındaki açıdır.

Torsiyon açısı: Bir molekülde dört atomdan oluşan iki düzlem arasındaki açıdır.

Atomlar arası mesafe ve bağ mesafesi: Atomlar arası mesafe atomlar arası denge mesafesidir ve itici ve çekici güçler arasında bir denge ile oluşur. Bağ mesafesi ise kovalent olarak bağlı iki atomun çekirdekleri arasındaki mesafedir

Bağ derecesi: iki atom arasındaki bağ sayısı olup bağ derecesi kaçlı bağ olduğunu gösterir. Diğer bir deyişle, iki atom 2 elektron paylaşırsa tek bağ, 4 elektron paylaşırsa çift bağ, 6 elektron paylaşırsa üçlü bağ meydana gelir. Bağ derecesi arttıkça bağ uzunluğu azalır.

$$\textbf{Bağ derecesi} = \frac{\textbf{Bağ MO deki é sayısı toplamı} - \textbf{Karşıbağ MO é sayısı toplamı}}{\textbf{2}} \tag{3.14}$$

Elektronlar, kovalent bağlarda atomları bir arada tutar. Bağ derecesinin yüksekliği, daha çok elektronun bulunduğunu ve böylece atomların daha sıkıca bir arada tutulduğunu gösterir.

Dik uzaklık: Bir moleküldeki en uzak atomun seçilen belli bir düzleme (farmakofor düzlemi) dik uzaklığıdır.

LogP: Yüklü ya da yüksüz küçük organik moleküllerdeki oktanol-su çözünme katsayısıdır. Oktanol-su çözünme katsayısı iki çözücü olan su ve oktanoldaki diferansiyel çözünürlüğün bir ölçüsüdür. Bu ise bir maddenin bağıl hidrifobisite ve hidrofilisitesinin bir ölçüsüdür.

İlaç benzeri maddelerin içeriğindeki moleküler büyüklük ile birlikte, dönen bağların sayısı ve hidrojen bağlarına yatkınlık, LogP değeri; apsorpsiyon, biyoyararlılık, hidrofobik ilaç reseptör etkileşimi metabolizma ve toksisite ile ilişkilidir.

$a_{ni}^{(j)}$ parametresinin yani değişkenlerin seçimi önemlidir. Çünkü aktiviteye etki eden grup özelliklerinin elde edilmesini sağlar. Aktivite üzerinde etkisi olduğu tahmin edilen ve QSAR çalışmalarında kullanılan bu parametreler, EMRE programı kullanılarak, S fonksiyonundaki formata uygun şekilde hazırlanır. EMRE programı, serideki bileşiklerin toplamda yüzlerce konformerinin olduğu düşünülürse bu konformerlerin her biri için oluşturulması gereken çok sayıda değişkenin, yaklaşık 2400–2500 sayfalık çıktılar halinde bulunan .txt dosyalarından yerlerinin tespit edilip istenilen formatta elle yazılması mümkün olmamasından dolayı geliştirilmiştir. Ayrıca bu programda .txt dosyalarında çıktı olarak bulunmayan ve aktivite hesaplamalarında çok büyük önem taşıyan torsiyon açısı, farmakofor olarak belirlenen düzleme diğer atomların dik uzaklığının hesaplanması ve en uzak atomların farmakofor düzlemine olan dik uzaklığının logaritması ile bu uzaklığa Hidrojen atomunun van der waals yarıçapı olan 1.5Å mesafesinin toplamı da matematiksel olarak ifade edilen bunun gibi değişkenlerin hesaplanması ve bulunması da EMRE programında gerçekleştirilmiştir. Son olarak yine bu programda düzlem ile seçilen iki atom arasındaki açıyı ölçmek için yine bu program kullanılmıştır**.**

Fukui Atomik Nükleofilik Reaktivite indeksi:

Karmaşık formül $: N_A = \sum_{i \in A} C_{iHOMO}^2 / (1 - \varepsilon_{HOMO})$ (3.15)

Basit formül $: N_A^{'} = \sum_{i \in A} C_{iHOMO}^2$ (3.16)

ε_{HOMO} : En yüksek enerjili dolu moleküler orbital enerjisi

C_{HOMO} : En yüksek enerjili dolu moleküler orbital katsayısı [76].

Fukui Atomik Elektrofilik Reaktivite indeksi:

Karmaşık formül $: N_A = \sum_{i \in A} C_{iLUMO}^2 / (\varepsilon_{LUMO} + 10)$ (3.17)

Basit formül $: N_A^{'} = \sum_{i \in A} C_{iLUMO}^2$ (3.18)

ε_{LUMO} : En düşük enerjili boş moleküler orbital enerjisi

C_{LUMO} : En düşük enerjili boş moleküler orbital katsayısı [76].

3.5.7. Adım 6

Alt matrisin belirlenmesinden sonra elektronik, geometrik ve fizikokimyasal özellikler kullanılarak teorik aktivite hesaplanacaktır. Bu özelliklerin etkisini aktiviteye etkisini hesaba katmak için S fonksiyonu (eşitlik 3.2) kullanılmıştır. Aktivite hesaplamaları genetik algoritma optimizasyon tekniği ile eşitlik 3.3 kullanılarak yapılmıştır.

Çalışmalarda κ_j adedi, populasyon sayısı ve κ_j limitleri değiştirilerek çeşitli denemeler yapılarak değişik bağıl hataya sahip birçok sonuç elde edildi ve bu sonuçların içinden en uygunlarını seçmek için istatistiksel olarak değerlendirme işlemine geçilir. İstatistiksel değerlendirme işlemi regresyon analizi yapılarak gerçekleştirilir. Regresyon işlemi ile elde edilen sonuçların değerlendirilmesi hesaplanan değerlerin birbirlerini ne oranda temsil ettiğini gösteren R^2 (regresyon katsayısı) ve doğru değerden ne kadar sapma olduğunu gösteren standart hata kavramları ile yapılır. İstatistiksel hesaplamalarda R kavramını hesaplayabilmek için ilk olarak eşitlik 3.19 kullanılır. Standart hatayı hesaplamak için eşitlik 3.20'de verilen formül kullanılır. İstatistiksel olarak elde edilen sonuçlara karar vermek için daha çok R^2 değerinin 1'e yakın, standart hata değerinin sıfıra yakın olması gerektiğinden ancak bu şartlar sağlandığında elde edilen sonuçların kesinliği ve doğruluğu iyidir denilebilir.

$$r^2=\left(\frac{\left(\sum xy-\frac{\sum x\sum y}{n}\right)^2}{\left(\sum x^2-\frac{(\sum x)^2}{n}\right)\left(\sum y^2-\frac{(\sum y)^2}{n}\right)}\right)^2 \tag{3.19}$$

$$se=\sqrt{\frac{1-r^2}{n-2}} \tag{3.20}$$

Ayrıca kurulan modelin doğruluğunu test etmek için çapraz doğrulama katsayısı (q2) eşitlik 3.23 ile hesaplanmıştır. q2 değerinin 0.35 den büyük değeri kabul edilebilir değerdir. Ayrıca q2 nin 1'e yakınlığı kurulan modelin doğruluğunu gösterir [77].

Yukarıdaki modelleme ve simülasyon işlemi başarıyla uygulandıktan sonra bulunan κ_j değerleri ve aktivitesi bilinmeyen bileşiğin konformerlerine ait parametreler denklemde yerine koyularak aktiviteleri hesaplanmıştır.

3.5.7.1. Optimizasyon

Optimizasyon bir problemin en iyi çözümünü bulma işlemidir. Optimizasyon bir tasarımda en uygun tasarımı bulmak için kullanılır. Tasarımda amaç sistemden minimum maliyetle maksimum performansı sağlamaktır. Klasik bir tasarımda tasarım tamamen tasarımcının bilgisi ve tecrübesi dâhilinde gerçekleştirilir. Tasarımcı tasarımın istenen performansa uygunluğunu deneyerek veya tecrübelerine dayanarak gözden geçirir ve tasarımı gerçekleştirir. Bu yaklaşım bazı yönleri ile tatmin edici olmasına karşın tasarım üzerindeki kısıtlamalar ve değişkenlerin fazla olması durumunda tasarımcı karmaşa yaşar ve tasarım istenen düzeyde olamaz. Fakat bu süreç bir optimizasyon yaklaşımı ile ele alındığında tasarım problemi bütün yönleri ile ele alınmış olur. Çünkü; optimizasyon teknikleri temel olarak matematiksel modellemelerle tasarımcıya büyük avantajlar sağlar. Şekil 3.12'de optimum tasarım prosesi gösterilmektedir. Şekilde tasarımla ilgili değişkenler ve tasarım kısıtlamaları belirlendikten sonra tasarımla ilgili bilgiler değerlendirilerek bir başlangıç tasarımı ortaya konur. Bu aşamadan sonra tasarım kısıtları kontrol edilir ve tasarımın istenen performansı sağlayıp sağlamadığı belirlenir. Eğer tasarım istenen performansa uygunsa tasarım başarıyla tamamlanmıştır. Eğer tasarım istenen performansa uygun değilse optimizasyona bağlı olarak tasarım yeniden analiz edilerek farklı çözümler üretilir.



Şekil 3.12. Optimum tasarım prosesi.

Tasarım yaparken veya üretim yaparken sistemi tanımlayan birçok parametre vardır ve bu parametrelerden en uygun olanların seçilmesi gerekir. İşte bu nedenlerle optimizasyon kullanılır ve sistemde parametrelerin optimum değerlerinin bulunması gerekir. Optimum kelime anlamı olarak en iyi, optimizasyon ise en iyileme anlamındadır. Optimizasyon, reel bir fonksiyonu maksimize veya minimize etme probleminin çözümünü, çözüm için izin verilen bir küme dâhilinde ki reel veya tamsayı değerlerini sistematik bir şekilde kullanarak arama işlemidir. Tanımdan da anlaşılacağı gibi bir işi yapmış olmakla en iyi yapmış olmak arasında fark vardır. Bu nedenle optimizasyon teknikleri kullanılır ve bu teknikler matematiksel metotlardır. Optimizasyonda amaç, optimizasyon tekniklerini kullanarak bir sistemin veya tasarımın optimum çözümünü bulmaktır. Günümüzde optimum çözümler için hâlâ diferansiyel hesap yöntemleri kullanılmaktadır [78].

Optimizasyonu problemlerinde dikkat etmemiz gereken üç önemli husus vardır. Bunlar:

1. Problemin hedefi içeren bir hedef fonksiyonu olacaktır.
2. Tasarım değişkenleri olmalıdır.
3. Problemdeki sınırlayıcı koşulları tanımlayan kısıtlar olmalıdır.

Hedef Fonksiyon (Objective Function)

Tasarımı yapılacak veya optimizasyonu yapılacak sistemin iyilik derecesini belirtir. Maksimum veya minimum olacak şekilde tanımlanır.

Tasarım-Problem Değişkenleri (Variables)

Tasarımı yapılacak veya optimizasyonu yapılacak problemin değişkenlerini gösterir.

Kısıtlayıcılar (Constraints)

Tasarım veya probleme getirilen sınırlamalara denir. Bir optimizasyon probleminde problemin uygun olduğunu gösteren ifadelere denir.

Kullanılan fonksiyonlara bağlı olarak çok sayıda optimizasyon yöntemleri bulunmaktadır. Aktivitede hesaplamasında kullanılan formül üstel ve nonlineer olduğu için en uygun optimizasyon fonksiyonunun lsqnonlin olduğu belirlenerek bu problemin çözümünde kullanılmıştır.

Lsqnonlin en küçük kareler yöntemini kullanarak lineer olmayan problemleri çözer ve sürekli fonksiyonların minimize edilmesinde kullanılır. Alınan sonuçlar lokal, yani belirli alt ve üst sınırlar içerisinde kalan sonuçlardır. En küçük kareler yöntemi, birbirine bağlı olarak değişen iki fiziksel büyüklük arasındaki matematiksel bağlantıyı, mümkün

olduğunca gerçeğe uygun bir denklem olarak yazmak için kullanılan, standart bir regresyon yöntemidir. Bir başka deyişle bu yöntem, ölçüm sonucu elde edilmiş veri noktalarına "mümkün olduğu kadar yakın" geçecek bir fonksiyon eğrisi bulmaya yarar.

3.5.7.2. Genetik Algoritma

Günümüzün karmaşık ve zor koşulları problemlere hızlı ve kolay çözüm veren yeni çözüm yöntemleri arayışına neden olmuştur. Özellikle sert(hard) optimizasyon teknikleri yerine, yumuşak hesaplama (soft computing) ve evrimsel algoritma (evolutionary algorithm) kullanımı ön plana çıkmıştır. Evrimsel yaklaşımlardan olan genetik algoritmalar da, bu arayışlar içinde önemli bir yer tutmaya başlamıştır. Uygulama başarıları artan ve sürekli geliştirilmeye çalışılan genetik algoritmalar diğer yumuşak hesaplama yöntemleri ile birlikte kullanılarak hibrid (hybrid) çözümler geliştirilmesine çalışılmaktadır. Genetik algoritmalar, doğal seçim ilkelerine dayanan bir arama ve optimizasyon yöntemidir. Temel ilkeleri John Holland tarafından ortaya atılmıştır. Temel ilkelerinin ortaya atılmasından sonra, genetik algoritmalar hakkında birçok bilimsel çalışma yayınlanmıştır. Ayrıca, genetik algoritmaların teorik kısmı ve uygulamaları hakkında birçok uluslararası konferans da düzenlenmektedir. Genetik algoritmaların, fonksiyon optimizasyonu, çizelgeleme, mekanik öğrenme, tasarım, hücresel üretim gibi alanlarda başarılı uygulamaları bulunmaktadır. Geleneksel optimizasyon yöntemlerine göre farklılıkları olan genetik algoritmalar, parametre kümesini değil kodlanmış biçimlerini kullanırlar. Olasılık kurallarına göre çalışan genetik algoritmalar, yalnızca amaç fonksiyonuna gereksinim duyar. Çözüm uzayının tamamını değil belirli bir kısmını tararlar. Böylece, etkin arama yaparak çok daha kısa bir sürede çözüme ulaşırlar [79]. Diğer bir önemli üstünlükleri ise çözümlerden oluşan populasyonu eş zamanlı incelemeleri ve böylelikle yerel en iyi çözümlere takılmamalarıdır. “Reeves; ulusal hükümetler ve organizasyonlar tarafından genetik algoritma tabanlı projelere, tavlama benzetimi (simulated annealing) ve yasaklı arama (tabu search) tabanlı projelere göre daha fazla kaynak ayrılmakta olduğunu belirtmektedir” [80]. Bu ifadeden anlaşılacağı üzere, genetik algoritmaların geleneksel optimizasyon yöntemlerine olduğu gibi sözü edilen yapay zeka yöntemlerine göre de çeşitli alanlarda üstünlükleri bulunmaktadır. Bu üstünlükler genetik algoritmaların arama yapısı ile ilgilidir. Genetik algoritmaların arama yapısı ise, alt diziler teoremi ve yapı blokları hipoteziyle açıklanmaktadır.

3.5.7.2.1. Genetik Algoritmaların Temel Teoremi ve Basit Genetik Algoritma

Genetik algoritmaların nasıl arama yaptığı alt dizi kavramıyla açıklanmaktadır. Alt diziler, genetik algoritmaların davranışlarını açıklamak için kullanılan teorik yapılardır. Bir alt dizi, belirli dizi kümeleri arasındaki benzerliği tanımlayan bir dizidir. Alt diziler, {0 ,1 , *} alfabesi kullanılarak tanımlanır. Örneğin H alt dizisi, ilk konumunda 0, ikinci ve dördüncü konumunda 1 değeri olan kromozomlar kümesi içindir.

H = 0 1 * 1 *

* sembolü dizinin o konumunun hangi değeri alıp almadığının önemli olmadığı anlamındadır. Dizi o konumda 0 veya 1 değeri alabilir. Eğer bir x dizisi, alt dizinin kalıbına uyarsa x dizisine "H'nin bir örneğidir" denir. Alt dizilerin iki özelliği mevcuttur. Bu özellikler aşağıda verilmiştir [79].

1. Alt dizi derecesi: Bir H alt dizisinin derecesi o(H) ile gösterilir ve mevcut alt dizi kalıbında bulunan sabit konumların sayısıdır. Bu sayı ikili alfabede 0 ve 1 değerlerinin sayısının toplamına eşittir.

2. Alt dizi uzunluğu: Bir H alt dizisinin uzunluğu δ(H) ile gösterilir ve mevcut alt dizi kalıbında bulunan belirli ilk ve son konumlar arasındaki uzaklıktır. Alt dizi derecesi ve alt dizi uzunluğu kavramlarının genetik algoritmaların temel teoreminde son derece önemli bir yeri vardır. Alt dizi derecesi düşük, alt dizi uzunluğu kısa olan diziler "yapı blokları" olarak adlandırılır. John Holland, genetik algoritmaların işleyişinde uygun yapı bloklarının tanımlanmasını ve bu yapı bloklarının daha uygun yapı blokları elde etmek amacıyla birleştirilmesini önermektedir. Bu fikir yapı blokları hipotezi olarak bilinmektedir. Genetik algoritmanın temel teoremi ise şöyle açıklanmaktadır [80]: Populasyon ortalamasının üstünde uyum gücü gösteren, kısa uzunluğa ve düşük dereceye sahip alt diziler zamanın ilerlemesiyle üstsel olarak çoğalırlar. Bu çoğalma, genetik işlemler aracılığı ile gerçekleşmektedir ve sonucunda ana-babadan daha üstün özellikler taşıyan bireyler ortaya çıkmaktadır. Bu çözüm kalitesinin kuşaktan kuşağa artması iki nedene bağlanmaktadır. Bu nedenler şöyle açıklanabilir:

- Başarısız olan bireylerin üreme şansları azaltıldığı için kötüye gidiş zorlaşmaktadır.
- Genetik algoritmaların yapısı kötüye gidişi engellemekle kalmamakta, genetik algoritmaların temel teoremi uyarınca, zaman içinde hızlı bir iyiye gidiş de sağlayabilmektedir.

Genetik algoritmaların işleme adımları incelendiğinde bu nedenler daha iyi anlaşılmaktadır. Genetik algoritmalar yapısı gereği, kötü bireyleri yani uygun olmayan çözümleri, operatörleri sayesinde elemektedir. Bu işlemler bir döngü içerisinde durdurma kriteri sağlanana kadar devam etmektedir.

Bir çok alanda uygulama imkânı ve uygulamaları olan genetik algoritmaların işleme adımları şöyle açıklanabilir [81]:

- Arama uzayındaki tüm mümkün çözümler dizi olarak kodlanır.
- Genellikle rastsal bir çözüm kümesi seçilir ve başlangıç populasyonu olarak kabul edilir.
- Her bir dizi için bir uygunluk değeri hesaplanır, bulunan uygunluk değerleri dizilerin çözüm kalitesini gösterir.
- Bir grup dizi belirli bir olasılık değerine göre rastsal olarak seçilip çoğalma işlemi gerçekleştirilir.
- Yeni bireylerin uygunluk değerleri hesaplanarak, çaprazlama ve mutasyon işlemlerine tabi tutulur.
- Önceden belirlenen kuşak sayısı boyunca yukarıdaki işlemler devam ettirilir.
- İterasyon, belirlenen kuşak sayısına ulaşınca işlem sona erdirilir. Amaç fonksiyonuna göre en uygun olan dizi seçilir.

Genetik algoritmalar bir çözüm uzayındaki her noktayı, kromozom adı verilen ikili bit dizisi ile kodlar. Her noktanın bir uygunluk değeri vardır. Tek bir nokta yerine, genetik algoritmalar bir populasyon olarak noktalar kümesini muhafaza eder. Her kuşakta, genetik algoritma, çaprazlama ve mutasyon gibi genetik operatörleri kullanarak yeni bir populasyon oluşturur. Birkaç kuşak sonunda, populasyon daha iyi uygunluk değerine sahip üyeleri içerir. Bu, Darwin'in rastsal mutasyona ve doğal seçime dayanan evrim modellerine benzemektedir. Genetik algoritmalar, çözümlerin kodlanmasını, uygunlukların hesaplanmasını, çoğalma, çaprazlama ve mutasyon operatörlerinin uygulanmasını içerir [82].

3.5.7.2.1.1. Çözümlerin Kodlanması

Bir problemin çözümü için genetik algoritma geliştirmenin ilk adımı, tüm çözümlerin aynı boyutlara sahip bitler dizisi biçiminde gösterilmesidir. Dizilerden her biri, problemin olası çözümler uzayındaki rastsal bir noktayı simgeler (Yeniay, 2001: 38).

Parametrelerin kodlanması, probleme özgü bilgilerin genetik algoritmanın kullanacağı şekle çevrilmesine olanak tanır [82].

3.5.7.2.1.2. İlk Populasyonun Oluşturulması

Olası çözümlerin kodlandığı bir çözüm grubu oluşturulur. Çözüm grubu populasyon, çözümlerin kodları da kromozom olarak adlandırılır. İkili alfabenin kullanıldığı kromozomların gösteriminde, ilk populasyonun oluşturulması için rastsal sayı üreticileri kullanılabilir. Rastsal sayı üreticisi çağrılır ve değer 0,5'den küçükse konum 0'a değilse 1 değerine ayarlanır [83]. Birey sayısının ve kromozom uzunluğunun az olduğu problemlerde yazı-tura ile de konum değerleri belirlenebilmektedir. Genetik algoritmalarda ikili kodlama yöntemi dışında, çözümü aranan probleme bağlı olarak farklı kodlama yöntemleri de kullanılmaktadır [79].

3.5.7.2.1.3. Uygunluk Değerinin Hesaplanması

Bir kuşak oluşturulduktan sonraki ilk adım, populasyondaki her üyenin uygunluk değerini hesaplama adımıdır. Örneğin, bir maksimizasyon problemi için i. üyenin uygunluk değeri f(i), genellikle o noktadaki amaç fonksiyonunun değeridir [82]. Çözümü aranan her problem için bir uygunluk fonksiyonu mevcuttur. Verilen belirli bir kromozom için uygunluk fonksiyonu, o kromozomun temsil ettiği çözümün kullanımıyla veya yeteneğiyle orantılı olan sayısal bir uygunluk değeri verir. Bu bilgi, her kuşakta daha uygun çözümlerin seçiminde yol göstermektedir. Bir çözümün uygunluk değeri ne kadar yüksekse, yaşama ve çoğalma şansı o kadar fazladır ve bir sonraki kuşakta temsil edilme oranı da o kadar yüksektir [80].

3.5.7.2.1.4. Çoğalma İşleminin Uygulanması

Çoğalma operatöründe diziler, amaç fonksiyonuna göre kopyalanır ve iyi kalıtsal özellikleri gelecek kuşağa daha iyi aktaracak bireyler seçilir. Üreme operatörü yapay bir seçimdir. Dizileri uygunluk değerlerine göre kopyalama, daha yüksek uygunluk değerine sahip dizilerin, bir sonraki kuşaktaki bir veya daha fazla yavruya daha yüksek bir olasılıkla katkıda bulunması anlamına gelmektedir. Çoğalma, bireyleri seçme işleminden, seçilmiş bireyleri bir eşleme havuzuna kopyalama işleminden ve havuzda bireyleri çiftler halinde gruplara ayırma işleminden oluşur [84]. Uygunluk değerinin

hesaplanması adımından sonra mevcut kuşaktan yeni bir populasyon yaratılmalıdır. Seçim işlemi, bir sonraki kuşak için yavru üretmek amacıyla hangi ailelerin yer alması gerektiğine karar vermektedir. Bu doğal seçimdeki en uygunun yaşaması durumuna benzerdir. Bu yöntemin amacı, ortalama uygunluğun üzerindeki değerlere çoğalma fırsatı tanımaktır. Bir dizinin kopyalanma şansı, uygunluk fonksiyonuyla hesaplanan dizinin uygunluk değerine bağlıdır [82]. Seçim yöntemlerine rulet tekerleği seçimi, turnuva seçimi ve sıralama seçimi gibi seçim yöntemleri örnek verilebilir.

3.5.7.2.1.5. Çaprazlama İşleminin Uygulanması

Mevcut gen havuzunun potansiyelini araştırmak üzere, bir önceki kuşaktan daha iyi nitelikler içeren yeni kromozomlar yaratmak amacıyla çaprazlama operatörü kullanılmaktadır. Çaprazlama genellikle, verilen bir çaprazlama oranına eşit bir olasılıkla seçilen aile çeşitlerine uygulanmaktadır [82]. Genetik algoritmanın performansını etkileyen önemli parametrelerden biri olan çaprazlama operatörü doğal populasyonlardaki çaprazlamaya karşılık gelmektedir. Çoğalma işlemi sonucunda elde edilen yeni populasyondan rastsal olarak iki kromozom seçilmekte ve karşılıklı çaprazlama işlemine tabi tutulmaktadır. Çaprazlama işleminde dizi uzunluğu L olmak üzere, $1 <= k <= L-1$ aralığında k tamsayısı seçilmektedir. Bu tamsayı değerine göre dizi çaprazlamaya uğratılır. En basit çaprazlama yöntemi tek noktalı çaprazlama yöntemidir. Tek noktalı çaprazlama yapılabilmesi için her iki kromozomun da aynı gen uzunluğunda olması gerekir. İki noktalı çaprazlamada ise kromozom iki noktadan kesilir ve karşılıklı olarak pozisyonlar yer değiştirilir [84].

3.5.7.2.1.6. Mutasyon İşleminin Uygulanması

Çaprazlama mevcut gen potansiyellerini araştırmak üzere kullanılır. Fakat populasyon gerekli tüm kodlanmış bilgiyi içermez ise, çaprazlama tatmin edici bir çözüm üretemez. Bundan dolayı, mevcut kromozomlardan yeni kromozomlar üretme yeteneğine sahip bir operatör gerekmektedir. Bu görevi mutasyon gerçekleştirir. Yapay genetik sistemlerde mutasyon operatörü, bir daha elde edilemeyebilir iyi bir çözümün kaybına karşı koruma sağlamaktadır [79]. İkili kodlama sisteminin kullanıldığı problemlerde mutasyon, düşük bir olasılık değeri altında bir bit değerini (0 veya 1 olabilir) diğer bit değerine dönüştürür. İkili kodlama sisteminin kullanılmadığı problemlerde ise daha farklı

mutasyon yöntemleri kullanılmaktadır. Hangi yöntem kullanılırsa kullanılsın, mutasyonun genel amacı, genetik çeşitliliği sağlamak veya korumaktır [85].

3.5.7.2.1.7. Yeni Kuşağın Oluşması ve Döngünün Durdurulması

Yeni kuşak çoğalma, çaprazlama ve mutasyon işlemlerinden sonra tanımlanmakta ve bir sonraki kuşağın ebeveynleri olmaktadırlar. Süreç yeni kuşakla çoğalma için belirlenen uygunluk ile devam eder. Bu süreç, önceden belirlenen kuşak sayısı kadar veya bir hedefe ulaşılıncaya kadar ya da başka bir durdurma kriteri sağlanana kadar devam eder. İstenen hassasiyet derecesine göre de maksimum iterasyon sayısı belirlenebilmekte ve iterasyon bu sayıya ulaştığında döngü durdurulabilmektedir. Durdurma kriteri iterasyon sayısı olabileceği gibi hedeflenen uygunluk değeri de olabilmektedir [86].

3.5.7.2.2. Genetik Algoritmalarda Parametre Seçimi

Parametreler, genetik algoritma performansı üzerinde önemli etkiye sahiptir. Optimal kontrol parametreleri bulmak için birçok çalışma yapılmıştır fakat tüm problemler için genel olarak kullanılabilecek parametreler bulunamamıştır [87]. Bu parametreler, kontrol parametreleri olarak adlandırılmaktadır. Kontrol parametreleri populasyon büyüklüğü, çaprazlama olasılığı, mutasyon olasılığı, kuşak aralığı, seçim stratejisi ve fonksiyon ölçeklemesi olarak sayılabilir. Bu parametreler aşağıda açıklanmıştır [80]:

Populasyon Büyüklüğü: Genetik algoritma kullanıcısı tarafından verilen en önemli kararlardan birisidir. Bu değer çok küçük olduğunda, genetik algoritma yerel bir optimuma takılabilmektedir. Populasyonun çok büyük olması ise çözüme ulaşma zamanını arttırmaktadır. Bu konuda Goldberg 1985'de, yalnızca kromozom uzunluğuna bağlı bir populasyon büyüklüğü hesaplama yöntemi önermiştir. Ayrıca Schaffer ve arkadaşları 1989'da çok sayıda test fonksiyonları üzerinde yaptıkları araştırmalar sonucunda, 20–30 arası bir populasyon büyüklüğünün iyi sonuçlar verdiğini belirtmişlerdir.

Çaprazlama Olasılığı: Çaprazlamanın amacı, mevcut iyi kromozomların özelliklerini birleştirerek daha uygun kromozomlar yaratmaktır. Kromozom çiftleri P(c) olasılığı ile çaprazlamaya uğramak üzere seçilirler. Çaprazlamanın artması, yapı bloklarının artmasına neden olmakta fakat aynı zamanda bazı iyi kromozomların da bozulma olasılığını arttırmaktadır.

Mutasyon Olasılığı: Mutasyonun amacı populasyondaki genetik çeşitliliği korumaktır. Mutasyon P(m) olasılığı ile bir kromozomdaki her bitte meydana gelebilir. Eğer mutasyon olasılığı artarsa, genetik arama rastsal bir aramaya dönüşür. Fakat bu aynı zamanda kayıp genetik malzemeyi tekrar bulmada yardımcı olmaktadır

3.5.7.2.3. Genetik Algoritmaların Uygulama Alanları

Karmaşık problemleri hızlı ve optimale yakın olarak çözebilen genetik algoritmalar, çeşitli problem tiplerine uygulanabilmektedir. Büyük çözüm uzaylarının geleneksel yöntemlerle taranması hesaplama zamanını arttırmaktadır. Ancak bu tip problemlere, genetik algoritmalar ile kısa sürede, kabul edilebilir çözümler bulunabilmektedir [88]. Genetik algoritmalar özellikle çözüm uzayının geniş, süreksiz ve karmaşık olduğu problem tiplerinde başarılı sonuçlar vermektedir.

3.5.7.3. Genetik Algoritmanın Elektron Konformasyonel Yönteme Uygulanışı

4D-QSAR çalışmalarımızda moleküllerin konformerleride dikkata alındığı için çok sayıda parametre ortaya çıkmakta aktiviteye etkisi fazla olan en iyi parametre ve parametre grubunu ortaya çıkarmak ve incelenen bileşik serisi için bir model oluşturabilmek için elektron konformasyonel QSAR çalışması ile genetik algoritma tekniği birleştirilerek EC-GA metodu geliştirlmiştir. Bu çalısmada aktivite denklemini çözmek ve aktiviteye en çok etki eden parametre grubunu bulmak için Matlab 7.0 programında optimizasyon metotlarından olan genetik algoritma tekniğini kullanan yeni bir kod yazılmıştır. Yazılan bu kodda genetik algoritma optimizasyon tekniği içerisinde yer alan lsqnonlin komutu kullanılmıştır.

Genetik algoritma optimizasyon tekniğinde kullanılacak olan akış şeması aşağıda şekil 3.13'deki gibidir. Karecikler içindeki sayılar gösterimin daha iyi anlaşılabilmesi için oluşturulmuş ve hesaplamalarda kullanılan rasgele seçilmiş moleküler parametrelerden bazılarını göstermektedir.

Programda sayısı kullanıcı tarafından belirlenen, satır sayısı popülasyon sayısına eşit ve sütun sayısı da kapa adedine eşit olan bir matris oluşturulur. EMRE programı kullanılarak hazırlanan parametre (özellik) listesinden rasgele seçilen özellikler ile bu matrisin elemanları oluşturulur. Matrisin her bir satırı (kapa seti) lsqnonlin fonksiyonuna gönderilerek κ değerleri hesaplanır.

Matrisin her bir satırı hesaplanan fonksiyon değerine göre küçükten büyüğe doğru sıraya dizilir ve kullanıcının belirlediği sayıda en iyi elemanlar (elit count) seçilir ve kalan satırlar silinir.

Matristen rasgele iki satır (κ seti) seçilir. Bu iki κ seti rasgele noktalardan çaprazlanarak yeni setler oluşturulur. Matrisin kalan kısmının %75'i çaprazlama işlemiyle oluşturulur. Kalan kısmın ise %0.015'i mutasyon olarak adlandırılan işlemle oluşturulur. Bu işlemde ise rasgele seçilen bir satırdan yine rasgele seçilen bir noktadan

çıkarılan κ yerine parametre listesinden yeni bir κ yerleştirilir. Matrisin kalan son kısmı da yine parametre listesinden rasgele seçilen parametrelerle doldurulur. Bütün bu işlemler ile yeniden oluşturulan matris için tekrar hesaplama yapılır. Böylece iyi olan κ setleri tutularak iyi olmayan setler elimine edilmiş olur. Jenerasyon sayısı, uygunluk değeri ifadelerinde yer alan limit değerleri tamamlanıncaya kadar döngü devam eder. Jenerasyon sayısı için limit değer 500, popülasyon boyutu 500 ve iterasyon sayısı da 500 olarak alınmıştır. Uygunluk fonksiyonu (fitness function) olarak toplam bağıl hata formülü kullanılır:

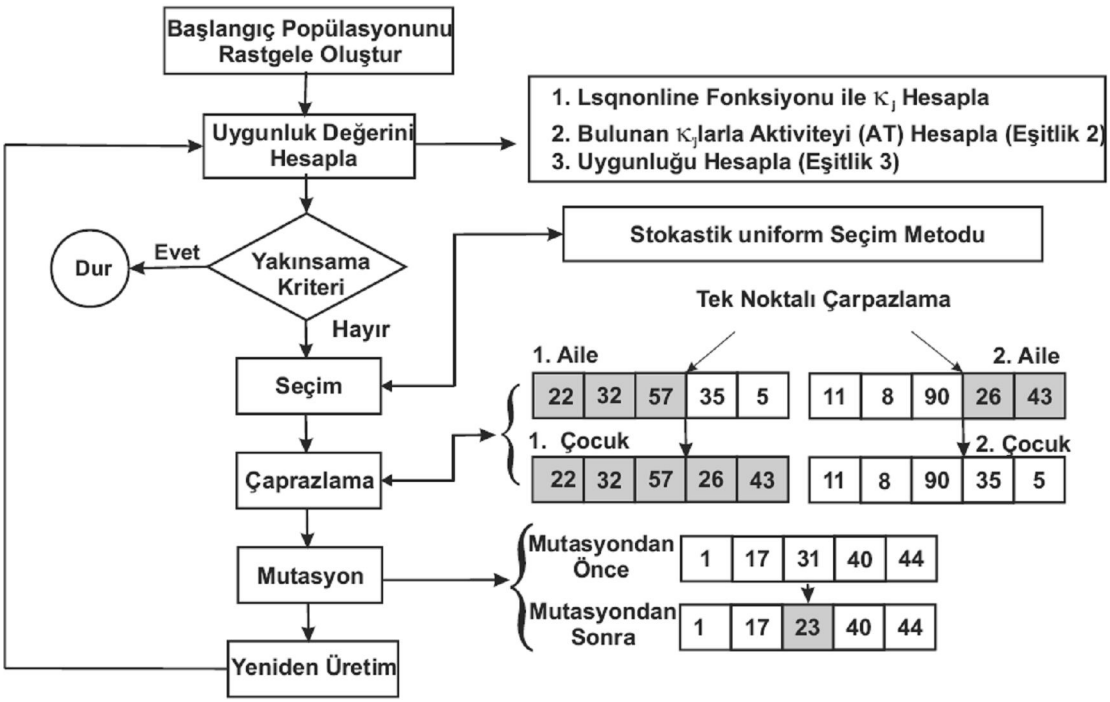
$$f = \mathrm{PRESS_N} = \sum_{n=1}^{N} \left| \mathrm{A_n^{den}} - \mathrm{A_n^{hes}} \right|^2 \qquad (3.21)$$

A_n^{den} : Deneysel aktivite değeri

A_n^{hes} : Hesaplanan aktivite değeri

n: Bileşik sayısı

Çalışmalarda κ_j adedi, popülasyon sayısı ve κ_j limitleri değiştirilerek çeşitli denemeler yapılarak değişik bağıl hataya sahip birçok sonuç elde edilir ve bu sonuçların içinden en uygunlarını seçmek için istatistiksel olarak değerlendirme işlemine geçilir. İstatistiksel değerlendirme işlemi regresyon analizi yapılarak gerçekleştirilir.



Şekil 3.13. Genetik algoritmanın aşamaları.

3.5.7.4. İstatiksel Analiz

Model parametrelerinin doğruluğu, çapraz doğrulama tekniği kullanılarak test edilmiştir. Çapraz doğrulama tüm veriyi kullanmayı sağlayan bir metottur. Veri seti rasgele 2 eşit kısma ayrılır. İlk veri setiyle model kurulur, kurulan model diğer veri setiyle test edilir ve doğruluk oranı hesaplanır.

3.5.7.4.1. Çapraz Doğrulama (Cross Validation)

QSAR modellerinin önemli karakteristiklerinden biri de tahmin gücüdür. Model geçerliliği dâhili (internal) ve harici (external) doğrulama (validation) teknikleri ile test edilmiştir [89].

3.5.7.4.2. Dâhili doğrulama (İnternal Validation)

PRESS (The predictive residual sum of squares) model geçerliliğini ölçmek için geçerli örnek sayısı için çapraz doğrulama (LOO cross validation) temelli standart bir indekstir. *PRESS* gözlemlenen aktivite değerleri ile tahmin edilen aktivite değerleri arasındaki farkın kareleri toplamıdır.

$$PRESS_N = \sum_{n=1}^{N} \left| A_n^{den} - A_n^{hes} \right|^2 \tag{3.22}$$

Burada A_n^{den} deneysel aktivite değerlerini, A_n^{hes} ise LOO çapraz doğrulama metodu ile hesaplanan aktivite değerlerini göstermektedir. Modelin tahmini performansının bir ölçümü olarak tahmini korelasyon katsayısı karesi (q^2) eşitlik 3.23 ile hesaplanmaktadır.

$$q^2 = 1 - \frac{\sum_{n=1}^{N} \left| A_n^{den} - A_n^{hes} \right|^2}{\sum_{n=1}^{N} \left| A_n^{den} - \overline{A}_n^{den} \right|^2} \equiv 1 - \frac{PRESS}{SSY} \tag{3.23}$$

Burada N eğitim setindeki bileşiklerin sayısı, $\overline{A}_n^{den}$ ise deneysel aktivite değerlerinin ortalamasını göstermektedir.

3.5.7.4.3. Harici Doğrulama (External Validation)

Harici Doğrulama (External Validation) metodu eğitim setinde bulunmayan bileşikler seçilerek oluşturulmuş test seti bileşiklerine uygulanır. QSAR modeli eğitim setindeki bileşikler kullanılarak oluşturulur ve modelin tahmin gücünü doğrulamak için test seti bileşiklerine uygulanır. Bu metot QSAR modelinin gücünü saptamak için sadece bir yoldur. Harici Doğrulamadan q^2 hesaplamak için schüüremann ve arkadaşlarının kullandığı iki farklı eşitlik vardır [90, 91].

$$q_{ext1}^2 = 1 - \frac{\sum_{n=1}^{N} \left| A_{n_{test}}^{den} - A_{n_{test}}^{hes} \right|^2}{\sum_{n=1}^{N} \left| A_{n_{test}}^{den} - \overline{A}_{n_{eğitim}}^{den} \right|^2} \tag{3.24}$$

$$q^2_{ext2} = 1 - \frac{\sum_{n=1}^{N} \left| A^{den}_{n_{test}} - A^{hes}_{n_{test}} \right|^2}{\sum_{n=1}^{N} \left| A^{den}_{n_{test}} - \overline{A}^{den}_{n_{test}} \right|^2} \tag{3.25}$$

Burada N test bileşiklerinin sayısını, A^{den}_n deneysel aktivite değerlerini, A^{hes}_n hesaplanan aktivite değerlerini, $\overline{A}^{den}_{eğitim}$ eğitim seti için aktivite değerleri ortalaması, $\overline{A}^{den}_{n_{test}}$ test seti için aktivite değerleri ortalamasını göstermektedir.

4. BÖLÜM

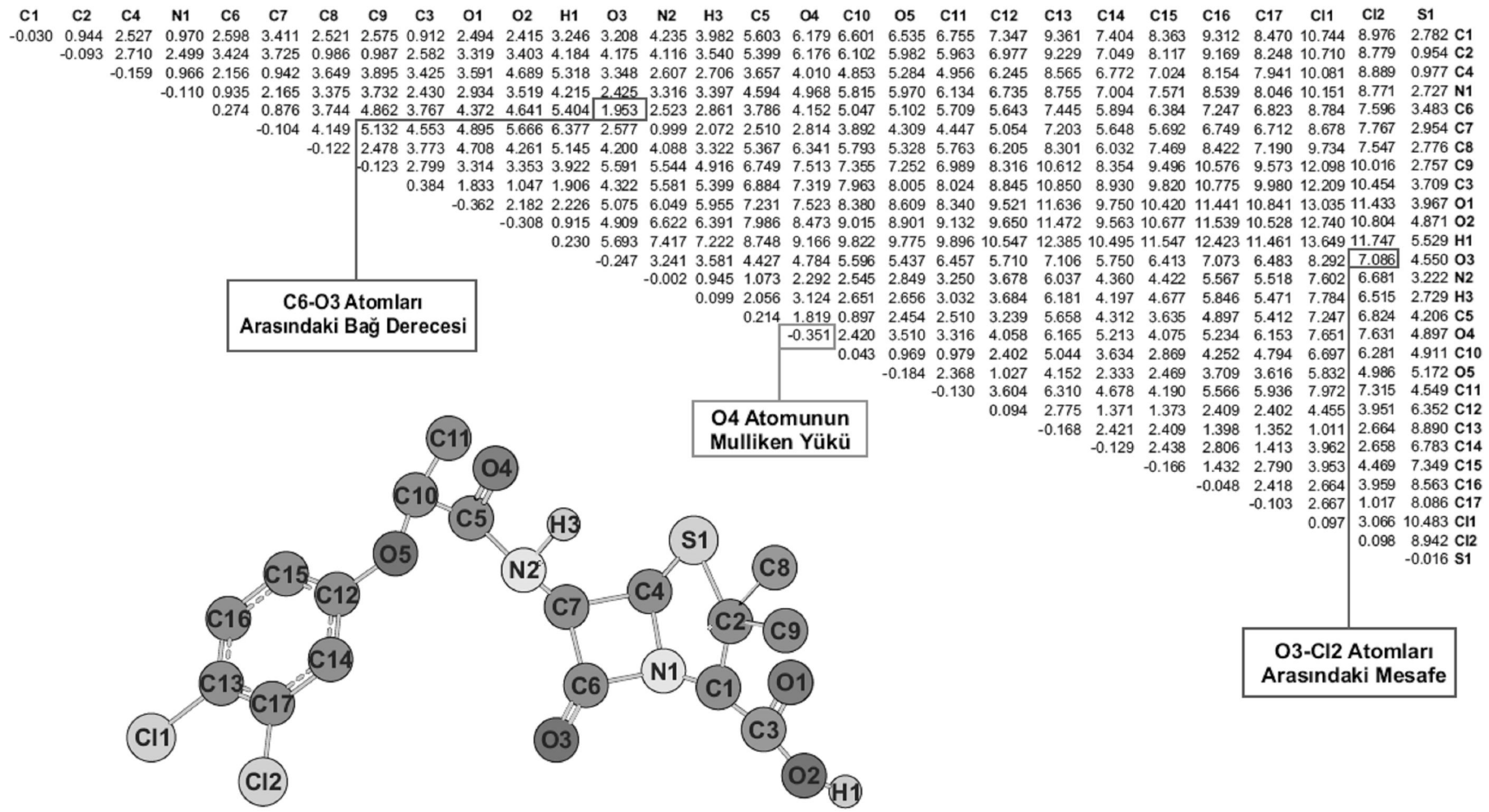
BULGULAR

4.1. ECMC Matrislerinin Hazırlanması ve Farmakofor Bulunması

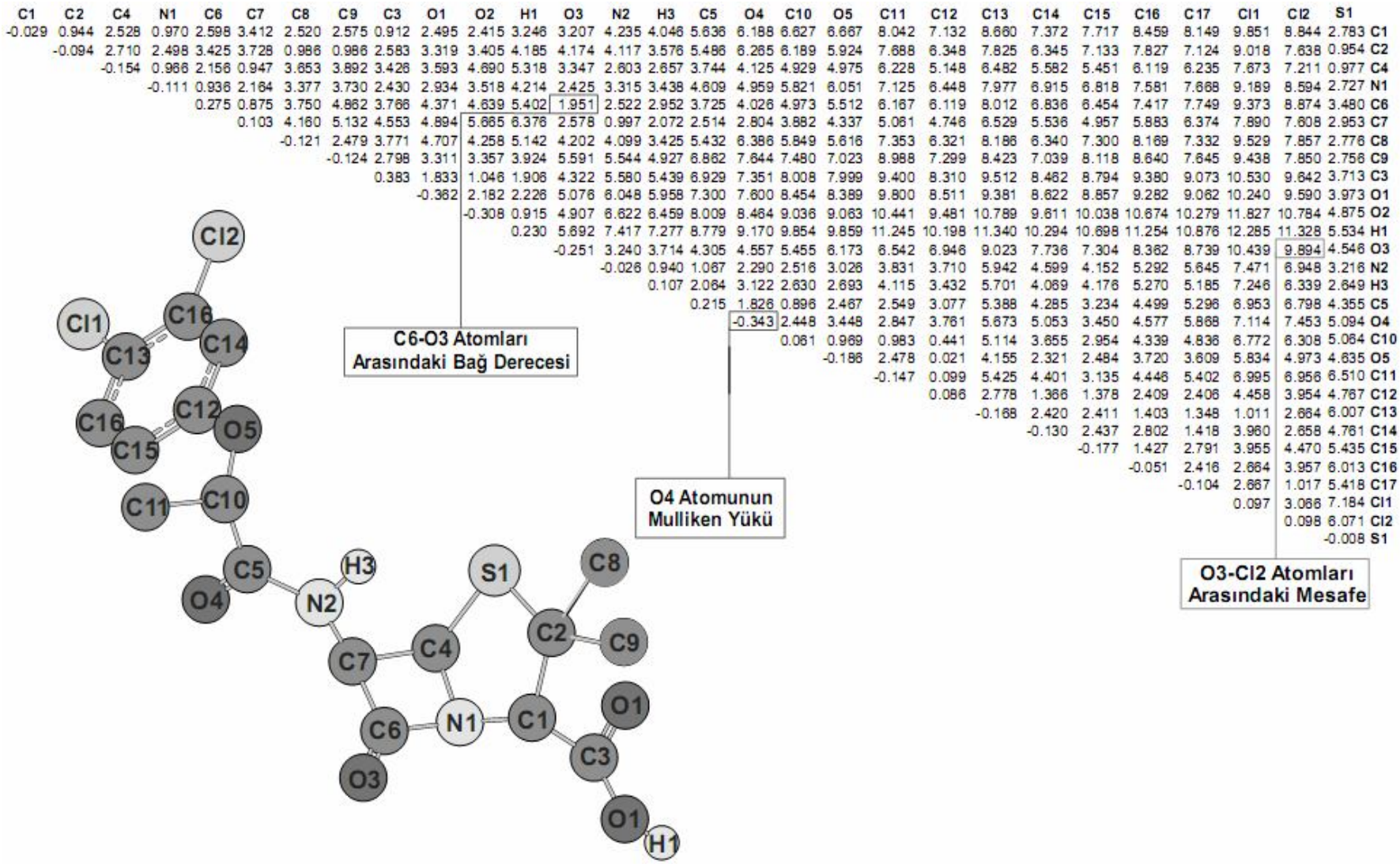
Bu çalışmada *.txt dosyalarındaki bilgiler içerisinden atomların yükleri, kartezyen koordinatları, bağ dereceleri ve atomlar arası mesafeler ile moleküle ait enerjiler kullanılarak 3. bölümde açıklandığı gibi ECMC oluşturulmuştur [4].

ECMC hesaplamalarında txt dosyalarından mulliken yükleri kullanılarak matris oluşturulmuştur ve farmakoforu belirleme işlemi bütün matrisler ayrı ayrı dikkate alınarak gerçekleştirilmiştir. Şekil 4.1 ve 4.2'de penisilin serisinde 73. bileşiğe ait 1. ve 15. konformerleri için mulliken yükleri kullanılarak oluşturulan ECMC'ler gösterilmiştir. Konformerler değiştikçe matris değişmektedir. Karbona bağlı hidrojen atomları reseptörle etkileşimde çok fazla etkiye sahip olmadığı için matris de dikkate alınmamıştır [4, 60, 61].

Şekil 4.1 ve 4.2'de penisilin serisindeki 73. bileşiğine ait 1. ve 15. konformerleri için verilen matrislerde köşegen elemanlardan biri olan O4 atomuna ait Mulliken yükü (1. konformer için -0.351; 15. konformer için -0.343) değerini göstermektedir. Köşegen olmayan diğer iki eleman ise C6-O3 atomları için bağ derecesini (1. konformer için 1.953; 15. konformer için 1.951) ve O3-Cl2 atomları arasındaki mesafeyi (1. konformer için 7.086; 15. konformer için 9.894) göstermektedir. Konformerlerdeki atomların farklı yönlenmeleri nedeniyle, her iki matrisin sayısal değerlerindeki farklılık atomlar arası mesafede daha belirgin olarak gözlenmektedir. Teorik aktivite hesaplama formülünde her bir konformerin katkısı Boltzmann dağılımlarına göre verilmektedir. Enerjisi düşük olan molekülün oda sıcaklığında bulunma ihtimali diğerlerine göre daha yüksektir. Ligand temelli QSAR çalışmalarında reseptör ile ligandın etkileşmesi dikkate alınmamaktadır.



Şekil 4.1. Penisilin serisinde 73. bileşiğinin 1. konformerine (73_1) ait ECMC.



Şekil 4.2. Penisilin serisinde 73. bileşiğinin 15. konformerine (73_15) ait ECMC.

Çünkü reseptörün aktiflenmiş durumda yapısının bilinmesi neredeyse imkânsızdır. Dikkat edildiğinde O4 atomunun yükü ve C6-O3 atomlarının bağ derecesi 73. bileşiğin 1. ve 15. konformerinde birbirine yakın değerdedirler. Her bir konformerin katkısı aktivite formülünde görüldüğü gibi boltzmann değerleri ile çarpılıp toplanmasıyla elde edilmektedir. O4-Cl2 atomları birbirlerine bağlı atomlar olmadıkları için burada bağ derecesinin gerçek değeri sıfırdır.

Düşük enerjili olan konformerin aktiviteye katkısı dolayısıyla resaptörle etkileşme ihtimali yüksek enerjili konformerden daha fazladır. Örneğin 73. bileşiğinde O4-Cl2 atomları arasındaki bağ uzunluğu konformerden konformere değişmektedir. Reseptör ile bu mesafenin etkileşmesinde her bir konformerin katkısı boltzman değeriyle çarpılıp toplanmasıyla elde edilmektedir.

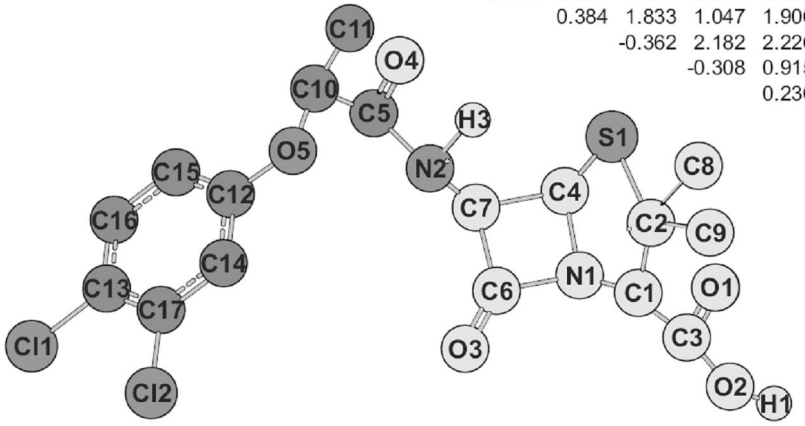
Hesaplamalar ile elde edilen sonuçların mevcut yeni yöntemlerle değerlendirilmesi ve hesaplamalar sonucunda bileşiklerin elde edilen kuantum kimyasal özelliklerin output dosyalarından matrislerin hazırlanması için EMRE veya ECSP programı kullanılmıştır. Bu programlar Spartan 02, Spartan 04, Spartan 06 ve Spartan 08 programı ile elde edilen oldukça karışık çıktı dosyalarının düzenli birer matris dosyası haline dönüştürür ve bu program kullanılarak penisilin serisine ait 88 bileşiğin 1222 adet konformerinin hepsi için matrisler oluşturulmuştur. Şekil 4.1 ve 4.2'de görülen matrisler tek konformer için üç boyutlu olarak hazırlanmıştır. Bileşik serilerine ait tüm konformerler için şekil 3.6'da gösterildiği gibi dört boyutlu matris hazırlanmıştır. Her bir konformerin. txt dosyasının büyüklüğü yaklaşık olarak 7.00 – 10.00 MB büyüklükte olup yaklaşık binlerce sayfa olduğundan matrislerin elle hazırlanması mümkün değildir. EMRE programı işimizi oldukça kolaylaştırmıştır.

Hazırlanan ECMC matrisleri kullanılarak atomik yükler ve atomlar arası mesafelerin karşılaştırılması ile ECSP (Elektron Konformasyon Alt Matris Programı) programı ile farmokofor grubu temsil eden alt matris (ECSA elektron konformasyonel alt matris) belirlenmiştir. ECSP programı serideki bütün bileşiklerin konformerlerinin matrislerinin yer aldığı .mat uzantılı dosyayı kullanarak verilen tolerans değerleri içerisinde tarama yaparak sonuçların bulunduğu çıktı dizinini oluşturur. Bu çıktı dizininde oluşan dosyalar incelendiğinde, elde edilen hesaplama sonuçları içerisinden analiz yapılarak en uygun olan farmakofor grup belirlenir.

ECSP programı kullanılarak penisilin serisi için aktiviteden sorumlu olan farmakofor grubu bulunmuştur. Birden fazla alt matris bulunma ihtimali olduğu için farmakoforu temsil eden matrisin (grubun) hangisi olacağına grubun hem düşük tolerans değerine ve aynı zamanda da bileşik serisinin bütün konformerlerin hepsinde o grubun çıkıp çıkmamasına bakılarak karar verilir. Buna göre çıkan guruplar tekrar analiz edilerek en uygun grup bulunmaya çalışılır ve bulunan bu alt matris farmakofor grubu olarak tanımlanır. Bersuker tarafından ortaya atılan metoda göre aktivitesi en yüksek yaklaşık 8 ila 10 molekül öncül olarak alınır, kendi içinde karşılaştırılarak muhtemel farmakofor grubunu içeren bir üst matris bulunur. Daha sonra geriye kalan bütün bileşiklerlede karşılaştırma yapılarak ortak farmakofor grubu bulunur [73]. İncelediğimiz penisilin serisinde 88 adet bileşik ve bunlara ait toplam 1222 konformer bulunmaktadır. Benzer hesaplama metodu bizim çalışmalırımızda da kullanılmıştır. Farmakofor grubunun tayininde ilk önce aktivite değeri yüksek olan 10 adet bileşik (bu bileşikler 73, 74, 72, 70, 71, 69, 67, 68, 87, 66) kendi arasında belli tolerans değerlerinde karşılaştırılarak bu 10 bileşiğin hepsinde çıkan düşük tolerans değerli en çok atom içeren grup üst grup olarak belirlenir. Bu işlem sonucunda aşağıda şekil 4.3'de verilen 15 li bir üst grup bulunmuştur.

Aşağıda 73. bileşiğin (referans) ilk konformerinde üst gruba ait matris ve 73. bileşiğin en düşük enerjili konformerinin şekli görülmektedir. Şekilde açık renkli olarak simgelenen atomlar üst grup atomlarıdır.

C1	C2	C4	N1	C6	C7	C8	C9	C3	O1	O2	H1	O3	H3	O4	
-0.030	0.944	2.527	0.970	2.598	3.411	2.521	2.575	0.912	2.494	2.415	3.246	3.208	3.982	6.179	C1
	-0.093	2.710	2.499	3.424	3.725	0.986	0.987	2.582	3.319	3.403	4.184	4.175	3.540	6.176	C2
		-0.159	0.966	2.156	0.942	3.649	3.895	3.425	3.591	4.689	5.318	3.348	2.706	4.010	C4
			-0.110	0.935	2.165	3.375	3.732	2.430	2.934	3.519	4.215	2.425	3.397	4.968	N1
				0.274	0.876	3.744	4.862	3.767	4.372	4.641	5.404	1.953	2.861	4.152	C6
					-0.104	4.149	5.132	4.553	4.895	5.666	6.377	2.577	2.072	2.814	C7
						-0.122	2.478	3.773	4.708	4.261	5.145	4.200	3.322	6.341	C8
							-0.123	2.799	3.314	3.353	3.922	5.591	4.916	7.513	C9
								0.384	1.833	1.047	1.906	4.322	5.399	7.319	C3
									-0.362	2.182	2.226	5.075	5.955	7.523	O1
										-0.308	0.915	4.909	6.391	8.473	O2
											0.230	5.693	7.222	9.166	H1
												-0.247	3.581	4.784	O3
													0.099	3.124	H3
														-0.351	O4



Şekil 4.3. 73. bileşik için üst gruba ait matris ve bileşiğin en düşük enerjili konformerinin şekli.

Aktivitesi en yüksek olan 73 nolu bileşiğin en düşük enerjili konformeri referans olarak seçilmiştir. Genel anlamda bütün bileşikler ve konformerler referans olarak alınabilir. Buda her referans için farklı matris değerleri ve tolerans değerleri karşımıza çıkarır ve bu durumda karışıklığa neden olur. Bu karışıklığı önlemek için literatürler de olduğu gibi tek bileşiğin tek konformeri referans alınmıştır.

Daha sonra ki adımda bu üst grup dikkate alınarak bileşiklerin en düşük enerjili konformerleri hesaba katılarak belirli tolerans değerlerinde hesaplar verilir. Bu hesaplamalar sonucunda üst grubun içinde olan tolerans değeri küçük 88 bileşiğin hepsinin en düşük enerjili konformerinde çıkan gruplar belirlenir. Sonraki adımda bileşiklerin bütün konformerleri hesaba katılarak belirli tolerans değerlerinde hesaplar verilir. Bu hesaplamalar sonucunda bileşik serisinin 1222 konformerin hepsinde çıkan tolerans değeri küçük gruplar belirler. Bu işlemler sonucunda bütün sonuçların analizleri yapılarak bir alt matris bulunur ve bu alt matrise farmakofor denir. Bu işlemler sonucunda penisilin serisinde 88 bileşik için farmakofor N1, C6, C3, O1, O2, H1 ve O3 atomlarından oluşan bir 7'li alt grup bulunmuştur. Tablo 4.1'de sırasıyla 73. bileşiğin (referans) ilk konformerinde farmakofor gruba ait alt matris, penisilin serisinde her bileşik için en düşük enerjili konformeri (88 konformer) dikkate alındığında ortaya çıkan alt matris elemanlarının tolerans değerleri ve penisilin serisinde bütün bileşiklerin

bütün konformerleri (1222 konformer) dikkate alındığında ortaya çıkan alt matris elemanlarının tolerans değerleri görülmektedir.

Tablo 4.1. a) 73. bileşiğin (referans) ilk konformerinde farmakofor gruba ait alt matris görülmektedir, b) 88 Bileşiğin en düşük enerjili konformerleri için minimum tolerans değerleri, c) 88 Bileşiğin 1222 konformeri için minimum tolerans değerleri.

a) Referans Bileşiğin ECSA (Pha)

N1	**C6**	**C3**	**O1**	**O2**	**H1**	**O3**	
-0.110	0.935	2.430	2.934	3.519	4.215	2.425	**N1**
	0.274	3.767	4.372	4.641	5.404	1.953	**C6**
		0.384	1.833	1.047	1.906	4.322	**C3**
			-0.362	2.182	2.226	5.075	**O1**
				-0.308	0.915	4.909	**O2**
					0.230	5.693	**H1**
						-0.247	**O3**

b) 88 Bileşiğin en düşük enerjili konformerleri için minimum tolerans değerleri

N1	**C6**	**C3**	**O1**	**O2**	**H1**	**O3**	
±0.061	±0.035	±0.089	±0.576	±0.681	±0.497	±0.015	**N1**
	±0.036	±0.380	±0.286	±1.061	±1.014	±0.031	**C6**
		±0.072	±0.051	±0.053	±0.524	±0.662	**C3**
			±0.063	±0.098	±1.190	±0.497	**O1**
				±0.074	±0.045	±1.410	**O2**
					±0.013	±1.535	**H1**
						±0.013	**O3**
m_max:±1.535 e88_1				y_max:±0.074 e77_1			

c) 88 Bileşiğin 1222 konformeri icin minimum tolerans değerleri

N1	**C6**	**C3**	**O1**	**O2**	**H1**	**O3**	
±0.064	±0.041	±0.093	±0.671	±0.772	±1.811	±0.019	**N1**
	±0.042	±0.405	±0.628	±1.068	±2.318	±0.046	**C6**
		±0.072	±0.051	±0.053	±0.534	±0.709	**C3**
			±0.063	±0.099	±1.192	±1.404	**O1**
				±0.074	±0.045	±1.469	**O2**
					±0.027	±2.898	**H1**
						±0.028	**O3**
m_max:±2.898 e88_10				y_max:±0.074 e77_1			

Yukarıdaki matriste m_max, mesafe için maksimum tolerans değerini; y_max ise yük için maksimum tolerans değerini belirtmektedir [73]. Yukarıdaki matriste 88 Bileşiğin en düşük enerjili konformerleri için maksimum tolerans değeri mesafede 88 nolu bileşikte olup değeri ±1.535, yüklerde ise maksimum tolerans değeri 77 nolu bileşikte olup değeri ±0.074 olarak bulunmuştur.

88 Bileşiğin 1222 konformeri için mesafe de 88 nolu bileşiğin 10. konformerinde (88_10) maksimum tolerans değerini veririken (m_max = ±2.898), 77 nolu bileşiğin 1. konformeri de (77_1) yük için maksimum tolerans değerini (y _max=±0.074) vermektedir. Bu seri içinde en düşük enerjili konformerler dikkate alındığında mesafe için maksimum tolerans değerleri 88 bileşiğin 1222 konformeri daikkate alındığındaki tolereans değerinden düşüktür.

Yedili bir alt matristen oluşan ECSA'da maksimum tolerans değerleri mesafe ve yük için sırasıyla ±2.898, yük için ±0.074'dir. Bu tolerans değerleri içerisinde belirlenen farmakofor atomları N1, C6, C3, O1, O2, H1 ve O3'dır. Yük için en yüksek tolerans değerini 77. bileşiğinin 10. konformeri verirken, mesafe için en yüksek tolerans değerini 88. bileşiğinin 1. konformeri vermiştir. Yani mesafe için maksimum tolerans değeri ±2.898 olduğunda 88. bileşiğinin l konformerinde de bu atomlar çıkmaktadır. Benzer şekilde yük için maksimum tolerans değeri ±0.074 olduğunda 77. bileşiğinin 10. konformerinde de farmakofor atomları çıkmaktadır.

4.2. Aktivite Hesaplamasında Kullanılan Parametreler

Farmakofor grubu bulunduktan sonra seçilen bileşik serisinde aktivite hesaplamalarına geçilmiştir. Bunun için aktiviteye etki ettiği düşünülen 646 adet parametre penisilin serisindeki 88 bileşiğin her bir konformeri için EMRE programında aktivite formülünde (eşitlik 3.3) kullanılabilmesi için uygun formatta hazırlanmıştır. Bileşik ve konformer sayısı arttıkça üstel olarak parametre sayısı da artacağından, penisilin serisinde bütün konformerler için 646 adet parametre hazırlanmıştır. Penisilin serisi için hesaplamalarda kullanılan parametreler ve 73. bileşiğinin en düşük enerjili konformerine ait sayısal değerler tablo 4.2'de verilmiştir. Bütün bileşiklerin bütün konformerlerinin parametrelerini ve sayısal değerlerini vermek çok yer kaplayacağından sadece referans bileşik olan 73 nolu bileşiğin en düşük enerjili konformerinin değeri verilmiştir.

Tablo 4.2. Penisilin serisi için hesaplamalarda kullanılan parametreler ve 73 bileşiğinin en düşük enerjili konformerine ait sayısal değerler.

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
1	Softness(eV)	**0.2297**	**2**	Elektronegatiflik (eV)	**4.8982**
3	HOMO(eV)	**-9.1837**	**4**	LUMO (eV)	**-0.5021**
5	LUMO-HOMO(eV)	**8.6816**	**6**	Kimyasal Potansiyel(eV)	**-4.8429**
7	Kimyasal Sertlik(eV)	**4.3408**	**8**	Elektrofilik index (eV)	**50.9032**
9	1-Nükleofilik (eV)	**0.3702**	**10**	HOMO(eigenvalue) (birimsiz)	**-0.3375**
11	LUMO(eigenvalue) (birimsiz)	**-0.0185**	**12**	Ovality	**1.7370**
13	Yüzey alanı ($Å^2$)	**422.4500**	**14**	Hacim($Å^3$)	**356.5300**
15	Moleküler ağırlık (g/mol)	**433.3100**	**16**	Olş. En. + Solv. En. (kcal/mol)	**-217.2290**
17	Solvasyon Enerjisi (kcal/mol)	**-59.1450**	**18**	Total Entalpi (kcal/mol)	**225.8870**
19	Translational Entalpi (kcal/mol)	**0.8890**	**20**	Rotational Entalpi (kcal/mol)	**0.8890**
21	Vibrational Entalpi (kcal/mol)	**224.1090**	**22**	Total Entropi (kcal/mol)	**194.9650**
23	Translational Entropi (kcal/mol)	**44.0800**	**24**	Rotational Entropi (kcal/mol)	**36.8860**
25	Vibrational Entropi (kcal/mol)	**113.9990**	**26**	Gibbs Enerjisi (kcal/mol)	**167.7580**
27	Q(minus) (e^-)	**-0.3623**	**28**	Q(plus) (e^-)	**0.2303**
29	Elektron Sayısı (birimsiz)	**146**	**30**	Zero Point (kcal/mol)	**208.6250**
31	Log P (birimsiz)	**2.0900**	**32**	Temel Titreşim Frekansı (cm-1)	**6.9200**
	Dipol (DB)				
33	Dipol (X)	**-0.3480**	**34**	Dipol (Y)	**-0.1865**
35	Dipol (Z)	**-3.5240**	**36**	Dipol (Toplam)	**3.5461**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
POLARLANABİLİRLİK (DİPOL)					
Polarlanabilirlik (Dipol) (alpha) (au^3)					
37	XX	**221.0316**	**38**	YY	**219.2634**
39	ZZ	**137.7064**	**40**	XY	**-18.2963**
41	XZ	**36.3379**	**42**	YZ	**19.4045**
43	Average	**192.6671**			
Polarlanabilirlik (beta) (au^5)					
44	XXX	**-223.4050**	**45**	XYY	**-25.7030**
46	XZZ	**-37.1230**	**47**	YYY	**550.9510**
48	YXX	**43.8530**	**49**	YZZ	**45.7960**
50	ZZZ	**-64.7390**	**51**	ZXX	**6.9310**
52	ZYY	**125.1040**			
Polarlanabilirlik (gamma) (au)					
53	XXXX	**95436.0900**	**54**	YYYY	**33376.7160**
55	ZZZZ	**9251.6060**	**56**	XXYY	**16734.3990**
57	XXZZ	**10322.7060**	**58**	YYZZ	**6521.0210**
C6-O3-C3 Düzlemine Atomun Dik Uzaklığı (Å)					
60	C5	**1.1217**	**61**	O4	**0.0285**
62	N2	**1.2131**	**63**	H3	**2.0770**
64	C7	**0.0575**	**65**	C4	**0.1060**
66	S1	**1.6528**	**67**	C1	**0.8069**
68	C2	**2.0614**	**69**	C8	**3.2385**
70	C9	**2.4632**	**71**	H1	**0.3323**
72	N1	**0.0981**	**73**	O2	**0.2145**
74	O1	**0.8006**	**75**	E	**3.9870**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
C6-O3-O1 Düzlemine Atomun Dik Uzaklığı (Å)					
76	C5	**0.4121**	**77**	O4	**0.7660**
78	N2	**0.7842**	**79**	H3	**1.7262**
80	C7	**0.1566**	**81**	C4	**0.1524**
82	S1	**1.7624**	**83**	C1	**1.2758**
84	C2	**2.4525**	**85**	C8	**3.5124**
86	C9	**3.1046**	**87**	H1	**0.8034**
88	N1	**0.1446**	**89**	O2	**1.1835**
90	C3	**0.7483**	**91**	E	**2.4902**
C6-C3-O1 Düzlemine Atomun Dik Uzaklığı (Å)					
92	C5	**2.6441**	**93**	O4	**2.1079**
94	N2	**2.0681**	**95**	H3	**2.7038**
96	C7	**1.0702**	**97**	C4	**1.4520**
98	S1	**2.9916**	**99**	C1	**0.5996**
100	C2	**2.1417**	**101**	C8	**2.4895**
102	C9	**2.6635**	**103**	H1	**1.0268**
104	N1	**0.2869**	**105**	O2	**0.6478**
106	O3	**0.8011**	**107**	E	**1.6609**
C3-O1-O2 Düzlemine Atomun Dik Uzaklığı (Å)					
108	C5	**1.0572**	**109**	O4	**1.9269**
110	N2	**0.8638**	**111**	H3	**0.0410**
112	C7	**1.4155**	**113**	C4	**0.4183**
114	S1	**1.3729**	**115**	C1	**0.0231**
116	C2	**1.3734**	**117**	C8	**1.5860**
118	C9	**2.5280**	**119**	H1	**0.0161**
120	N1	**1.0118**	**121**	C6	**1.8556**
122	O3	**2.5756**	**123**	E	**3.4286**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
O1-O2-O3 Düzlemine Atomun Dik Uzaklığı (Å)					
124	C5	**2.8076**	**125**	O4	**2.0083**
126	N2	**2.4280**	**127**	H3	**3.1915**
128	C7	**1.2506**	**129**	C4	**1.4620**
130	S1	**3.0979**	**131**	C1	**1.1456**
132	C2	**2.6833**	**133**	C8	**3.4398**
134	C9	**3.1394**	**135**	H1	**0.5110**
136	N1	**0.5976**	**137**	C6	**0.5066**
138	C3	**0.3595**	**139**	E	**3.5904**
N1-C6-O3 Düzlemine Atomun Dik Uzaklığı (Å)					
140	C5	**0.8394**	**141**	O4	**0.2942**
142	N2	**1.0454**	**143**	H3	**1.9458**
144	C7	**0.0292**	**145**	C4	**0.1254**
146	S1	**1.7064**	**147**	C1	**1.0019**
148	C2	**2.2316**	**149**	C8	**3.3676**
150	C9	**2.7374**	**151**	H1	**0.1275**
152	C3	**0.3041**	**153**	O2	**0.6095**
154	O1	**0.4796**	**155**	E	**3.4004**
***C6-O3-C3 Düzlemine Atomun Dik Uzaklığı + Van der Waals (Å)**					
160	C5(O4)	**3.1021**	**161**	O4(O4)	**0.7885**
162	N2(H3)	**2.8148**	**163**	H3(H3)	**2.6770**
164	C7(H9)	**1.7737**	**165**	C4(H19)	**1.8162**
166	C2(H7)	**4.8523**	**167**	C8(H7)	**4.9374**
168	C9(H4)	**4.1624**	**169**	H1(H1)	**0.9323**
170	O2(H1)	**1.7666**	**171**	O1(O1)	**1.5606**
172	E(E)	**4.8370**			

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
***C6-O3-O1 Düzlemine Atomun Dik Uzaklığı + Van der Waals (Å)**					
173	C5(O4)	**2.3925**	**174**	O4(O4)	**1.5260**
175	N2(H3)	**2.3860**	**176**	H3(H3)	**2.3262**
177	C7(H9)	**1.8727**	**178**	C4(H19)	**1.8626**
179	C2(H7)	**5.2434**	**180**	C8(H7)	**5.2112**
181	C9(H4)	**4.8037**	**182**	H1(H1)	**1.4034**
183	O2(H1)	**2.7356**	**184**	C3(O1)	**2.7228**
185	E(E)	**3.3402**			
***C6-C3-O1 Düzlemine Atomun Dik Uzaklığı + Van der Waals (Å)**					
186	C5(O4)	**4.6245**	**187**	O4(O4)	**2.8679**
188	N2(H3)	**3.6699**	**189**	H3(H3)	**3.3038**
190	C7(H9)	**2.7864**	**191**	C4(H19)	**3.1622**
192	C2(H7)	**4.9326**	**193**	C8(H7)	**4.1884**
194	C9(H4)	**4.3627**	**195**	H1(H1)	**1.6268**
196	O2(H1)	**2.1999**	**197**	O3(O3)	**1.5611**
198	E(E)	**2.5109**			
***C3-O1-O2 Düzlemine Atomun Dik Uzaklığı + Van der Waals (Å)**					
199	C5(O4)	**3.0376**	**200**	O4(O4)	**2.6869**
201	N2(H3)	**2.4656**	**202**	H3(H3)	**0.6410**
203	C7(H9)	**3.1317**	**204**	C4(H19)	**2.1285**
205	C2(H7)	**4.1643**	**206**	C8(H7)	**3.2849**
207	C9(H4)	**4.2271**	**208**	H1(H1)	**0.6161**
209	C6(O3)	**3.8122**	**210**	O3(O3)	**3.3356**
211	E(E)	**4.2786**			
***O1-O2-O3 Düzlemine Atomun Dik Uzaklığı + Van der Waals (Å)**					
212	C5(O4)	**4.7880**	**213**	O4(O4)	**2.7683**
214	N2(H3)	**4.0297**	**215**	H3(H3)	**3.7915**
216	C7(H9)	**2.9668**	**217**	C4(H19)	**3.1722**
218	C2(H7)	**5.4742**	**219**	C8(H7)	**5.1386**
220	C9(H4)	**4.8385**	**221**	H1(H1)	**1.1110**
222	C6(O3)	**2.4632**	**223**	C3(O1)	**2.3340**
224	E(E)	**4.4404**			

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
	***N1-C6-O3 Düzlemine Atomun Dik Uzaklığı + Van der Waals (Å)**				
225	C5(O4)	**2.8198**	**226**	O4(O4)	**1.0542**
227	N2(H3)	**2.6472**	**228**	H3(H3)	**2.5458**
229	C7(H9)	**1.7453**	**230**	C4(H19)	**1.8356**
231	C2(H7)	**5.0225**	**232**	C8(H7)	**5.0664**
233	C9(H4)	**4.4365**	**234**	H1(H1)	**0.7275**
235	C3(O1)	**2.2786**	**236**	O2(H1)	**2.1617**
237	E(E)	**4.2504**			
	C6-O3-C3 Düzlemi İle Doğru Arasındaki Açı (Derece)				
240	C5 O4	**63.6009**	**241**	N2 H3	**59.5927**
242	N2 C7	**52.2543**	**243**	N1 C1	**37.7372**
244	C4 N1	**7.7286**	**245**	S1 C1	**17.7028**
246	C2 C8	**50.4839**	**247**	N1 S1	**39.9488**
248	C8 N1	**81.3964**	**249**	O1 O2	**27.7231**
250	H1 O2	**35.0472**	**251**	N1 E	**23.7300**
252	C4 E	**22.6430**	**253**	S1 E	**12.8658**
	C6-O3-O1 Düzlemi İle Doğru Arasındaki Açı (Derece)				
254	C5 O4	74.8601	**255**	N2 H3	**70.1067**
256	N2 C7	**40.0758**	**257**	N1 C1	**49.9071**
258	C4 N1	**0.2953**	**259**	S1 C1	**10.0752**
260	C2 C8	**43.9953**	**261**	N1 S1	**36.3923**
262	C8 N1	**86.3649**	**263**	C3 O2	**18.7201**
264	H1 O2	**23.5298**	**265**	N1 E	**13.3598**
266	C4 E	**13.4089**	**267**	S1 E	**3.9807**
	C6-C3-O1 Düzlemi İle Doğru Arasındaki Açı (Derece)				
268	C5 O4	**26.0650**	**269**	N2 H3	**39.3877**
270	N2 C7	**43.0691**	**271**	N1 C1	**12.2096**
272	C4 N1	**50.1594**	**273**	S1 C1	**59.3018**
274	C2 C8	**13.1752**	**275**	N1 S1	**82.6920**
276	C8 N1	**40.7459**	**277**	O2 O3	**1.7897**
278	H1 O2	**23.4614**	**279**	N1 E	**7.7791**
280	C4 E	**1.1874**	**281**	S1 E	**7.2926**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
C3-O1-O2 Düzlemi İle Doğru Arasındaki Açı (Derece)					
282	C5 O4	**45.4503**	**283**	N2 H3	**64.5897**
284	N2 C7	**22.1813**	**285**	N1 C1	**41.9609**
286	C4 N1	**23.0267**	**287**	S1 C1	**30.1219**
288	C2 C8	**8.0098**	**289**	N1 S1	**60.9896**
290	C8 N1	**50.3379**	**291**	C6 O3	**36.9895**
292	C6 H1	**19.9010**	**293**	N1 E	**13.7729**
294	C4 E	**17.3741**	**295**	S1 E	**27.2597**
O1-O2-O3 Düzlemi İle Doğru Arasındaki Açı (Derece)					
296	C5 O4	**40.9118**	**297**	N2 H3	**49.6570**
298	N2 C7	**53.6737**	**299**	N1 C1	**21.7550**
300	C4 N1	**34.7274**	**301**	S1 C1	**44.5731**
302	C2 C8	**29.7189**	**303**	N1 S1	**66.4836**
304	C8 N1	**57.3773**	**305**	C6 C3	**2.2380**
306	C3 H1	**27.1683**	**307**	N1 E	**17.1469**
308	C4 E	**12.1885**	**309**	S1 E	**2.6926**
N1-C6-O3 Düzlemi İle Doğru Arasındaki Açı (Derece)					
310	C5 O4	**68.2643**	**311**	N2 H3	**64.0042**
312	N2 C7	**47.3360**	**313**	N1 C1	**42.6537**
314	C4 N1	**4.7409**	**315**	S1 C1	**14.6702**
316	C2 C8	**48.1105**	**317**	N1 S1	**38.7399**
318	C8 N1	**86.3008**	**319**	C3 O1	**40.1888**
320	H1 O2	**30.4146**	**321**	N1 E	**19.5712**
322	C4 E	**18.9577**	**323**	S1 E	**9.2996**
C6-O3-C3 Düzlemi İle Doğru Arasındaki Açı (Radyan)					
330	C5 O4	**1.1100**	**331**	N2 H3	**1.0401**
332	N2 C7	**0.9120**	**333**	N1 C1	**0.6586**
334	C4 N1	**0.1349**	**335**	S1 C1	**0.3090**
336	C2 C8	**0.8811**	**337**	N1 S1	**0.6972**
338	C8 N1	**1.4206**	**339**	O1 O2	**0.4839**
340	H1 O2	**0.6117**	**341**	N1 E	**0.4142**
342	C4 E	**0.3952**	**343**	S1 E	**0.2246**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
C6-O3-O1 Düzlemi İle Doğru Arasındaki Açı (Radyan)					
344	C5 O4	**1.3066**	**345**	N2 H3	**1.2236**
346	N2 C7	**0.6995**	**347**	N1 C1	**0.8710**
348	C4 N1	**0.0052**	**349**	S1 C1	**0.1758**
350	C2 C8	**0.7679**	**351**	N1 S1	**0.6352**
352	C8 N1	**1.5074**	**353**	C3 O2	**0.3267**
354	H1 O2	**0.4107**	**355**	N1 E	**0.2332**
356	C4 E	**0.2340**	**357**	S1 E	**0.0695**
C6-C3-O1 Düzlemi İle Doğru Arasındaki Açı (Radyan)					
358	C5 O4	**0.4549**	**359**	N2 H3	**0.6874**
360	N2 C7	**0.7517**	**361**	N1 C1	**0.2131**
362	C4 N1	**0.8754**	**363**	S1 C1	**1.0350**
364	C2 C8	**0.2299**	**365**	N1 S1	**1.4432**
366	C8 N1	**0.7111**	**367**	O2 O3	**0.0312**
368	H1 O2	**0.4095**	**369**	N1 E	**0.1358**
370	C4 E	**0.0207**	**371**	S1 E	**0.1273**
C3-O1-O2 Düzlemi İle Doğru Arasındaki Açı (Radyan)					
372	C5 O4	**0.7933**	**373**	N2 H3	**1.1273**
374	N2 C7	**0.3871**	**375**	N1 C1	**0.7324**
376	C4 N1	**0.4019**	**377**	S1 C1	**0.5257**
378	C2 C8	**0.1398**	**379**	N1 S1	**1.0645**
380	C8 N1	**0.8786**	**381**	C6 O3	**0.6456**
382	C6 H1	**0.3473**	**383**	N1 E	**0.2404**
384	C4 E	**0.3032**	**385**	S1 E	**0.4758**
O1-O2-O3 Düzlemi İle Doğru Arasındaki Açı (Radyan)					
386	C5 O4	**0.7140**	**387**	N2 H3	**0.8667**
388	N2 C7	**0.9368**	**389**	N1 C1	**0.3797**
390	C4 N1	**0.6061**	**391**	S1 C1	**0.7779**
392	C2 C8	**0.5187**	**393**	N1 S1	**1.1604**
394	C8 N1	**1.0014**	**395**	C6 C3	**0.0391**
396	C3 H1	**0.4742**	**397**	N1 E	**0.2993**
398	C4 E	**0.2127**	**399**	S1 E	**0.0470**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
	N1-C6-O3 Düzlemi İle Doğru Arasındaki Açı (Radyan)				
400	C5 O4	**1.1914**	**401**	N2 H3	**1.1171**
402	N2 C7	**0.8262**	**403**	N1 C1	**0.7444**
404	C4 N1	**0.0827**	**405**	S1 C1	**0.2560**
406	C2 C8	**0.8397**	**407**	N1 S1	**0.6761**
408	C8 N1	**1.5062**	**409**	C3 O1	**0.7014**
410	H1 O2	**0.5308**	**411**	N1 E	**0.3416**
412	C4 E	**0.3309**	**413**	S1 E	**0.1623**
	Bağ Dereceleri				
420	C5 O4	**1.8187**	**421**	N2 H3	**0.9451**
422	N2 C7	**0.9990**	**423**	C7 C4	**0.9418**
424	C4 N1	**0.9665**	**425**	N1 C6	**0.9351**
426	C6 O3	**1.9529**	**427**	N1 C1	**0.9696**
428	S1 C4	**0.9766**	**429**	S1 C2	**0.9535**
430	C1 C3	**0.9121**	**431**	C3 O1	**1.8330**
432	C3 O2	**1.0466**	**433**	O2 H1	**0.9150**
	FUKUİ ATOMİK NÜKLEOFİLİK REAKTİVİTE İNDEKSİ (eV)				
440	C5	**0.0039**	**441**	O4	**0.0034**
442	N2	**0.0011**	**443**	H3	**0.0000**
444	C7	**0.0002**	**445**	C4	**0.0000**
446	S1	**1.6535**	**447**	C2	**0.0000**
448	C1	**0.0000**	**449**	N1	**0.0001**
450	C6	**0.0001**	**451**	O3	**0.0001**
452	C3	**0.0000**	**453**	O1	**0.0000**
454	O2	**0.0000**	**455**	H1	**0.0000**
456	C8	**0.0000**			

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
		FUKUİ ATOMİK ELEKTROFİLİK REAKTİVİTE İNDEKSİ (eV)			
460	C5	**0.0263**	**461**	O4	**0.0128**
462	N2	**0.0052**	**463**	H3	**0.0026**
464	C7	**0.0218**	**465**	C4	**0.2132**
466	S1	**1.7860**	**467**	C2	**0.2676**
468	C1	**0.0019**	**469**	N1	**0.0021**
470	C6	**0.0024**	**471**	O3	**0.0016**
472	C3	**0.0005**	**473**	O1	**0.0001**
474	O2	**0.0002**	**475**	H1	**0.0002**
476	C8	**0.0060**			
		MULLİKEN YÜKÜ (e^-)			
480	C5	**0.2142**	**481**	O4	**-0.3506**
482	N2	**-0.0019**	**483**	H3	**0.0994**
484	C7	**-0.1039**	**485**	C4	**-0.1586**
486	S1	**-0.0162**	**487**	C2	**-0.0931**
488	C1	**-0.0296**	**489**	N1	**-0.1104**
490	C6	**0.2737**	**491**	O3	**-0.2471**
492	C3	**0.3836**	**493**	O1	**-0.3623**
494	O2	**-0.3078**	**495**	H1	**0.2304**
496	C8	**-0.1219**			
		ELEKTROSTATİK YÜK (e^-)			
500	C5	**0.7359**	**501**	O4	**-0.5633**
502	N2	**-0.5294**	**503**	H3	**0.2374**
504	C7	**-0.0336**	**505**	C4	**0.5300**
506	S1	**-0.3982**	**507**	C2	**0.0783**
508	C1	**0.1600**	**509**	N1	**-0.7611**
510	C6	**0.7304**	**511**	O3	**-0.4535**
512	C3	**0.9285**	**513**	O1	**-0.6047**
514	O2	**-0.6976**	**515**	H1	**0.4124**
516	C8	**0.0117**			

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
		VALENS YÜKÜ (e^-)			
520	C5	**3.8554**	**521**	O4	**2.0488**
522	N2	**3.2607**	**523**	H3	**0.9901**
524	C7	**3.9335**	**525**	C4	**3.9293**
526	S1	**2.0541**	**527**	C2	**3.9411**
528	C1	**3.9396**	**529**	N1	**3.1036**
530	C6	**3.8283**	**531**	O3	**2.1644**
532	C3	**3.8318**	**533**	O1	**2.0471**
534	O2	**2.1418**	**535**	H1	**0.9469**
536	C8	**3.9675**			
		MESAFE (Å)			
540	O4 C6	**4.1523**	**541**	O4 O3	**4.7844**
542	H3 O2	**6.3911**	**543**	N2 O1	**6.0493**
544	O3 O2	**4.9088**	**545**	C6 C3	**3.7672**
546	C6 O2	**4.6408**	**547**	S1 N1	**2.7268**
548	S1 C3	**3.7089**	**549**	S1 C6	**3.4825**
550	O3 O1	**5.0751**	**551**	O3 C3	**4.3224**
552	N1 C6	**1.4815**	**553**	N1 O3	**2.4253**
554	N1 C3	**2.4301**	**555**	N1 O1	**2.9340**
556	N1 O2	**3.5187**	**557**	N1 H1	**4.2151**
558	H1 C6	**5.4042**	**559**	H1 O3	**5.6933**
560	C4 C6	**2.1563**	**561**	O3 C4	**3.3475**
562	C6 E	**8.7837**	**563**	O3 E	**8.2921**
564	C3 E	**12.2093**	**565**	O1 E	**13.0353**
566	O2 E	**12.7400**	**567**	H1 E	**13.6492**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
	*** MESAFE + VAN DER WAALS (Å)**				
570	O4(O4) C6(O3)	**6.8689**	**571**	O4(O4) O3(O3)	**6.3044**
572	H3(H3) O2(H1)	**8.5432**	**573**	N2(H3) O1(O1)	**8.4111**
574	O3(O3) O2(H1)	**7.2210**	**575**	C6(O3) C3(O1)	**7.6982**
576	C6(O3) O2(H1)	**8.1496**	**577**	O3(O3) O1(O1)	**6.5951**
578	O3(O3) C3(O1)	**7.0568**	**579**	N1 C6(O3)	**3.4381**
580	N1 O3(O3)	**3.1852**	**581**	N1 C3(O1)	**4.4046**
582	N1 O1(O1)	**3.6940**	**583**	N1 O2(H1)	**5.0709**
584	N1 H1(H1)	**4.8151**	**585**	H1(H1) C6(O3)	**7.9608**
586	H1(H1) O3(O3)	**7.0533**	**587**	C4 C6(O3)	**4.1129**
588	O3(O3) C4	**4.1075**	**589**	S1 C3(O1)	**5.6834**
590	S1 C6(O3)	**5.4391**	**591**	C6(O3) E(E)	**11.5903**
592	O3(O3) E(E)	**9.9021**	**593**	C3(O1) E(E)	**15.0338**
594	O1(O1) E(E)	**14.6453**	**595**	O2(H1) E(E)	**15.1421**
596	H1(H1) E(E)	**15.0992**			
	AÇILAR (DERECE)				
600	C6-O3-C3	**55.0684**	**601**	C6-O3-O1	**48.3373**
602	C6-O3-O2	**70.1393**	**603**	C6-O3-H1	**70.0898**
604	C3-O1-O3	**46.0913**	**605**	C3-O1-O2	**33.9331**
606	C3-O1-H1	**58.8507**	**607**	O2-O1-C6	**82.9583**
608	O2-O1-O3	**73.1455**	**609**	H1-O1-O2	**24.9198**
610	N1-O3-O1	**20.6806**	**611**	N1-C6-O3	**129.5026**
612	N1-C3-O1	**102.0051**	**613**	N1-O2-H1	**131.9452**
614	H3-N2-N1	**86.0027**	**615**	O4-C5-O1	**99.2054**
616	S1-N1-C6	**108.0064**	**617**	S1-N1-C3	**91.7935**
618	S1-C4-O3	**120.4196**	**619**	S1-C4-O2	**84.6779**

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
		AÇILAR (RADYAN)			
625	C6-O3-C3	**0.9611**	**626**	C6-O3-O1	**0.8436**
627	C6-O3-O2	**1.2242**	**628**	C6-O3-H1	**1.2233**
629	C3-O1-O3	**0.8044**	**630**	C3-O1-O2	**0.5922**
631	C3-O1-H1	**1.0271**	**632**	O2-O1-C6	**1.4479**
633	O2-O1-O3	**1.2766**	**634**	H1-O1-O2	**0.4349**
635	N1-O3-O1	**0.3609**	**636**	N1-C6-O3	**2.2602**
637	N1-C3-O1	**1.7803**	**638**	N1-O2-H1	**2.3029**
639	H3-N2-N1	**1.5010**	**640**	O4-C5-O1	**1.7315**
641	S1-N1-C6	**1.8851**	**642**	S1-N1-C3	**1.6021**
643	S1-C4-O3	**2.1017**	**644**	S1-C4-O2	**1.4779**
650	Boltzmann Dist.(aq)	**0.1889**	**651**	Boltzmann Dist.	**0.1889**
652	CFD (NEG)	**1.0000**	**653**	CFD (POS)	**1.0000**
654	CFD(HBA)	**7.0000**	**655**	CFD(HBD)	**2.0000**
656	PSA (Å²)	**83.5610**	**657**	rel. Eaq (kcal/mol)	**0.0000**
658	Eaq (kcal/mol)	**-217.2308**	**659**	E (kcal/mol)	**-158.0860**
660	rel. E (kcal/mol)	**0.0000**			
		Dihedral Açısı (Derece)			
661	O4-C5-N2-H3	**158.5168**	**662**	C5-N2-H3-N7	**-144.5531**
663	H3-N2-C7-C6	**71.3726**	**664**	C7-C6-O3-N1	**178.3432**
665	C4-N1-C6-O3	**-175.2564**	**666**	C1-N1-C6-O3	**-53.6265**
667	C6-O3-C3-O1	**-51.1957**	**668**	C6-O3-O2-H1	**-88.4561**
669	C3-O1-O2-H1	**179.0190**	**670**	C1-C3-O2-H1	**178.0254**
671	O3-N1-O1-O2	**-42.0435**	**672**	S1-C4-N1-C6	**116.6310**
673	C2-C1-C3-O1	**-78.9587**			

Tablo 4.2'nin devamı

Özellik No	Parametreler	Sayısal değerleri	Özellik No	Parametreler	Sayısal değerleri
		Dihedral Açısı (Radyan)			
674	O4-C5-N2-H3	**2.7666**	**675**	C5-N2-H3-N7	**-2.5229**
676	H3-N2-C7-C6	**1.2457**	**677**	C7-C6-O3-N1	**3.1127**
678	C4-N1-C6-O3	**-3.0588**	**679**	C1-N1-C6-O3	**-0.9360**
680	C6-O3-C3-O1	**-0.8935**	**681**	C6-O3-O2-H1	**-1.5438**
682	C3-O1-O2-H1	**3.1245**	**683**	C1-C3-O2-H1	**3.1071**
684	O3-N1-O1-O2	**-0.7338**	**685**	S1-C4-N1-C6	**2.0356**
686	C2-C1-C3-O1	**-1.3781**			
700	C6 Atomunun Mulliken Yükü	**0.2737**	**701**	C3Atomunun Mulliken Yükü	**0.3836**
702	O1 Atomunun Mulliken Yükü	**-0.3623**	**703**	O2 Atomunun Mulliken Yükü	**-0.3078**
704	O3 Atomunun Mulliken Yükü	**-0.2471**	**705**	C6-C3 Atomları Arasındaki Mesafe	**3.7672**
706	C6-O1 Atomları Arasındaki Mesafe	**4.3721**	**707**	C6-O2 Atomları Arasındaki Mesafe	**4.6408**
708	C6-O3 Atomları Arasındaki Bağ Derecesi	**1.9529**	**709**	C3-O1 Atomları Arasındaki Bağ Derecesi	**1.8330**
710	C3-O2 Atomları Arasındaki Bağ Derecesi	**1.0466**	**711**	C3-O3 Atomları Arasındaki Mesafe	**4.3224**
712	O1-O2 Atomları Arasındaki Mesafe	**2.1821**	**713**	O1-O3 Atomları Arasındaki Mesafe	**5.0751**
714	O2-O3 Atomları Arasındaki Mesafe	**4.9088**			

Not 1: * A(B) C(D) ve buna benzer gösterimlerde parantez içinde belirtilen atomlar kendilerinden önceki atomlarla ilişkilendirilir ve parantez içindeki atomun van der waals yarıçapı hesaplamaya katılmıştır.

Not 2: E ile gösterilen atom R substituentinin en uzak atomu simgelemektedir. 73. Bileşikte Cl1 atomunu simgeler.

Not 3: Tabloda 0.0000 şeklinde görünen sayısal değerin sıfır olduğu anlamına gelmemmektedir sadece sayının 4 haneye yuvarlama sonucunda olmaktadır.

Penisilin bileşik serisinde 88 bileşiğin herbir konformeri için belirlenen değişkenler hazırlandıktan sonra Matlab 7.0 paket programında yazılan kod ile eşitlik 3.3'deki formüle göre aktivite hesaplamaları yapıldı. Aktivite hesaplamaları bileşiklerin yüzde bağlanma (percent bound drug, PAB) değerlerine göre yapılmaktadır. PAB değeri biyolojik veya biyokimyasal fonksiyonların engellenmesinde bir bileşiğin etkinliğinin bir ölçümü olarak tarif edilmektedir.

4.3. Penisilin Serinin Aktivite Hesaplaması

Penisilin serisi için hazırlanan 646 parametre içerisinden aktivite üzerine en çok etkisi olan parametrelerin seçimi için genetik algoritma optimizasyon tekniği kullanılarak yapılmıştır. En iyi parametreler ve bu parametrelere ait κ_j değerleri hesaplandıktan sonra aktivite hesaplamaları eşitlik 3.3 kullanılarak yapılmıştır. Kullanılan bu denklem lineer olmadığı için bu denklemin çözümü lsqnonlin fonksiyonu ile gerçekleştirilmiştir. Penisilin seri için yapılan hesaplamalar ilerleyen bölümlerde açıklanmıştır. Aktivite hesaplamalarında kullanılan parametre sayısının az, R^2 değerinin maksimum, standart hata değerinin minimum ve q^2 değerinin maksimum olduğu durum en uygun durumdur. Bu nedenle çalışmamızda en yüksek R^2, en düşük standart hata ve q^2 değerlerinin maksimum olduğu sonuçlar hem kimyasal açıdan hem de istatistiksel açıdan en uygun sonuçlar olduğuna karar verilmiştir. Taramalar sonucunda parametre sayısı arttıkça R^2 ve q^2 değerlerinde bir artış olduğu gözlenmektedir. Ayrıca Penisilin serisinde oluşturulan modelin etkinliğini arttırmak ve modelin performansını ölçmek için 2 farklı bileşik setinde hesaplamalar yapılarak modelleme geliştirilmiştir. Bu hesaplamalara ilişkin sonuçlar aşağıda verilmektedir.

4.3.1. Penisilin Serisinin Eğitim ve Test Seti Olarak Hesaplanması

Bir Modelini geliştirilmesinin en önemli parçası verilerin eğitim ve test seti şeklinde ayrılmasıdır. İncelenen serilerin eğitim ve test seti şeklinde sınıflandırılmasının nedeni geliştirilen optimizasyon modelinin biyolojik aktivite hesaplamalarında deneye ihtiyaç duyulmadan aktivitesi bilinmeyen bileşiklerin aktivite hesaplamalarını tahmin etmek ve modelin performansını ölçmektir. Aktivite değerlerinin sınıflandırılması ya da tahmin edilebilmesinde modelin oluşturulması için kullanılan veri setinin bir kısmı eğitim seti olarak tarif edilir. Eğitim seti farklı modelleri değerlendirmek için kullanılan test seti ile

bağlantılı olarak kullanılır. Verilerin büyük kısmını eğitim seti ve küçük kısmını ise test seti oluşturmaktadır. Eğitim ve test seti için benzer veriler kullanılarak veri farklılıklarının etkisi minimize edilebilir ve modelin daha iyi anlanması sağlanır. QSAR çalışmalarında eğitim ve test seti için sınıflandırma oranları ortalama olarak sırasıyla %70, %30 olarak alınmaktadır. Çalışmalarımızda bu oranlar dikkate alınarak hesaplamalar yapıldığında oluşturulan modelin uygunluğu regresyon katsayısı ve q^2 ile belirlenir. Eğitim ve test seti için bileşik sayıları belirlendikten sonra yazılan Matlab kodlarıyla bu bileşikler rastgele seçilerek hesaplamalar başlatılmıştır. Hazırlanan parametreler arasından aktivite üzerine en çok etki eden parametrelerin seçimi genetik algoritma optimizasyon tekniği kullanılarak gerçekleştirilmiştir. En iyi parametrelerin ve bu parametreler karşılık gelen kappa değerlerinin bulunmasıyla işleme başlanmıştır. Başlangıç değeri sıfır olacak şekilde işlem başlamış ve çok geniş bir uzayda tarama yapılmıştır. En iyi değerin bulunması için çok sayıda (ortalama 2000) iterasyon işlemi yapılmıştır. Regresyon katsayısı ve q^2 nin maksimum olduğu eğitim seti bizim için en uygun olandır. Bu şekilde öğrenmesi sağlanan eğitim seti aktivitesi bilinmeyen bileşiklerin aktivitelerinin tahmin edilebilmesi için test edilmelidir. Bu amaç için eğitim setinde olmayan bileşikler kullanılır. Eğitim seti için bulunan moleküler parametreler ve kappa değerleri değiştirilmeden aynı aktivite formülü (eşitlik 3.3) kullanılarak test seti için yeni hesaplamalar yapılır. Test seti için de regresyon katsayısı ve q^2 nin maksimum olduğu hesaplama sonucu en iyi değerdir. Bu şekilde eğitim setinin uygunluğu test edilmiş olur.

Ayrıca hesaplama sonuçlarının doğruluğunu test etmek için tek bir hesaplama seti ile değil 2 farklı hesaplama seti ile çalışılmıştır. Aşağıda sırasıyla bu setlerin sonuçları yer almaktadır.

4.3.1.1. Penisilin Serisinde 60 Bileşik Eğitim ve 28 Bileşik Test Seti Olarak Aktivite Hesaplaması

Penisilin serisindeki bileşikler rastgele (random) 60 bileşik eğitim, 28 bileşik test seti olarak seçildikten sonra bileşik seti sabit tutulup 1 parametreden (1 κ_j) 13 parametreye (13 κ_j) kadar taramalar yapılarak elde edilen hesaplamalara ilişkin sonuçlar aşağıda tablo 4.3 ve 4.4'de verilmektedir. 13 paramatreden sonrası için taramalar yapılmıştır fakat sonuçlarda önemli bir değişiklik olmadığı için 13 parametreden sonraki

hesaplamalar verilmemiştir. Ayrıca literatürlerde baş parmak (Thumb approach) kuralına göre, iyi bir model elde etmek için data seti (eğitim seti) seçilen parametre sayısının yaklaşık 5-7 katı olmalıdır yani [92] aktivite hesaplamalarında kullanılan parametre sayısının az olması gerektiğinden 13 parametreden sonrası için olan hesaplama sonuçları tabloda verilmemiştir.

Tablolarda yıldızla simgelenen bileşikler test bileşikleridir. Aşağıdaki tablo 4.3'de 1 parametreden (1 κ_j) 7 parametreye (7 κ_j) kadar eğitim ve test seti için hesaplanan aktivite değerleri, regresyon katsayısı, standart hata değerleri ve q^2 değerleri verilmiştir.

Tablo 4.3. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için elde edilen 1-7 κj değerlerine göre hesaplanan aktivite değerleri.

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**
1	14.013	15.957	15.283	15.186	15.602	15.539	15.375
2*	16.983	18.526	22.050	18.614	10.616	11.492	11.907
3	16.983	17.159	15.994	17.456	13.560	13.399	12.465
4*	21.533	23.344	25.919	21.157	16.531	17.120	17.235
5	14.579	15.295	13.936	15.681	11.734	11.633	10.432
6*	24.385	26.261	25.038	21.907	23.589	22.184	20.789
7	20.008	22.001	22.103	19.757	21.999	21.229	22.030
8	31.447	33.541	33.782	26.839	18.604	18.707	17.342
9	16.983	18.671	18.935	17.915	19.326	19.099	18.215
10*	13.700	15.443	18.538	15.100	17.177	17.395	17.680
11	20.235	26.579	22.151	17.682	14.218	13.445	13.192
12	24.942	26.252	32.693	23.705	20.771	20.843	20.870
13*	32.349	38.608	52.678	29.022	38.784	36.726	38.794
14	35.211	39.389	36.311	35.393	41.123	39.521	42.245
15	33.276	33.970	36.068	37.847	45.184	44.491	44.323
16*	36.633	43.960	41.829	38.045	29.835	30.850	32.366
17	48.323	50.945	51.800	45.484	46.619	46.526	52.760
18	22.915	24.940	25.209	24.771	29.558	29.199	30.626
19	28.245	31.482	30.693	28.554	36.222	34.148	36.996
20*	20.815	22.748	26.783	25.768	39.669	39.974	39.387
21	35.211	38.460	37.093	33.854	24.984	25.302	25.536
22	41.483	47.186	49.722	43.934	63.679	61.916	67.678
23	20.008	22.104	25.833	24.503	32.866	35.482	40.028
24	40.786	45.088	42.702	37.717	47.153	43.460	40.168
25	44.900	44.782	54.738	41.088	51.893	48.992	48.423
26*	44.395	45.507	46.955	52.781	65.892	65.156	60.528
27	41.250	56.098	76.911	69.779	64.429	63.942	62.648
28	66.318	66.197	61.066	77.832	78.848	77.095	78.901
29*	48.323	52.513	52.527	57.495	58.307	61.540	70.551
30	47.243	55.178	46.557	52.212	59.307	57.873	58.311

Tablo 4.3'ün devamı

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**
31	54.107	63.590	57.935	61.677	65.310	68.078	80.620
32*	41.719	52.630	46.345	53.572	55.299	59.480	71.148
33	66.318	73.547	58.241	68.735	67.515	65.756	60.648
34	56.610	66.949	57.977	64.999	62.250	65.264	79.123
35	81.284	82.350	74.426	73.215	66.323	65.767	77.640
36	42.914	46.127	46.327	52.866	52.739	56.291	60.431
37	51.715	53.756	51.582	60.709	53.396	56.350	64.155
38	114.749	118.015	100.871	107.152	82.759	81.308	92.510
39	50.274	57.435	55.456	57.913	44.779	47.590	50.768
40*	70.972	72.847	72.245	77.311	89.656	86.001	89.470
41	55.658	59.453	57.517	59.153	63.209	63.692	67.394
42	60.926	65.085	64.694	58.855	70.009	68.228	72.125
43	42.914	47.997	46.134	54.631	54.327	58.456	61.447
44*	30.057	33.585	39.968	47.408	61.894	69.727	80.955
45	76.384	72.389	75.608	79.320	78.858	78.376	87.299
46	56.610	95.466	70.289	75.018	94.197	92.175	81.308
47	56.610	58.873	55.804	59.268	63.886	62.960	66.479
48	86.988	91.980	77.467	89.411	69.744	70.874	82.069
49	51.715	59.200	52.465	63.908	58.191	61.668	70.272
50*	49.708	46.370	43.818	59.502	54.780	54.839	56.960
51	65.202	70.350	68.547	65.237	96.083	89.693	88.484
52*	98.507	95.496	89.682	94.255	84.445	84.427	102.824
53*	36.221	34.936	48.521	49.667	89.826	92.382	99.878
54*	127.759	142.980	111.931	109.932	110.494	99.544	108.493
55*	42.914	44.802	43.430	62.903	67.790	72.017	73.464
56	76.384	74.577	76.859	94.291	95.635	101.693	104.363
57	79.917	87.405	71.172	89.730	84.923	81.203	83.501
58	64.469	62.771	59.153	80.275	76.997	80.265	83.357
59*	207.736	201.175	159.662	124.574	133.111	104.048	106.887
60	102.483	98.027	78.109	101.325	85.936	80.714	80.424
61*	91.012	99.410	89.244	120.690	106.926	114.393	120.295
62*	124.903	134.218	120.557	140.810	133.724	136.816	144.232
63*	91.012	81.050	93.817	96.270	91.628	92.181	92.392
64	70.972	82.507	74.090	77.484	77.574	79.540	97.548
65	127.759	123.609	114.435	113.348	110.184	105.390	108.146
66*	86.988	91.064	80.932	98.693	81.583	84.768	103.139
67*	70.972	62.694	70.734	87.501	75.213	79.376	84.956
68	97.400	82.904	94.117	106.972	105.763	105.983	110.769
69	104.827	107.332	106.807	105.933	115.253	115.238	126.012
70	124.903	137.341	122.218	121.896	117.020	115.264	121.815
71	97.400	106.146	96.508	98.487	102.317	102.353	112.495
72	114.103	109.765	110.589	106.318	102.241	99.769	103.845
73	97.400	97.400	97.400	97.400	97.400	97.400	97.400
74	91.012	71.900	93.588	96.260	88.005	88.244	90.484

Tablo 4.3'ün devamı

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**
75*	21.533	24.110	26.599	33.977	27.102	31.213	31.965
76	26.693	27.649	29.041	34.727	27.221	29.046	28.183
77	64.835	32.834	35.662	26.598	25.276	23.230	25.005
78	20.008	21.219	31.995	19.389	36.229	35.272	35.573
79*	42.914	43.399	44.908	41.205	42.504	41.959	46.384
80	39.204	52.519	44.148	42.473	51.295	50.170	43.341
81	48.051	53.509	45.828	48.002	50.887	48.087	49.590
82	22.657	25.007	27.416	31.260	40.586	43.340	39.878
83	39.204	41.540	40.700	49.175	44.642	48.176	52.252
84*	98.507	124.584	94.468	77.845	70.476	66.410	75.677
85	36.426	39.840	41.905	36.576	51.090	48.595	56.923
86*	49.708	62.996	103.694	72.487	102.288	99.464	94.772
87*	68.219	59.557	80.088	63.972	76.623	72.835	90.137
88	7.483	0.815	1.574	1.791	1.941	1.975	1.913
$R^2_{eğitim}$	0.580	0.646	0.660	0.717	0.752	0.774	0.790
$S.Hata_{eğitim}$	0.085	0.078	0.077	0.070	0.065	0.063	0.060
R^2_{test}	0.392	0.381	0.507	0.621	0.721	0.761	0.790
$S.Hata_{test}$	0.153	0.154	0.138	0.121	0.104	0.096	0.090
q^2	-0.131	0.383	0.389	0.460	0.551	0.547	0.593
q^2_{ext1}	-0.452	-0.469	0.210	0.407	0.597	0.680	0.690
q^2_{ext2}	-0.468	-0.485	0.201	0.401	0.592	0.677	0.686
Not: "*" ile simgelenen bileşikler test bileşikleridir.							

Aşağıdaki tablo 4.4'de 8 parametreden (8 κ_j) 13 parametreye (13 κ_j) kadar eğitim ve test seti için hesaplanan aktivite değerleri, regresyon katsayısı, standart hata değerleri ve q^2 değerleri verilmiştir.

Tablo 4.4. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için elde edilen 8-13κj değerlerine göre hesaplanan aktivite değerleri.

Bileşikler	**8 κ_j**	**9 κ_j**	**10 κ_j**	**11 κ_j**	**12 κ_j**	**13 κ_j**	**PAB_{den}**
1	15.048	15.300	15.016	14.975	15.098	15.269	7.200
2*	11.932	10.296	11.798	10.997	8.165	7.476	12.000
3	11.554	9.399	11.018	10.741	9.414	9.674	15.000
4*	19.033	15.092	18.290	17.932	13.256	12.488	16.800
5	9.724	8.297	9.229	9.003	8.368	8.920	18.000
6*	22.923	18.449	22.337	21.828	18.992	20.417	20.000
7	25.521	21.763	26.796	27.299	21.898	21.363	25.000
8	16.749	15.483	17.111	16.444	15.746	14.949	26.000
9	17.714	17.082	17.240	17.395	17.141	18.030	28.000
10*	19.966	19.075	19.752	19.734	18.568	17.798	32.000
11	14.990	12.562	14.268	14.160	12.113	10.600	33.000

Tablo 4.4'ün devamı

Bileşikler	8 κ_j	9 κ_j	10 κ_j	11 κ_j	12 κ_j	13 κ_j	PAB_{den}
12	21.426	16.407	21.182	20.692	15.073	15.903	38.000
13*	35.312	32.852	28.479	26.064	28.967	30.467	42.000
14	42.690	44.751	45.669	46.029	43.844	41.171	47.000
15	41.086	44.779	41.608	42.859	47.538	49.843	53.200
16*	34.694	32.886	33.046	32.840	29.690	27.040	55.000
17	54.361	43.315	50.996	50.008	42.705	41.010	57.000
18	31.983	33.690	33.805	34.725	36.773	35.304	58.000
19	39.544	40.536	43.677	44.100	42.796	41.401	58.800
20*	42.796	47.029	46.022	48.159	48.246	48.629	59.000
21	27.699	25.284	29.398	29.119	25.751	20.620	60.000
22	67.253	56.672	54.979	56.481	55.126	57.824	60.000
23	46.692	51.781	46.817	48.920	54.203	52.891	61.700
24	41.224	37.487	41.845	41.602	40.885	40.846	62.000
25	46.399	47.585	50.651	51.053	53.901	49.680	63.000
26*	62.817	67.601	62.059	66.023	68.775	74.356	65.000
27	76.245	56.630	75.850	71.656	57.553	62.964	66.200
28	73.581	76.362	77.299	77.922	83.107	81.036	68.000
29*	73.188	71.761	69.644	70.170	78.161	82.231	69.700
30	61.229	64.820	62.258	65.078	65.219	63.668	74.000
31	89.598	94.923	85.201	86.701	89.084	81.699	74.500
32*	68.703	90.296	74.255	74.062	87.655	87.121	77.000
33	58.550	62.435	53.953	54.362	62.042	64.409	78.000
34	74.412	96.589	77.809	77.529	88.725	84.976	80.000
35	83.896	80.854	86.402	85.714	92.580	89.884	80.000
36	55.131	68.215	59.306	60.110	65.249	63.457	81.500
37	60.300	68.896	63.381	62.527	68.292	68.660	81.500
38	88.071	76.183	86.419	84.143	89.402	95.345	82.100
39	51.115	61.963	47.592	51.422	59.104	55.433	82.200
40*	91.878	75.742	82.254	86.249	76.054	82.695	82.500
41	64.580	71.645	68.378	69.484	71.972	67.462	83.000
42	69.736	69.703	69.321	70.254	74.586	74.001	83.000
43	55.411	68.970	55.924	57.431	70.232	70.500	83.500
44*	86.298	97.671	85.719	88.291	84.750	85.893	83.600
45	93.852	98.954	92.510	95.654	85.103	87.553	84.000
46	73.033	95.284	73.192	75.015	96.073	93.881	84.000
47	63.190	77.559	65.498	66.814	81.937	81.833	86.000
48	80.269	75.141	77.371	76.197	87.479	89.460	86.000
49	69.438	81.683	71.731	72.188	79.746	79.340	86.100
50*	57.574	59.229	58.365	58.780	64.169	71.341	87.000
51	92.249	96.936	96.787	91.586	100.001	100.095	88.000
52*	108.697	119.239	110.541	106.799	122.865	122.245	89.300
53*	107.729	107.631	108.169	110.821	92.361	93.849	89.700
54*	96.566	92.762	100.729	96.901	96.799	91.740	91.000
55*	79.999	97.734	79.869	84.044	84.904	88.115	91.500

Tablo 4.4'ün devamı

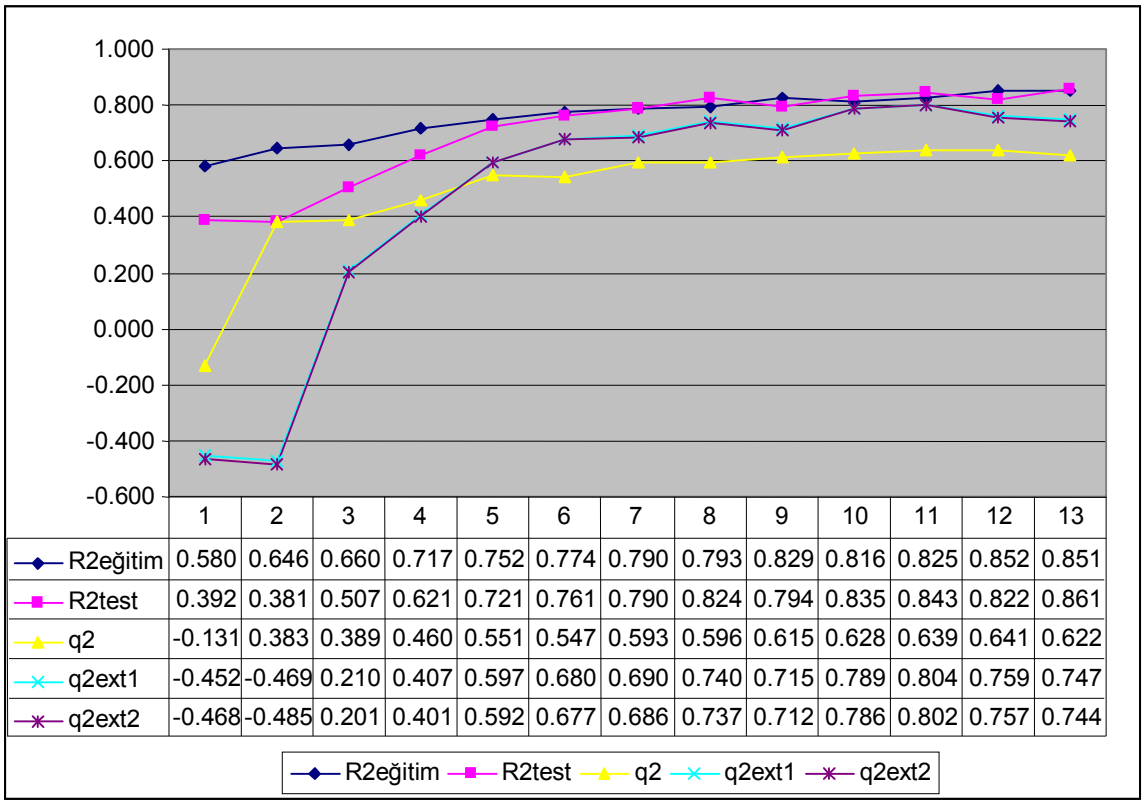
Bileşikler	**8 κ_j**	**9 κ_j**	**10 κ_j**	**11 κ_j**	**12 κ_j**	**13 κ_j**	**PAB_{den}**
56	100.680	95.706	102.244	91.335	84.454	81.098	92.000
57	88.912	83.598	96.664	97.375	91.817	87.382	92.400
58	97.399	89.964	97.136	99.869	88.333	88.208	92.500
59*	106.653	90.861	106.808	101.817	99.601	105.364	93.300
60	80.408	72.348	76.172	76.556	73.816	76.830	93.600
61*	110.114	115.496	101.026	101.001	99.662	99.488	94.000
62*	131.801	122.318	120.732	117.733	117.732	121.766	94.000
63*	95.083	92.290	95.974	92.741	109.834	122.523	94.700
64	91.339	108.969	96.034	92.094	97.962	94.038	94.800
65	109.310	89.857	100.800	99.998	85.237	89.364	95.200
66*	107.461	102.152	106.021	103.726	108.359	122.714	95.600
67*	86.529	88.268	85.927	84.664	86.386	91.161	96.000
68	117.006	93.603	108.571	108.852	92.213	102.741	96.000
69	113.158	118.269	107.987	111.039	112.130	110.041	96.500
70	111.371	93.894	101.504	97.743	91.946	91.493	97.000
71	101.138	110.218	102.876	102.325	100.150	100.539	97.000
72	99.773	94.300	101.617	101.672	89.229	92.868	97.200
73	97.400	97.400	97.400	97.400	97.400	97.400	97.400
74	96.236	90.409	98.813	99.837	88.783	92.184	97.400
75*	34.139	38.060	35.851	35.701	33.596	32.131	18.000
76	27.044	29.374	26.140	26.654	27.998	26.688	18.000
77	24.777	19.853	25.237	25.476	18.986	22.036	20.000
78	34.562	39.847	35.439	35.419	42.311	41.604	30.000
79*	51.704	45.181	54.383	53.808	46.482	47.616	39.000
80	40.722	37.954	34.867	33.674	39.155	40.644	50.000
81	48.425	47.153	50.977	50.105	50.958	48.704	60.000
82	42.520	47.350	39.785	41.538	47.576	50.744	65.000
83	48.489	58.178	51.890	51.556	60.574	61.416	80.000
84*	73.384	64.665	73.060	70.211	84.486	92.105	89.000
85	60.053	59.465	63.169	63.007	66.064	66.524	92.000
86*	113.427	67.988	96.947	96.792	67.608	97.312	95.000
87*	103.082	100.702	93.393	93.823	81.693	89.335	96.000
88	2.038	1.940	2.008	1.930	1.962	1.854	2.000
$R^2_{eğitim}$	0.793	0.829	0.816	0.825	0.852	0.851	
$S.Hata_{eğitim}$	0.060	0.054	0.056	0.055	0.051	0.051	
R^2_{test}	0.824	0.794	0.835	0.843	0.822	0.861	
$S.Hata_{test}$	0.082	0.089	0.080	0.078	0.083	0.073	
q^2	0.596	0.615	0.628	0.639	0.641	0.622	
q^2_{ext1}	0.740	0.715	0.789	0.804	0.759	0.747	
Not: "*" ile simgelenen bileşikler test bileşikleridir.							

Tablo 4.5'de Penisilin serisinde 60 bileşik eğitim, 28 bileşik test seti için aktivite hesaplamalarında aktiviteye en fazla etkisi olan parametre numaraları ile buna karşılık gelen κ_j değerleri görülmektedir.

Tablo 4.5. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için hesaplamalardan elde edilen parametreler ve κj değerleri.

κ_j adedi	Parametre No	Kj değeri
1	31	-0.565
2	31, 133	-0.547, 8.736
3	24, 31, 375	-0.119, -0.441, 6.050
4	31, 375, 446, 451	-0.386, 4.534, -0.283, 57.144
5	24, 31, 375, 446, 653	-0.260, -0.254, 6.233, -0.262, 0.360
6	24, 31, 375, 446, 451, 653	-0.254, -0.203, 6.085, -0.332, 27.355, 0.344
7	24, 31, 375, 446, 451, 653, 704	-0.283, -0.195, 6.807, -0.347, 43.951, 0.347, 28.582
8	24, 31, 341, 375, 446, 451, 653, 704	-0.308, -0.151, -0.348, 7.151, -0.329, 38.916, 0.427, 37.046
9	2, 17, 24, 81, 375, 446, 451, 653, 704	0.435, -0.003, -0.388, -5.733, 7.145, -0.533, 57.790, 0.491, 33.053
10	2, 24, 31, 341, 375, 442, 446, 451, 653, 704	0.390, -0.348, -0.109, -0.304, 6.754, -10.843, -0.352, 100.000, 0.534, 37.776
11	2, 17, 24, 31, 341, 375, 442, 446, 451, 653, 704	0.380, -0.002, -0.362, -0.085, -0.338, 6.686, -10.699, -0.373, 100.000, 0.534, 37.183
12	2, 17, 24, 40, 43, 81, 375, 442, 446, 451, 653, 704	0.520, -0.004, -0.304, 0.012, -0.004, -6.048, 6.166, -9.536, -0.508, 100.000, 0.584, 24.981
13	2, 17, 24, 40, 43, 81, 375, 442, 443, 446, 653, 704, 712	0.460, -0.004, -0.257, 0.018, -0.005, -5.397, 7.660, 5.153, -100.000, -0.512, 0.628, 22.348, 10.625

Penisilin serisin de hem eğitim hem de test setine ait tüm hesaplamalar için deneysel değerlere karşılık gelen hesaplanan aktivite değerleri yukarıdaki tablo 4.3 ve 4.4'de verilmiştir. Şekil 4.4'de Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği verilmiştir. Bu hesaplamalardan görüldüğü üzere belli bir parametre sayısına kadar regresyon katsayısının ve q^2 değerlerinin artığı görülmektedir. Hem baş parmak kuralına göre hemde şekil 4.4'de yapılan hesaplamalara ilişkin sonuçlarda penisilin serisinde eğitim ve test seti için 12 parametre sayısından sonra istatistiksel hesaplamalarda önemli bir değişiklik olmadığı için en uygun parametre sayısı olarak 12 parametre belirlenmiştir.



	1	2	3	4	5	6	7	8	9	10	11	12	13
R2eğitim	0.580	0.646	0.660	0.717	0.752	0.774	0.790	0.793	0.829	0.816	0.825	0.852	0.851
R2test	0.392	0.381	0.507	0.621	0.721	0.761	0.790	0.824	0.794	0.835	0.843	0.822	0.861
q2	-0.131	0.383	0.389	0.460	0.551	0.547	0.593	0.596	0.615	0.628	0.639	0.641	0.622
q2ext1	-0.452	-0.469	0.210	0.407	0.597	0.680	0.690	0.740	0.715	0.789	0.804	0.759	0.747
q2ext2	-0.468	-0.485	0.201	0.401	0.592	0.677	0.686	0.737	0.712	0.786	0.802	0.757	0.744

Şekil 4.4. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği.

Penisilin serisi için tablo 5.2'de verilen 12 parametre içerisinde eğitim setinde aktiviteye katkısı en fazla olan parametreyi bulmak için bir parametre dışarıda bırakılarak (leave one out) hesaplamaları yapılmıştır. Yani 12 parametrenin içinde bir tanesi dışarıda bırakılarak 11 parametre için eşitlik 3.3 kullanılarak hesaplamalar yapılmış ve R^2 değerleri literatürlere benzer şekilde hesaplanmıştır [72]. Ayrıca E istatistik değerleri hesaplanarak 12 parametre içerisinde hangi parametrenin aktivite üzerinde daha etkili olduğunu bulunmuştur. E ve q^2 değelerindeki düşüş $a^{(7)}$ ile simgelenen 375 nolu parametrenin aktivite üzerindeki etkisinin fazla olduğunu, E ve q^2 değelerindeki artış $a^{(5)}$ ile simgelenen 43 nolu parametrenin aktivite üzerindeki etkisinin en az olduğunu göstermektedir. Diğer parametrelerin etkileri tablo 4.6'da görülmektedir.

Tablo 4.6. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için leave one out hesaplama sonuçları.

ELENEN PARAMETRELER	İndeks	Parametreler	E	$R^2_{eğitim}$	R^2_{test}	q^2	q^2_{ext1}	q^2_{ext2}
	$a^{(1)}$	2	0.966	0.837	0.817	0.628	0.728	0.725
	$a^{(2)}$	17	0.865	0.828	0.796	0.585	0.714	0.711
	$a^{(3)}$	24	0.708	0.795	0.723	0.493	0.518	0.513
	$a^{(4)}$	40	0.951	0.838	0.815	0.622	0.746	0.743
	$a^{(5)}$	**43**	**1.045**	**0.849**	**0.810**	**0.656**	**0.757**	**0.755**
	$a^{(6)}$	81	0.894	0.828	0.781	0.598	0.692	0.689
	$a^{(7)}$	**375**	**0.204**	**0.636**	**0.599**	**-0.759**	**0.176**	**0.167**
	$a^{(8)}$	442	0.868	0.842	0.814	0.587	0.738	0.736
	$a^{(9)}$	446	0.337	0.652	0.549	-0.066	0.241	0.233
	$a^{(10)}$	451	0.689	0.821	0.842	0.479	0.776	0.774
	$a^{(11)}$	653	0.517	0.761	0.548	0.305	0.428	0.422
	$a^{(12)}$	704	0.915	0.842	0.808	0.608	0.743	0.740

Aktivite değerleri arasında çok büyük farklılıklar olduğundan hesaplamaları lineerleştirmek için bu tür çalışmalarda aktivite değerinin logaritması alınarak yapılan hesaplamalar yaygın olarak kullanılmaktadır. Bunun için tablo 4.3 ve 4.4'de penisilin serisindeki bileşiklerden elde edilen 60 bileşik eğitim ve 28 bileşik test setinin aynısı aşağıdaki tablo 4.7 ve 4.8'de verilen hesaplamalarda da kullanılmıştır. Bu hesaplamada penisilin bileşik serisinin deneysel aktivite değerlerinin logaritması (log PAB) alınarak hesaplamalar yapılmıştır. Aşağıdaki tablo 4.7'de penisilin serisindeki 1 parametreden 7 parametreye kadar taramalar yapılarak eğitim ve test seti için hesaplanan aktivite değerleri, regresyon katsayısı, standart hata değerleri ve q^2 değerleri verilmiştir.

Tablo 4.7. Penisilin serisi 60 bileşik eğitim, 28 bileşik test için elde edilen 1–7 κj değerlerine göre hesaplanan logaritma cinsinden aktivite değerleri.

Bileşikler	1 κ_j	2 κ_j	3 κ_j	4 κ_j	5 κ_j	6 κ_j	7 κ_j
1	1.315	1.322	1.373	1.312	1.279	1.249	1.241
2*	1.370	1.385	1.326	1.253	1.237	1.333	1.323
3	1.370	1.375	1.297	1.234	1.192	1.054	1.051
4*	1.441	1.442	1.361	1.272	1.264	1.362	1.343
5	1.326	1.339	1.242	1.186	1.136	0.942	0.935
6*	1.480	1.473	1.488	1.415	1.392	1.448	1.388
7	1.418	1.435	1.502	1.424	1.409	1.459	1.472
8	1.562	1.547	1.518	1.362	1.388	1.491	1.429
9	1.370	1.380	1.437	1.368	1.344	1.354	1.356
10*	1.308	1.307	1.541	1.485	1.486	1.577	1.538

Tablo 4.7'nin devamı

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**
11	1.422	1.436	1.352	1.290	1.259	1.305	1.301
12	1.487	1.519	1.549	1.403	1.535	1.529	1.531
13*	1.572	1.587	1.483	1.575	1.669	1.459	1.437
14	1.600	1.606	1.614	1.639	1.728	1.691	1.726
15	1.581	1.595	1.531	1.631	1.698	1.655	1.719
16*	1.614	1.623	1.609	1.559	1.658	1.642	1.639
17	1.712	1.720	1.777	1.664	1.704	1.779	1.801
18	1.460	1.460	1.652	1.615	1.612	1.689	1.697
19	1.527	1.539	1.558	1.585	1.663	1.609	1.619
20*	1.430	1.444	1.791	1.808	1.808	1.892	1.898
21	1.600	1.583	1.505	1.463	1.559	1.549	1.536
22	1.657	1.681	1.519	1.634	1.564	1.782	1.780
23	1.418	1.405	1.806	1.698	1.634	1.791	1.764
24	1.651	1.659	1.714	1.726	1.707	1.809	1.752
25	1.686	1.672	1.773	1.834	1.709	1.757	1.775
26*	1.682	1.682	1.785	1.937	1.847	1.970	1.880
27	1.655	1.849	1.743	1.884	1.893	1.911	1.870
28	1.832	1.846	1.707	1.927	1.930	1.913	1.942
29*	1.712	1.725	1.786	1.796	1.759	1.861	1.886
30	1.704	1.705	1.672	1.809	1.827	1.836	1.815
31	1.754	1.750	1.826	1.843	1.827	1.887	1.846
32*	1.659	1.677	1.778	1.808	1.882	1.877	1.877
33	1.832	1.805	1.766	1.896	1.942	1.928	1.852
34	1.771	1.779	1.781	1.790	1.931	1.877	1.888
35	1.913	1.883	1.818	1.773	1.791	1.868	1.821
36	1.669	1.678	1.714	1.722	1.851	1.807	1.826
37	1.737	1.743	1.705	1.717	1.842	1.800	1.831
38	2.059	2.076	1.865	1.851	1.878	1.954	1.977
39	1.727	1.726	1.720	1.763	1.785	1.766	1.784
40*	1.859	1.891	2.023	2.157	2.018	2.151	2.177
41	1.765	1.769	1.844	1.851	1.852	1.921	1.944
42	1.799	1.805	1.923	1.898	1.897	1.952	1.963
43	1.669	1.680	1.718	1.747	1.843	1.822	1.850
44*	1.547	1.579	1.960	2.005	1.970	2.027	2.074
45	1.888	1.889	2.023	2.040	2.071	2.036	1.935
46	1.771	1.641	1.634	1.787	1.702	1.799	1.833
47	1.771	1.762	1.757	1.822	1.859	1.889	1.926
48	1.941	1.946	1.790	1.799	1.790	1.892	1.920
49	1.737	1.752	1.712	1.759	1.858	1.829	1.871
50*	1.723	1.708	1.552	1.734	1.755	1.672	1.699
51	1.825	1.814	2.016	2.172	2.281	2.072	2.003
52*	1.993	1.967	1.880	1.889	1.881	1.945	1.884
53*	1.610	1.625	2.117	2.273	2.236	2.100	2.138
54*	2.107	2.130	1.930	2.045	2.115	2.090	2.054

Tablo 4.7'nin devamı

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**
55*	1.669	1.663	1.708	1.953	1.951	1.941	1.919
56	1.888	1.873	1.978	2.087	2.082	2.040	1.961
57	1.906	1.936	1.694	1.921	1.862	1.895	1.916
58	1.821	1.796	1.745	1.933	1.897	1.945	1.880
59*	2.338	2.336	1.885	2.003	2.100	2.057	2.010
60	2.010	1.976	1.671	1.902	1.917	1.875	1.853
61*	1.960	1.993	1.967	2.115	2.194	2.119	2.128
62*	2.097	2.122	2.103	2.211	2.202	2.145	2.136
63*	1.960	1.987	1.949	1.950	1.837	2.011	2.015
64	1.859	1.880	1.864	1.866	1.882	1.939	1.945
65	2.107	2.092	2.028	2.075	2.115	2.009	1.999
66*	1.941	1.950	1.796	1.853	1.825	1.929	1.935
67*	1.859	1.879	1.878	1.939	1.933	1.990	1.992
68	1.989	1.997	2.011	2.125	2.143	2.076	2.055
69	2.020	2.030	2.136	2.158	2.206	2.015	2.050
70	2.097	2.135	2.116	2.112	2.031	2.095	2.094
71	1.989	1.993	1.986	2.020	2.051	2.002	2.011
72	2.057	2.066	1.985	1.973	2.011	2.014	1.975
73	1.989	1.989	1.989	1.989	1.989	1.989	1.989
74	1.960	1.997	1.965	1.956	1.848	2.025	2.027
75*	1.441	1.448	1.367	1.498	1.420	1.501	1.522
76	1.508	1.500	1.458	1.570	1.505	1.583	1.574
77	1.823	1.448	1.377	1.360	1.391	1.302	1.238
78	1.418	1.406	1.760	1.736	1.726	1.467	1.503
79*	1.669	1.663	1.695	1.621	1.677	1.736	1.699
80	1.637	1.673	1.568	1.645	1.726	1.697	1.727
81	1.710	1.725	1.644	1.718	1.766	1.738	1.752
82	1.457	1.457	1.544	1.654	1.583	1.690	1.678
83	1.637	1.641	1.632	1.642	1.716	1.711	1.725
84*	1.993	2.022	1.938	1.799	1.845	1.916	1.920
85	1.612	1.625	1.704	1.746	1.713	1.792	1.813
86*	1.723	1.920	2.149	2.025	2.068	2.079	2.042
87*	1.843	1.880	2.029	1.963	2.015	1.857	1.857
88	1.150	0.298	0.298	0.300	0.300	0.303	0.305
$R^2_{eğitim}$	0.568	0.754	0.749	0.795	0.807	0.869	0.880
$S.Hata_{eğitim}$	0.086	0.065	0.066	0.059	0.058	0.048	0.046
R^2_{test}	0.510	0.539	0.722	0.774	0.814	0.784	0.784
$S.Hata_{test}$	0.137	0.133	0.103	0.093	0.085	0.091	0.091
q^2	-0.546	0.681	0.646	0.721	0.730	0.787	0.802
q^2_{ext1}	0.433	0.484	0.724	0.746	0.786	0.734	0.747
q^2_{ext2}	0.428	0.479	0.721	0.744	0.784	0.731	0.744
Not: "*" ile simgelenen bileşikler test bileşikleridir.							

Aşağıdaki tablo 4.8'de penisilin serisindeki 8 parametreden 13 parametreye kadar

taramalar yapılarak eğitim ve test seti için hesaplanan aktivite değerleri, regresyon katsayısı, standart hata değerleri ve q^2 değerleri verilmiştir.

Tablo 4.8. Penisilin serisi 60 bileşik eğitim, 28 bileşik test için elde edilen 8–13 κj değerlerine göre hesaplanan logaritma cinsinden aktivite değerleri.

Bileşikler	8 κ_j	9 κ_j	10 κ_j	11 κ_j	12 κ_j	13 κ_j	$LogPAB_{den}$
1	1.225	1.238	1.226	1.206	1.231	1.221	0.857
2*	1.363	1.272	1.262	1.291	1.270	1.246	1.079
3	1.072	1.036	1.054	1.071	1.047	1.063	1.176
4*	1.404	1.338	1.324	1.351	1.344	1.327	1.225
5	0.962	0.934	0.940	1.012	0.953	0.989	1.255
6*	1.376	1.403	1.383	1.350	1.384	1.339	1.301
7	1.460	1.427	1.458	1.486	1.434	1.387	1.398
8	1.458	1.427	1.415	1.414	1.422	1.402	1.415
9	1.339	1.336	1.336	1.282	1.329	1.321	1.447
10*	1.481	1.561	1.526	1.510	1.563	1.598	1.505
11	1.358	1.320	1.306	1.276	1.330	1.335	1.519
12	1.537	1.569	1.560	1.509	1.588	1.609	1.580
13*	1.387	1.523	1.713	1.520	1.525	1.564	1.623
14	1.720	1.732	1.766	1.752	1.744	1.714	1.672
15	1.742	1.778	1.759	1.716	1.826	1.834	1.726
16*	1.677	1.757	1.713	1.712	1.844	1.932	1.740
17	1.786	1.794	1.774	1.787	1.760	1.783	1.756
18	1.611	1.716	1.678	1.667	1.711	1.771	1.763
19	1.613	1.644	1.676	1.651	1.669	1.653	1.769
20*	1.756	1.899	1.834	1.846	1.850	1.846	1.771
21	1.587	1.673	1.632	1.692	1.714	1.747	1.778
22	1.743	1.643	1.778	1.624	1.681	1.689	1.778
23	1.662	1.766	1.707	1.797	1.734	1.717	1.790
24	1.720	1.791	1.728	1.604	1.761	1.735	1.792
25	1.783	1.779	1.692	1.786	1.818	1.823	1.799
26*	1.819	1.877	1.798	1.756	1.865	1.840	1.813
27	1.913	1.866	1.870	1.847	1.820	1.918	1.821
28	1.926	1.958	1.938	1.888	1.941	1.961	1.833
29*	1.856	1.855	1.805	1.889	1.831	1.878	1.843
30	1.801	1.888	1.838	1.758	1.851	1.816	1.869
31	1.839	1.912	1.855	1.986	1.877	1.887	1.872
32*	1.820	1.889	1.876	1.869	1.857	1.787	1.887
33	1.816	1.926	1.879	1.863	1.871	1.906	1.892
34	1.867	1.946	1.951	2.018	1.947	1.888	1.903
35	1.845	1.895	1.855	1.853	1.898	1.896	1.903
36	1.809	1.845	1.864	1.952	1.863	1.848	1.911
37	1.823	1.905	1.894	1.907	1.920	1.933	1.911
38	1.995	2.006	1.952	2.045	1.951	2.004	1.914
39	1.805	1.863	1.792	1.877	1.920	2.014	1.915
40*	2.083	2.063	2.021	2.068	1.938	2.092	1.917

Tablo 4.8'in devamı

Bileşikler	8 κ_j	9 κ_j	10 κ_j	11 κ_j	12 κ_j	13 κ_j	$LogPAB_{den}$
41	1.906	1.931	1.906	1.961	1.906	1.884	1.919
42	1.914	1.958	1.906	1.943	2.023	1.926	1.919
43	1.829	1.853	1.859	1.832	1.855	1.829	1.922
44*	1.954	2.021	1.992	2.026	2.021	2.043	1.922
45	1.897	1.934	1.932	2.000	1.922	1.923	1.924
46	1.810	1.791	1.714	1.896	1.880	1.872	1.924
47	1.913	1.967	1.928	2.000	1.998	1.937	1.935
48	1.937	1.904	1.880	1.926	1.877	1.924	1.935
49	1.884	1.903	1.929	1.972	1.944	1.948	1.935
50*	1.711	1.712	1.730	1.768	1.744	1.810	1.940
51	1.941	2.027	2.047	2.069	1.994	1.986	1.945
52*	1.902	1.966	1.902	1.988	1.979	1.955	1.951
53*	1.973	2.156	2.074	2.194	2.109	2.171	1.953
54*	2.050	2.086	2.052	2.098	1.990	2.102	1.959
55*	1.886	1.959	1.946	1.998	1.971	1.984	1.961
56	1.914	1.958	1.924	1.985	1.993	1.930	1.964
57	1.918	1.938	1.878	1.981	1.898	1.864	1.966
58	1.898	1.880	1.874	1.865	1.926	1.951	1.966
59*	2.099	2.130	2.098	2.216	2.077	2.024	1.970
60	1.910	1.958	1.921	1.872	1.983	1.964	1.971
61*	2.082	2.114	2.105	2.102	2.050	2.139	1.973
62*	2.084	2.094	2.043	2.112	2.097	2.133	1.973
63*	2.022	2.042	1.933	1.992	2.031	1.975	1.976
64	1.918	1.910	1.920	1.874	1.906	1.881	1.977
65	2.022	2.067	2.021	1.975	2.042	2.026	1.979
66*	1.961	1.820	1.886	2.014	1.897	1.938	1.981
67*	1.988	2.083	2.015	1.982	2.054	2.079	1.982
68	2.042	2.074	2.056	1.969	2.062	2.047	1.982
69	2.003	2.050	2.018	2.047	2.039	2.075	1.985
70	2.060	1.963	1.960	2.007	1.950	1.949	1.987
71	2.012	1.920	1.977	2.120	1.927	1.975	1.987
72	2.009	1.840	1.945	1.980	1.855	1.880	1.988
73	1.989	1.989	1.989	1.989	1.989	1.989	1.989
74	2.035	1.952	1.957	1.943	1.951	1.958	1.989
75*	1.577	1.449	1.484	1.548	1.477	1.426	1.255
76	1.596	1.436	1.539	1.510	1.437	1.423	1.255
77	1.240	1.282	1.263	1.305	1.302	1.295	1.301
78	1.466	1.572	1.529	1.539	1.535	1.497	1.477
79*	1.701	1.689	1.714	1.642	1.697	1.636	1.591
80	1.737	1.751	1.721	1.736	1.682	1.664	1.699
81	1.731	1.731	1.766	1.725	1.696	1.705	1.778
82	1.656	1.658	1.631	1.672	1.638	1.568	1.813
83	1.708	1.694	1.746	1.789	1.683	1.656	1.903
84*	1.949	1.803	1.874	1.963	1.828	1.898	1.949

Tablo 4.8'in devamı

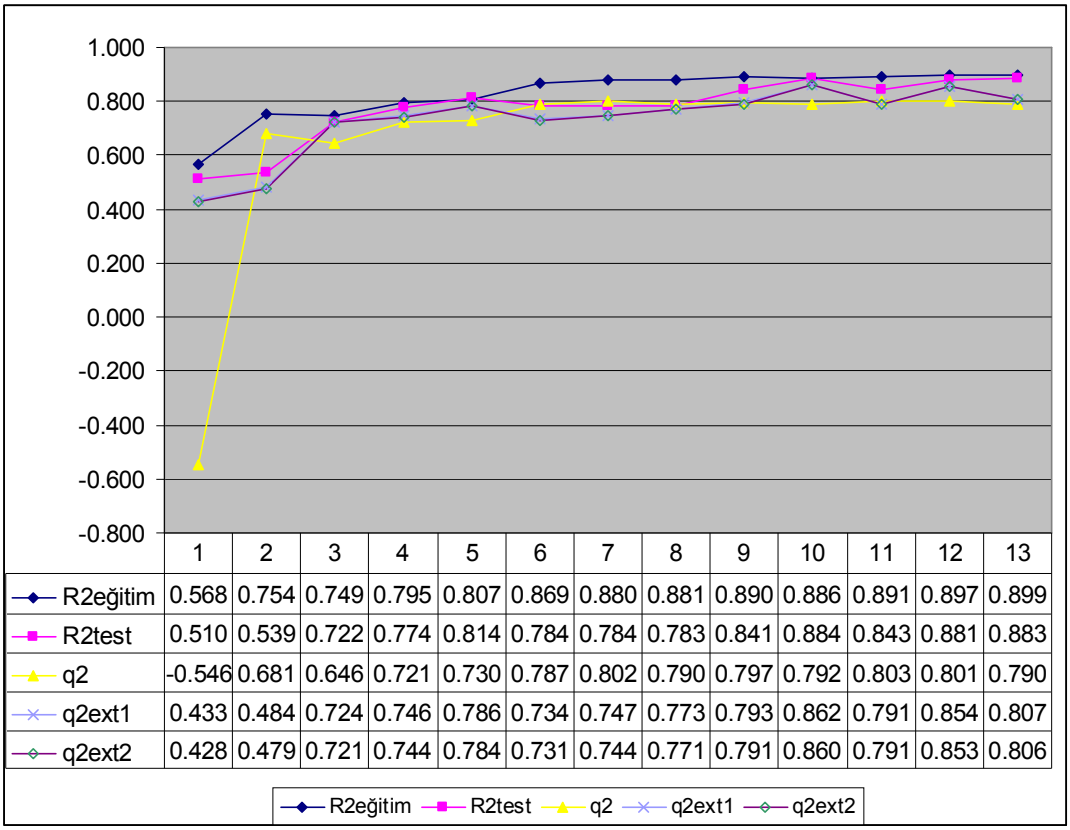
Bileşikler	8 κ_j	9 κ_j	10 κ_j	11 κ_j	12 κ_j	13 κ_j	$LogPAB_{den}$
85	1.770	1.765	1.760	1.890	1.720	1.801	1.964
86*	1.982	2.084	2.004	1.933	1.945	2.167	1.978
87*	1.862	1.992	1.936	1.946	1.939	2.095	1.982
88	0.306	0.303	0.305	0.301	0.300	0.303	0.301
$R^2_{eğitim}$	0.881	0.890	0.886	0.891	0.897	0.899	
$S.Hata_{eğitim}$	0.045	0.044	0.044	0.043	0.042	0.042	
R^2_{test}	0.783	0.841	0.884	0.843	0.881	0.883	
$S.Hata_{test}$	0.091	0.078	0.067	0.078	0.068	0.067	
q^2	0.790	0.797	0.792	0.803	0.801	0.790	
q^2_{ext1}	0.773	0.793	0.862	0.791	0.854	0.807	
q^2_{ext2}	0.771	0.791	0.860	0.789	0.853	0.806	
Not: "*" ile simgelenen bileşikler test bileşikleridir.							

Tablo 4.9'da Penisilin serisinde 60 bileşik eğitim, 28 bileşik test seti için aktivite değeri logaritmik olan hesaplamalarda aktiviteye en fazla etkisi olan parametre numaraları ile buna karşılık gelen κ_j değerleri görülmektedir.

Tablo 4.9. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için aktivite değeri logaritmik olan hesaplamalardan elde edilen parametreler ve κj değerleri.

κ_j adedi	Parametre No	Kj değeri
1	31	-0.121
2	31, 380	-0.120, 2.333
3	23, 380, 656	-0.301, 2.924, 0.006
4	23, 380, 446, 656	-0.274, 2.714, -0.092, 0.004
5	23, 380, 446, 464, 656	-0.331, 2.764, -0.006, -3.570, 0.005
6	15, 23, 27, 380, 446, 656	0.010, -1.570, 0.408, 2.870, -0.053, 0.005
7	15, 23, 27, 369, 380, 446, 656	0.011, -1.617, 0.275, 0.101, 2.851, -0.066, 0.005
8	15, 23, 31, 369, 380, 397, 446, 656	0.010, -1.491, -0.033, 0.237, 2.707, -0.119, -0.061, 0.002
9	15, 23, 344, 380, 397, 446, 464, 466, 656	0.009, -1.390, -0.375, 2.686, 0.243, -0.066, -2.718, 0.057, 0.005
10	15, 23, 27, 344, 369, 380, 446, 464, 656, 704	0.001, -1.490, 0.466, -0.155, 0.159, 2.786, -0.058, -2.258, 0.005, 2.712
11	2, 6, 15, 23, 165, 369, 380, 446, 464, 656, 714	0.608, 0.633, 0.007, -1.120, -1.576, 0.150, 2.339, -0.083, -1.545, 0.004, -0.287
12	15, 23, 56, 344, 369, 380, 446, 464, 466, 634, 656, 709	0.010, -1.468, 0.000*, -0.299, 0.264, 1.752, -0.076, -2.984, 0.070, 0.546, 0.004, 1.189
13	2, 15, 16, 23, 56, 344, 369, 380, 446, 464, 466, 512, 656	-0.005, 0.008, -0.001, -1.304, 0.000**, -0.242, 0.264, 2.746, -0.078, -3.707, 0.078, 0.842, 0.004
* Gerçek değeri -0.000004		**Gerçek Değeri -0.000003

Penisilin serisin aktivite değeri logaritma cinsinden olan hesaplamalarda hem eğitim hem de test setine ait tüm hesaplamalar için deneysel değerlere karşılık gelen hesaplanan aktivite değerleri yukarıda tablo 4.7 ve 4.8'de verilmiştir. Şekil 4.5'de penisilin serisi aktivite değeri logaritmik cinsinden olan 60 bileşik eğitim, 28 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği verilmiştir. Bu hesaplamalardan görüldüğü üzere belli bir parametre sayısına kadar regresyon katsayısının ve q^2'nin yükseldiği görülmektedir. Hem baş parmak kuralına göre hem de şekil 4.5'de yapılan hesaplamalara ilişkin sonuçlarda penisilin serisinde eğitim ve test seti için 12 parametre sayısından sonra istatistiksel hesaplamalarda önemli bir değişiklik olmadığı için en uygun parametre sayısı olarak 12 parametre belirlenmiştir.



	1	2	3	4	5	6	7	8	9	10	11	12	13
R2eğitim	0.568	0.754	0.749	0.795	0.807	0.869	0.880	0.881	0.890	0.886	0.891	0.897	0.899
R2test	0.510	0.539	0.722	0.774	0.814	0.784	0.784	0.783	0.841	0.884	0.843	0.881	0.883
q2	-0.546	0.681	0.646	0.721	0.730	0.787	0.802	0.790	0.797	0.792	0.803	0.801	0.790
q2ext1	0.433	0.484	0.724	0.746	0.786	0.734	0.747	0.773	0.793	0.862	0.791	0.854	0.807
q2ext2	0.428	0.479	0.721	0.744	0.784	0.731	0.744	0.771	0.791	0.860	0.791	0.853	0.806

Şekil 4.5. Penisilin serisi 60 bileşik eğitim, 28 bileşik test seti için aktivite değeri logaritma cinsinden olan hesaplamalarda optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği.

Penisilin serisi için tablo 5.4'de verilen 12 parametre içerisinde eğitim setinde aktiviteye katkısı en fazla olan parametreyi bulmak için bir parametre dışarıda bırakılarak (leave one out) hesaplamaları yapılmıştır. 12 parametrenin içinde 11 parametre hesaplamaları (leave one out) yapılarak R^2 değerleri hesaplanmıştır. E istatistik değerleri de

hesaplanarak 12 parametre içerisinde hangi parametrenin aktivite üzerinde daha etkili olduğu bulunmuştur. E değerindeki düşüş $a^{(11)}$ ile simgelenen 656 nolu parametrenin aktivite üzerindeki etkisinin fazla olduğunu, E ve q^2 değelerindeki artış $a^{(12)}$ ile simgelenen 709 nolu parametrenin aktivite üzerindeki etkisinin en az olduğunu göstermektedir. Diğer parametrelerin etkileri tablo 4.10'da verilmiştir.

Tablo 4.10. Penislin serisi 60 bileşik eğitim, 28 bileşik test için aktivite değeri logaritma cinsinden olan hesaplamalarda leave one out hesaplama sonuçları.

ELENEN PARAMETRELER	İndeks	Parametreler	E	$R^2_{eğitim}$	R^2_{test}	q^2	q^2_{ext1}	q^2_{ext2}
	$a^{(1)}$	15	0.614	0.836	0.827	0.685	0.791	0.789
	$a^{(2)}$	23	0.517	0.811	0.782	0.626	0.745	0.743
	$a^{(3)}$	56	0.934	0.893	0.862	0.793	0.827	0.826
	$a^{(4)}$	344	0.764	0.886	0.817	0.747	0.781	0.779
	$a^{(5)}$	369	0.833	0.876	0.799	0.768	0.762	0.760
	$a^{(6)}$	380	0.494	0.875	0.869	0.608	0.849	0.847
	$a^{(7)}$	446	0.762	0.863	0.801	0.746	0.785	0.783
	$a^{(8)}$	464	0.726	0.892	0.789	0.733	0.756	0.753
	$a^{(9)}$	466	0.885	0.886	0.828	0.781	0.819	0.818
	$a^{(10)}$	634	0.904	0.895	0.873	0.786	0.844	0.843
	$a^{(11)}$	**656**	**0.387**	**0.859**	**0.747**	**0.501**	**0.706**	**0.703**
	$a^{(12)}$	**709**	**1.019**	**0.897**	**0.879**	**0.810**	**0.853**	**0.852**

4.3.1.2. Penisilin Serisinde 74 Bileşik Eğitim ve 13 Bileşik Test Seti Olarak Aktivite Hesaplaması

Bu hesaplama setinde penisilin serisinde 87 bileşik için elde edilen aktivite hesaplama sonuçları verilmiştir. Bileşik serisinin alındığı makaleye göre serisideki ilk 74 bileşik eğitim sonraki 13 bileşik test seti olarak alındı. Sonra bileşik seti sabit tutularak 1 parametreden (1 κ_j) başlayıp 14 parametreye (14 κ_j) kadar taramalar yapılarak elde edilen hesaplamalara ilişkin sonuçlar aşağıda tablo 4.11 ve 4.12'de verilmiştir. 14 paramatreden sonrası için taramalar yapılmıştır fakat sonuçlarda önemli bir değişiklik olmadığı ve aktivite hesaplamalarında kullanılan parametre sayısının az olması gerektiğinden 14 parametreden sonraki hesaplamalar verilmemiştir. Aşağıdaki tablo 4.11'de 1 parametreden (1 κ_j) 8 parametreye (8 κ_j) kadar eğitim ve test seti için

hesaplanan aktivite değerleri, regresyon katsayısı, standart hata değerleri ve q^2 değerleri verilmiştir.

Tablo 4.11. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için elde edilen 1–8 κj değerlerine göre hesaplanan aktivite değerleri.

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**	**8 κ_j**
1	16.501	16.450	15.452	15.117	16.131	15.671	15.889	15.547
2	19.676	19.723	18.856	18.445	14.280	14.849	14.847	13.499
3	19.676	19.578	20.062	15.678	13.281	13.414	13.333	13.990
4	24.454	24.031	23.045	20.686	18.962	19.261	18.504	19.252
5	17.109	17.418	18.150	14.071	11.286	11.609	11.475	11.849
6	27.403	26.653	26.218	20.569	23.234	21.509	22.273	22.290
7	22.862	23.003	21.583	19.519	22.428	21.134	21.540	22.859
8	34.591	33.597	33.974	23.690	21.340	20.923	22.683	23.107
9	19.676	19.775	18.870	17.130	19.862	18.680	19.255	19.743
10	16.162	15.810	16.054	19.982	21.757	21.745	21.328	20.370
11	23.100	23.181	19.232	17.578	16.557	15.683	16.128	14.821
12	27.977	29.278	30.381	23.136	23.680	23.138	23.224	26.613
13	35.498	35.888	33.977	46.584	57.563	54.586	56.154	61.041
14	38.364	37.673	31.803	37.232	43.380	40.516	42.070	49.202
15	36.428	36.737	38.411	50.213	47.941	50.517	47.116	59.176
16	39.779	39.254	35.021	38.483	34.612	35.037	32.744	35.039
17	51.264	50.957	56.960	57.964	46.471	53.380	50.253	52.393
18	25.887	25.404	28.003	30.536	31.792	32.371	30.939	32.503
19	31.351	31.458	30.315	31.748	38.209	35.487	37.338	44.868
20	23.706	23.724	23.216	35.562	43.251	43.687	41.683	42.739
21	38.364	37.144	38.218	37.101	29.903	31.241	28.552	32.833
22	44.577	45.205	40.906	57.199	68.206	67.629	72.341	67.233
23	22.862	22.136	31.533	44.652	36.886	46.736	40.547	33.516
24	43.890	43.882	42.682	49.702	48.656	49.758	46.756	46.367
25	47.927	46.671	47.085	70.830	67.407	71.368	66.483	58.111
26	47.434	46.711	54.671	71.034	69.878	72.812	69.134	65.974
27	44.347	67.202	59.728	64.049	55.643	56.480	55.564	48.557
28	68.501	68.659	75.658	96.286	82.725	86.303	84.287	78.314
29	51.264	50.917	62.449	78.920	58.699	70.917	65.746	57.639
30	50.213	49.250	51.613	57.845	63.774	60.708	63.383	60.802
31	56.855	56.122	80.924	90.484	68.562	82.692	76.458	61.737
32	44.809	44.466	47.043	54.112	56.740	57.884	56.158	55.264
33	68.501	65.181	78.198	78.671	71.645	72.444	71.145	73.056
34	59.259	58.495	53.996	58.228	64.358	61.873	62.895	76.285
35	82.532	78.630	95.047	90.128	67.198	76.492	74.175	65.933
36	45.983	45.728	44.233	45.797	55.386	51.847	56.289	62.065
37	54.549	54.113	52.921	56.749	54.482	55.276	54.534	68.890
38	113.175	114.189	128.918	100.345	77.835	82.266	86.639	92.519
39	53.155	52.451	52.557	61.649	55.066	56.164	57.132	52.006

Tablo 4.11'in devamı

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**	**8 κ_j**
40	72.891	74.099	90.713	85.943	89.851	87.576	94.962	83.436
41	58.346	57.614	73.991	62.851	65.942	65.556	62.359	66.508
42	63.383	62.722	66.569	63.697	72.977	70.302	70.564	72.458
43	45.983	45.768	43.577	49.405	57.575	54.926	55.180	60.419
44	33.188	33.830	31.428	46.834	67.150	63.296	69.153	65.119
45	77.964	78.029	98.212	78.585	79.467	78.404	75.282	79.296
46	59.259	68.732	59.473	91.302	92.054	94.575	90.208	86.949
47	59.259	57.891	68.998	65.344	67.445	66.648	65.171	75.157
48	87.821	86.822	95.630	81.194	68.232	70.319	77.197	76.843
49	54.549	54.624	51.339	60.395	59.888	60.189	60.091	69.393
50	52.608	50.600	59.368	62.239	56.570	56.381	58.347	67.625
51	67.445	66.025	78.019	94.316	105.870	104.297	102.180	107.328
52	98.414	93.960	129.587	109.993	82.657	93.268	91.416	68.518
53	39.370	39.436	48.330	64.501	102.233	93.986	106.037	105.117
54	124.871	124.697	128.692	147.812	104.973	117.330	100.564	95.831
55	45.983	44.818	51.314	87.722	74.181	82.750	76.227	78.239
56	77.964	75.861	109.352	100.601	102.838	104.448	106.501	112.737
57	81.261	83.118	89.453	93.426	85.933	83.783	87.704	84.904
58	66.751	63.748	71.334	105.523	84.433	94.008	85.621	82.520
59	194.890	193.385	161.315	116.416	120.379	103.343	114.962	126.667
60	102.045	97.702	95.846	98.105	88.629	83.634	91.073	98.019
61	91.534	92.683	81.988	106.836	106.261	104.934	103.188	104.691
62	122.312	122.319	116.400	121.879	131.145	124.517	133.884	118.516
63	91.534	94.119	83.884	115.574	86.037	98.810	94.767	70.863
64	72.891	73.439	59.928	69.232	78.601	74.713	79.182	83.720
65	124.871	121.248	107.788	111.145	110.433	107.011	103.092	111.575
66	87.821	86.869	94.455	68.248	79.125	70.252	90.056	91.062
67	72.891	74.057	77.105	80.630	72.829	75.238	71.640	73.189
68	97.400	98.293	91.786	113.790	107.460	109.240	103.215	104.635
69	104.180	103.339	98.296	101.247	118.253	111.615	117.665	124.855
70	122.312	125.211	120.625	102.275	113.523	105.428	113.263	97.950
71	97.400	95.986	92.772	96.142	104.728	101.104	99.373	104.968
72	112.591	113.533	118.932	110.813	99.459	103.209	93.680	94.722
73	97.400	97.400	97.400	97.400	97.400	97.400	97.400	97.400
74	91.534	96.610	89.885	112.317	83.557	95.644	92.368	66.174
75*	24.454	24.313	26.055	42.833	34.049	38.526	36.946	29.393
76*	29.769	29.165	34.077	47.364	34.866	39.129	35.518	34.595
77*	67.097	35.620	6.909	43.249	38.048	37.658	35.442	35.509
78*	22.862	22.131	24.203	44.990	50.794	56.452	45.694	47.047
79*	45.983	44.849	57.238	43.423	41.728	43.077	38.753	42.988
80*	42.329	43.554	43.000	49.293	52.926	52.842	50.507	49.268
81*	50.999	50.795	55.774	58.073	52.175	53.547	51.693	55.856
82*	25.620	25.375	35.770	43.161	46.007	49.002	45.095	47.395
83*	42.329	41.893	55.081	43.867	46.535	45.693	47.395	49.995

Tablo 4.11'in devamı

Bileşikler	**1 κ_j**	**2 κ_j**	**3 κ_j**	**4 κ_j**	**5 κ_j**	**6 κ_j**	**7 κ_j**	**8 κ_j**
84*	98.414	98.675	109.083	79.842	63.616	68.606	67.578	69.694
85*	39.574	39.820	48.720	65.748	54.328	63.943	51.468	51.297
86*	52.608	80.097	95.537	63.105	66.268	68.798	61.778	59.728
87*	70.297	74.431	103.814	74.450	70.881	74.827	69.273	72.568
$R^2_{eğitim}$	0.550	0.550	0.588	0.709	0.763	0.774	0.775	0.769
$S.Hata_{eğitim}$	0.079	0.079	0.076	0.064	0.057	0.056	0.056	0.057
R^2_{test}	0.223	0.507	0.653	0.571	0.727	0.702	0.748	0.748
$S.Hata_{test}$	0.266	0.212	0.178	0.197	0.158	0.165	0.151	0.151
q^2	0.046	0.041	0.104	0.466	0.557	0.596	0.561	0.500
q^2_{ext1}	0.196	0.472	0.650	0.535	0.530	0.578	0.513	0.557
q^2_{ext2}	0.037	0.368	0.581	0.443	0.437	0.495	0.417	0.470
Not: "*" ile simgelenen bileşikler test bileşikleridir.								

Aşağıdaki tablo 4.12'de 9 parametreden (9 κ_j) 14 parametreye (14 κ_j) kadar eğitim ve test seti için hesaplanan aktivite değerleri, regresyon katsayısı, standart hata değerleri ve q^2 değerleri verilmiştir.

Tablo 4.12. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için elde edilen 9–14 κj değerlerine göre hesaplanan aktivite değerleri.

Bileşikler	**9 κ_j**	**10 κ_j**	**11 κ_j**	**12 κ_j**	**13 κ_j**	**14 κ_j**	**PAB_{den}**
1	14.506	14.618	14.102	13.482	13.961	13.837	7.200
2	16.930	15.822	16.323	15.781	15.795	15.697	12.000
3	14.983	12.003	13.089	12.870	12.116	12.268	15.000
4	18.083	19.265	21.179	22.258	20.590	21.033	16.800
5	13.759	10.448	12.238	12.013	11.241	11.401	18.000
6	20.818	19.782	20.089	21.834	19.422	19.650	20.000
7	24.130	20.837	22.859	25.559	22.266	22.478	25.000
8	24.321	19.053	18.699	22.373	18.821	19.584	26.000
9	16.499	16.871	15.757	15.334	15.726	15.626	28.000
10	22.858	20.337	24.148	22.446	22.791	22.847	32.000
11	13.459	16.731	15.470	11.772	14.705	14.137	33.000
12	22.497	25.038	17.631	24.145	21.222	21.814	38.000
13	46.176	53.586	44.361	48.444	50.017	49.653	42.000
14	42.305	50.716	46.014	46.695	50.904	50.974	47.000
15	48.812	65.611	56.187	58.913	63.238	62.964	53.200
16	31.667	44.062	47.043	41.875	48.447	48.802	55.000
17	49.149	51.036	47.621	55.724	53.175	53.317	57.000
18	31.749	33.964	36.414	36.981	35.936	36.118	58.000
19	37.018	44.814	38.503	41.380	43.579	43.573	58.800
20	46.779	38.873	44.009	53.253	44.082	45.168	59.000
21	28.574	40.532	47.946	46.548	49.092	49.962	60.000

Tablo 4.12'nin devamı

Bileşikler	**9 κ_j**	**10 κ_j**	**11 κ_j**	**12 κ_j**	**13 κ_j**	**14 κ_j**	**PAB_{den}**
22	57.272	57.230	61.421	56.961	60.150	60.033	60.000
23	42.871	41.231	45.737	50.966	45.609	47.143	61.700
24	38.712	39.993	36.059	35.140	36.158	36.203	62.000
25	68.448	49.884	66.958	55.977	58.301	56.544	63.000
26	63.875	56.328	58.948	59.765	54.372	55.728	65.000
27	59.290	59.879	55.754	51.717	53.167	53.322	66.200
28	78.788	82.133	73.720	72.797	76.071	75.315	68.000
29	66.395	67.576	68.883	65.902	70.154	69.150	69.700
30	65.467	70.783	59.974	55.005	61.906	60.509	74.000
31	79.665	87.581	91.266	86.256	92.333	92.876	74.500
32	68.812	71.620	52.782	54.711	60.819	59.572	77.000
33	63.106	68.531	68.327	63.265	65.493	67.019	78.000
34	66.279	86.100	86.379	78.639	94.354	91.914	80.000
35	77.684	75.673	77.139	86.200	83.011	84.098	80.000
36	60.226	62.430	61.568	72.846	69.720	71.457	81.500
37	58.431	75.308	65.508	72.639	75.742	75.745	81.500
38	91.239	87.899	88.363	104.271	94.562	96.427	82.100
39	55.489	62.819	60.898	60.029	62.150	62.075	82.200
40	87.299	85.984	85.117	78.855	78.699	78.125	82.500
41	69.033	88.434	89.680	78.008	85.603	83.321	83.000
42	69.902	73.267	80.437	73.725	76.589	75.098	83.000
43	54.443	63.416	64.262	61.680	66.475	66.680	83.500
44	79.293	59.783	71.641	66.880	68.781	67.942	83.600
45	73.966	78.265	94.024	91.560	87.984	87.767	84.000
46	83.628	88.091	77.807	98.277	85.185	83.948	84.000
47	70.558	81.402	83.720	80.215	84.642	83.581	86.000
48	78.473	76.443	74.936	75.692	77.753	77.416	86.000
49	66.081	77.335	67.631	70.801	76.641	75.394	86.100
50	59.643	61.581	65.729	68.387	66.137	66.539	87.000
51	102.947	106.100	92.647	107.340	98.332	99.643	88.000
52	95.847	88.671	90.959	84.018	91.378	87.603	89.300
53	117.559	87.944	109.488	113.723	101.707	101.463	89.700
54	95.758	101.661	98.863	92.533	100.593	99.049	91.000
55	93.855	94.519	109.890	104.153	105.472	109.369	91.500
56	112.095	107.681	101.937	108.661	102.082	105.506	92.000
57	107.238	93.906	90.374	94.540	90.401	90.699	92.400
58	96.471	97.439	107.320	86.002	105.984	103.063	92.500
59	102.137	100.355	111.469	114.001	111.309	110.885	93.300
60	78.695	87.663	83.453	94.190	84.901	86.457	93.600
61	99.006	99.160	96.978	105.606	100.458	100.796	94.000
62	121.219	102.671	101.684	106.584	101.178	100.173	94.000
63	102.005	77.110	87.288	80.434	85.371	82.301	94.700
64	80.949	83.224	69.282	66.827	76.642	73.366	94.800
65	90.475	90.124	98.661	100.736	97.338	98.039	95.200

Tablo 4.12'nin devamı

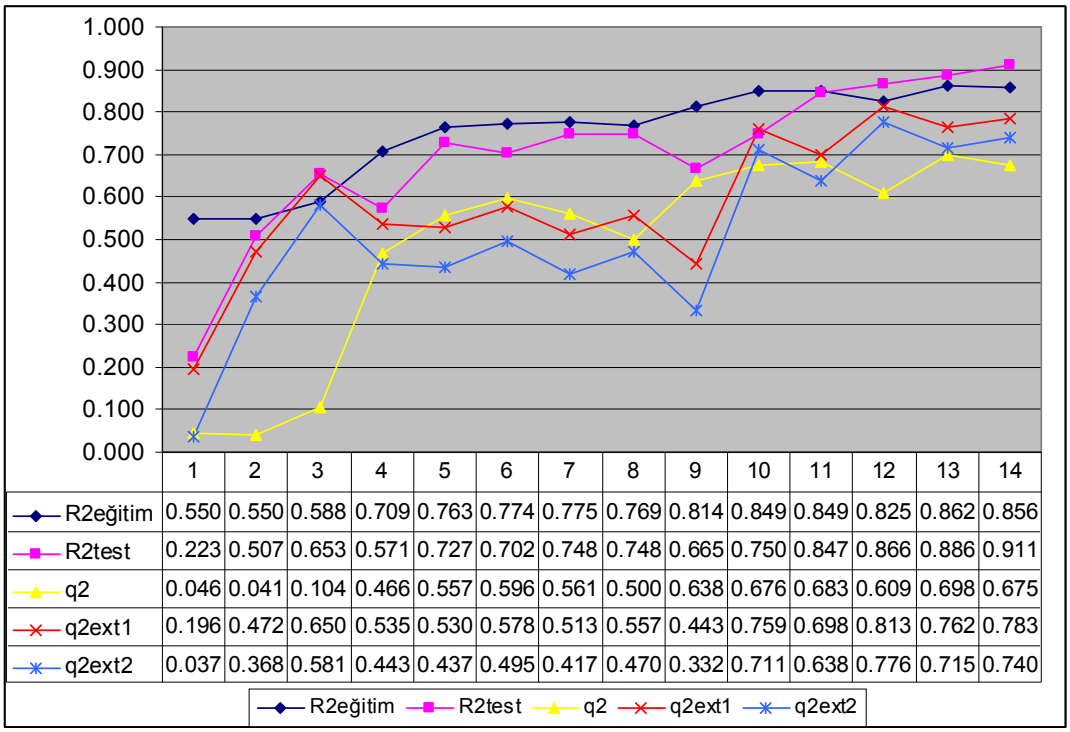
Bileşikler	**9 κ_j**	**10 κ_j**	**11 κ_j**	**12 κ_j**	**13 κ_j**	**14 κ_j**	**PAB_{den}**
66	93.674	86.024	93.489	95.402	92.249	91.252	95.600
67	72.432	83.192	85.864	83.715	85.041	85.576	96.000
68	96.741	84.271	96.727	103.557	94.434	96.000	96.000
69	104.306	110.717	100.412	110.394	104.774	104.508	96.500
70	100.008	86.632	88.820	82.480	83.649	81.511	97.000
71	103.182	101.721	110.357	123.412	112.206	112.872	97.000
72	85.452	78.953	80.593	85.941	82.401	80.793	97.200
73	97.400	97.400	97.400	97.400	97.400	97.400	97.400
74	99.547	75.246	91.237	89.764	86.339	83.854	97.400
75*	41.830	41.098	43.892	38.602	42.182	42.261	18.000
76*	33.091	36.983	38.656	30.930	35.546	34.721	18.000
77*	28.080	32.374	34.935	20.143	31.666	35.240	20.000
78*	46.094	45.778	40.637	39.780	42.165	41.781	30.000
79*	38.144	53.600	38.741	40.768	42.986	41.835	39.000
80*	45.700	48.340	42.256	42.744	42.959	43.460	50.000
81*	47.036	69.700	55.035	48.916	61.052	59.447	60.000
82*	52.971	54.754	62.861	71.258	58.506	63.257	65.000
83*	51.694	59.460	52.146	56.960	57.060	57.440	80.000
84*	71.400	82.962	75.131	68.892	80.779	77.711	89.000
85*	54.678	62.025	70.551	78.794	70.332	72.572	92.000
86*	50.632	83.391	71.330	84.969	71.533	76.360	95.000
87*	55.551	104.444	71.817	75.435	78.042	79.000	96.000
$R^2_{eğitim}$	0.814	0.849	0.849	0.825	0.862	0.856	
$S.Hata_{eğitim}$	0.051	0.046	0.046	0.049	0.044	0.045	
R^2_{test}	0.665	0.750	0.847	0.866	0.886	0.911	
$S.Hata_{test}$	0.175	0.151	0.118	0.110	0.102	0.090	
q^2	0.638	0.676	0.683	0.609	0.698	0.675	
q^2_{ext1}	0.443	0.759	0.698	0.813	0.762	0.783	
q^2_{ext2}	0.332	0.711	0.638	0.776	0.715	0.740	
Not: "*" ile simgelenen bileşikler test bileşikleridir.							

Tablo 4.13'de Penisilin serisinde 74 bileşik eğitim, 13 bileşik test seti için aktivite hesaplamalarında aktiviteye en fazla etkisi olan parametre numaraları ile buna karşılık gelen κ_j değerleri görülmektedir.

Tablo 4.13. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için aktivite hesaplamalarında elde edilen parametre numaraları ve κj değerleri.

κ_j adedi	Parametre No	Kj değeri
1	31	-0.518
2	31, 375	-0.516, 4.692
3	31, 462, 583	-0.467, -45.212, 0.903
4	31, 38, 375, 446	-0.218, -0.007, 2.260, -0.399
5	24, 31, 375, 446, 653	-0.302, -0.199, 3.302, -0.317, 0.212
6	24, 31, 38, 375, 446, 653	-0.214, -0.147, -0.004, 3.172, -0.375, 0.1930
7	1, 24, 31, 375, 378, 446, 653	-6.164, -0.275, -0.187, 3.088, -0.037, -0.364, 0.187,
8	24, 31, 123, 429, 446, 470, 583, 653	-0.279, -0.226, -0.040, 81.075, -0.256, -45.827, 0.233, 0.247,
9	2, 9, 24, 31, 32, 375, 378, 446, 653	1.271, 6.343, -0.365, -0.081, -0.021, 1.856, 4.411, -0.478, 0.251
10	2, 24, 31, 378, 446, 462, 470, 562, 563, 653	0.350, -0.323, -0.135, 9.430, -0.340, -40.703, -55.712, -0.218, 0.244, 0.145
11	1, 2, 9, 24, 31, 378, 446, 462, 562, 563, 653	21.877, 2.217, 17.031, -0.245, -0.137, 5.941, -0.434, -23.753, -0.263, 0.273, 0.117
12	1, 2, 9, 24, 31, 398, 429, 446, 470, 563, 583, 653	17.817, 2.381, 17.435, -0.400, -0.093, 0.024, 100.000, -0.386, -55.718, 0.060, 0.462, 0.332
13	1, 2, 9, 24, 31, 38, 378, 446, 462, 470, 562, 563, 653	15.982, 1.863, 13.352, -0.262, -0.123, -0.001, 7.837, -0.404, -21.076, -42.422, -0.255, 0.260, 0.181
14	1, 2, 9, 24, 31, 38, 378, 429, 446, 462, 470, 562, 563, 653	16.773, 1.978, 14.305, -0.273, -0.115, -0.001, 6.421, 44.432, -0.402, -18.674, -49.144, -0.225, 0.235, 0.207

Penisilin serisinde 74 bileşik eğitim, 13 bileşik test seti olarak aktivite hesaplamalarında hem eğitim hem de test setine ait tüm hesaplamalar için deneysel değerlere karşılık gelen hesaplanan aktivite değerleri yukarıdaki tablo 4.11 ve 4.12'de verilmiştir. Şekil 4.6'da penisilin serisi için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği verilmiştir. Aktivite hesaplamalarında kullanılan parametre sayısının az, regresyon katsayısının ve q^2 değerinin maksimum olduğu durum en uygun durum olduğundan hesaplamalarda da görüldüğü üzere belli bir parametre sayısına kadar regresyon katsayısının ve q^2'nin yükseldiği görülmektedir. Şekil 4.6'da yapılan hesaplamalara ilişkin sonuçlarda penisilin serisinde eğitim ve test seti için 14 parametre en uygun parametre sayısı olarak belirlenmiştir.



	1	2	3	4	5	6	7	8	9	10	11	12	13	14
R2eğitim	0.550	0.550	0.588	0.709	0.763	0.774	0.775	0.769	0.814	0.849	0.849	0.825	0.862	0.856
R2test	0.223	0.507	0.653	0.571	0.727	0.702	0.748	0.748	0.665	0.750	0.847	0.866	0.886	0.911
q2	0.046	0.041	0.104	0.466	0.557	0.596	0.561	0.500	0.638	0.676	0.683	0.609	0.698	0.675
q2ext1	0.196	0.472	0.650	0.535	0.530	0.578	0.513	0.557	0.443	0.759	0.698	0.813	0.762	0.783
q2ext2	0.037	0.368	0.581	0.443	0.437	0.495	0.417	0.470	0.332	0.711	0.638	0.776	0.715	0.740

Şekil 4.6. Penisilin serisi 74 bileşik eğitim, 13 bileşik test için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği.

Penislin serisi için tablo 5.6'da verilen 14 parametre içerisinde eğitim setinde aktiviteye katkısı en fazla olan parametreyi bulmak için bir parametre dışarıda bırakılarak (leave one out) hesaplamaları yapılmıştır. Yani 14 parametrenin içinde bir tanesi dışarıda bırakılarak 13 parametre için eşitlik 3.3 kullanılarak hesaplamalar yapılmış ve R^2 değerleri literatürlere benzer şekilde hesaplanmıştır [72]. Ayrıca E istatistik değerleri hesaplanarak 14 parametre içerisinde hangi parametrenin aktivite üzerinde daha etkili olduğunu bulunmuştur. E ve q^2 değelerindeki düşüş $a^{(9)}$ ile simgelenen 446 nolu parametrenin aktivite üzerindeki etkisinin fazla olduğunu, E ve q^2 değelerindeki artış $a^{(8)}$ ile simgelenen 429 nolu parametrenin aktivite üzerindeki etkisinin en az olduğunu göstermektedir. Diğer parametrelerin etkileri tablo 4.14'de gösterilmektedir.

Tablo 4.14. Penisilin serisi 74 bileşik eğitim, 13 bileşik test seti için leave one out hesaplama sonuçları.

ELENEN PARAMETRELER	İndeks	Parametreler	E	$R^2_{eğitim}$	R^2_{test}	q^2	q^2_{ext1}	q^2_{ext2}
	$a^{(1)}$	1	1.017	0.859	0.809	0.680	0.718	0.663
	$a^{(2)}$	2	0.900	0.844	0.717	0.639	0.734	0.682
	$a^{(3)}$	9	0.946	0.853	0.759	0.656	0.738	0.686
	$a^{(4)}$	24	0.738	0.830	0.709	0.559	0.595	0.515
	$a^{(5}$	31	0.723	0.831	0.821	0.550	0.728	0.675
	$a^{(6)}$	38	1.018	0.855	0.909	0.681	0.796	0.756
	$a^{(7)}$	378	0.835	0.834	0.725	0.611	0.703	0.644
	$a^{(8)}$	**429**	**1.078**	**0.862**	**0.886**	**0.698**	**0.762**	**0.715**
	$a^{(9)}$	**446**	**0.492**	**0.732**	**0.669**	**0.340**	**-0.090**	**-0.306**
	$a^{(10)}$	462	0.986	0.852	0.868	0.671	0.692	0.631
	$a^{(11)}$	470	0.958	0.852	0.844	0.661	0.693	0.632
	$a^{(12)}$	562	0.818	0.833	0.820	0.602	0.750	0.701
	$a^{(13)}$	563	0.737	0.825	0.791	0.559	0.651	0.582
	$a^{(14)}$	653	0.904	0.852	0.780	0.640	0.695	0.635

5. BÖLÜM

SONUÇ

5.1. EMRE, ECSP ve Matlab Yazılımları

Günümüzde bilgisayar-destekli ilaç tasarımı konusunda birçok çalışma yapılmasına rağmen, ülkemizde henüz bu konuda herhangi bir yerli yazılım bulunmaması nedeniyle araştırmacılar yabancı kaynaklı yazılımlara yönelmektedirler ve bu paket programları yurtdışından yüksek ücretlerle temin etmek zorunda kalmaktadırlar. Bu çalışmamızda algoritması bizim tarafımızdan ortaya koyulan yeni yazılımlar geliştirilmiş ve bu çalışma kapsamında penisilin serisi için başarıyla uygulanarak iyi sonuçlar elde edilmiştir. Ayrıca geliştirilen bu programların ticarileştirilmesi ve patent alma gibi potansiyelleri de bulunmaktadır. Böylece geliştirilen bu yeni yazılımlar hem ülke ekonomisine hem de bilime katkı sağlayacaktır. Hazırlanan yazılımlarla; deneysel aktivitesi bilinen bileşiklerin teorik aktivitelerinin öngörülmesi; bilinmeyenlerin aktivitelerinin öngörülmesi ve bilinmeyen bileşiklerin tasarlanması amaçlanmaktadır.
Çalışma grubumuz tarafından yazılımları gerçekleştirilen programlar 3 aşamadan oluşmaktadır:

EMRE programı (Delphi 7.0 programlama dili kullanılarak hazırlanmış) bileşiklerin ve konformerlerin Spartan, Gaussian ve Codessa programlarının hesaplamaları sonucu elde edilen çıktı dosyalarından bileşiklerin elektronik, geometrik ve fizikokimyasal özellikleri kullanılarak aktivite üzerinde etkisi olduğu tahmin edilen bu değişkenleri hazırlamaktadır. Ayrıca hazırlanan parametreler sadece aktivite hesaplamalarında (SAR ve QSAR çalışmaları) değil araştırmacıya bağlı olarak istatistiksel (SPSS programında, kümeleme analizlerinde, excel hesaplamalarında, regresyon ve korelasyon) hesaplamalarda kullanılabilmektedir. Bileşiklerin ve konformerlerin mulliken, elektrostatik yükleri ve atomik valensleri kullanarak matrisin köşegen olan elemanlarını;

kartezyen koordinatlar, mesafe ve bağ dereceside kullanılarak köşegen olmayan elemanlarını hesaplayarak bileşiklerin üç boyutlu matrisini (ECMC) hazırlamaktadır. C programlama dili kullanılarak yazılmış ve arayüzü java ile yapılmış olan ECSP programı konformasyon matrislerini kullanarak alt matrisleri yani farmakofor grubunu hesaplamaktadır. Her iki programa ait arayüzler hem fonksiyonel hemde kullanılırlık açısından başarılıdır.

Matlab, teknik hesaplamalar ve matematiksel problemlerin çözümü ve analizi için tasarlanmış bir yazılım geliştirme aracıdır. Özellikle mühendislik alanındaki sistemlerin analizinde kullanılan Matlab, yapay sinir ağları, optimizasyon, veri elde etme, veritabanı, gibi araçları ile hesaplamalar için mükemmel bir ortam sunar. Matlab'da yazılan programlar, Matlab'in kendine özgü dili kullanılarak yazılır ve Matlab içinde çalıştırılır. Çalışmalarımızda kullanmak üzere yazılan kodlar ile hesaplamarımız başarılı bir şekilde gerçekleştirilmiştir.

Bu çalışma farklı disiplinler ile işbirliği içerisinde olması açısından özgün bir çalışmadır. Kimya bilimi dışında denklem, üç boyutlu matris çözümleri ve yazılımlar için matematik ve bilgisayar mühendisliği gibi disiplinlerde bu çalışmanın içinde bulunmaktadır.

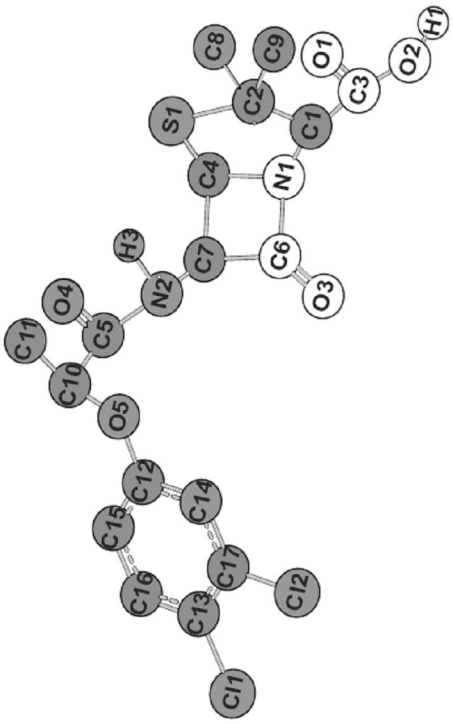
Deneysel çalışmalara temel oluşturması amacıyla aktivite gösterebilecek yeni bileşikleri çeşitli ilaç firmalarına önermek ve firmaların AR-GE çalışmalarına katkıda bulunulmasını sağlaması açısından oldukça önemli bir çalışmadır.

Konformerler arasındaki enerji farkı çok az olduğu için (oda sıcaklığında ortalama 0-2 kcal/mol) biyoalıcıyla ilacın hangi konformerin veya konformerlerin etkileştiği bilinemediği için hesaplamalarda bileşiklerin bütün konformerleri geliştirilen bu programlarda Boltzmann dağılımlarıyla orantılı olarak dikkate alınarak hesaplamalar gerçekleştirldiği için bu çalışma bir 4D-QSAR çalışmasıdır.

5.2. Penisilin Türevlerinin Teorik Aktivite Hesaplama Sonuçları

Penisilin serisinde ilk kez Pha belirlenerek ve bu seriler için EC-GA yöntemi ilk kez kullanılmıştır. Penisilin serisinde aktivitesi en yüksek olan 73. bileşiğin en düşük enerjili konformerine (referans bileşik) göre ECMC'lerin karşılaştırılması ile şekil 5.1'de gösterildiği gibi N1, C6, C3, O1, O2, H1 ve O3 atomları farmakofor grup olarak

bulunmuştur. Penisilin serisi için ECSP programı kullanılarak elde edilen farmakofor grubu ve buna ait ECSA şekil 5.1'de gösterilmiştir.



73. (Referans) Bileşiğin ECSA (PHA)

N1	C6	C3	O1	O2	H1	O3	
-0.110	0.935	2.430	2.934	3.519	4.215	2.425	**N1**
	0.274	3.767	4.372	4.641	5.404	1.953	**C6**
		0.384	1.833	1.047	1.906	4.322	**C3**
			-0.362	2.182	2.226	5.075	**O1**
				-0.308	0.915	4.909	**O2**
					0.230	5.693	**H1**
						-0.247	**O3**

O1 Atomunun Mulliken Yükü

O1-O2 Atomları Arasındaki Mesafe

C6-O3 Atomları Arasındaki Bağ Dercesi

88 Bileşiğin 1222 konformeri icin minimum tolerans değerleri

N1	C6	C3	O1	O2	H1	O3	
±0.064	±0.041	±0.093	±0.671	±0.772	±1.811	±0.019	**N1**
	±0.042	±0.405	±0.628	±1.068	±2.318	±0.046	**C6**
		±0.072	±0.051	±0.053	±0.534	±0.709	**C3**
			±0.063	±0.099	±1.192	±1.404	**O1**
				±0.074	±0.045	±1.469	**O2**
					±0.027	±2.898	**H1**
						±0.028	**O3**

m_max:±2.8982 e88j y_max:±0.0735 e77a

Şekil 5.1. Penisilin serisi için Referans bileşiğin ECSA, farmakofor grubu temsil eden atomlar, 1222 Konformer için tolerans değeri.

Yedili bir alt matristen oluşan ECSA'da maksimum tolerans değerleri mesafe ve yük için sırasıyla ±2.898, yük için ±0.074'dir. Bu tolerans değerleri içerisinde belirlenen farmakofor atomları N1, C6, C3, O1, O2, H1 ve O3 şekil 5.1'de açık renkli şekilde gösterilmiştir. Yük için en yüksek tolerans değerini 77. bileşiğinin 10. konformeri verirken, mesafe için en yüksek tolerans değerini 88. bileşiğinin 1. konformeri vermiştir. Yani mesafe için maksimum tolerans değeri ±2.898 olduğunda 88. bileşiğinin l konformerinde de bu atomlar çıkmaktadır. Benzer şekilde yük için maksimum tolerans değeri ±0.074 olduğunda 77. bileşiğinin 10. konformerinde de farmakofor atomları çıkmaktadır.

Farmakofor grubu belirlendikten sonra 88 bileşiğin 1222 konformerinin herbiri için 646 adet moleküler parametre hazırlandı. Penisilin serisi için hazırlanan 646 moleküler parametre içerisinden aktivite üzerine en çok etkisi olan parametrelerin seçimi genetik algoritma optimizasyon tekniği kullanılarak yapılmıştır.

Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti olarak aktivite hesaplama sonuçlarında tablo 4.3-4.4 ve şekil 4.4'de verilen hesaplamalara bakıldığında 12 parametre ile verilen sonuç en uygun parametre sayısı olarak belirlenmiştir. Tablo 5.1'de 60 bileşik eğitim ve 28 bileşik test seti için 12 parametreye göre ilgili bileşiklerin konformer sayıları, deneysel ve hesaplanan aktivite değerleri verilmiştir.

Tablo 5.1. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için 12 parametreye göre deneysel ve hesaplanan aktivite değerleri.

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
1	14	7.200	15.098
2*	13	12.000	8.165
3	8	15.000	9.414
4*	20	16.800	13.256
5	15	18.000	8.368
6*	13	20.000	18.992
7	15	25.000	21.898
8	22	26.000	15.746
9	9	28.000	17.141
10*	12	32.000	18.568
11	11	33.000	12.113
12	18	38.000	15.073
13*	18	42.000	28.967
14	15	47.000	43.844
15	10	53.200	47.538
16*	14	55.000	29.690
17	24	57.000	42.705
18	25	58.000	36.773
19	19	58.800	42.796
20*	16	59.000	48.246
21	11	60.000	25.751
22	17	60.000	55.126
23	14	61.700	54.203
24	8	62.000	40.885
25	14	63.000	53.901
26*	14	65.000	68.775
27	9	66.200	57.553
28	17	68.000	83.107
29*	21	69.700	78.161
30	17	74.000	65.219
31	14	74.500	89.084
32*	13	77.000	87.655
33	10	78.000	62.042
34	10	80.000	88.725
35	11	80.000	92.580

Tablo 5.1'in devamı

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
36	13	81.500	65.249
37	16	81.500	68.292
38	11	82.100	89.402
39	17	82.200	59.104
40*	11	82.500	76.054
41	25	83.000	71.972
42	32	83.000	74.586
43	24	83.500	70.232
44*	12	83.600	84.750
45	12	84.000	85.103
46	14	84.000	96.073
47	26	86.000	81.937
48	9	86.000	87.479
49	13	86.100	79.746
50*	8	87.000	64.169
51	9	88.000	100.001
52*	8	89.300	122.865
53*	11	89.700	92.361
54*	9	91.000	96.799
55*	10	91.500	84.904
56	10	92.000	84.454
57	13	92.400	91.817
58	12	92.500	88.333
59*	15	93.300	99.601
60	20	93.600	73.816
61*	24	94.000	99.662
62*	19	94.000	117.732
63*	9	94.700	109.834
64	8	94.800	97.962
65	12	95.200	85.237
66*	13	95.600	108.359
67*	11	96.000	86.386
68	12	96.000	92.213
69	14	96.500	112.130
70	13	97.000	91.946
71	21	97.000	100.150
72	9	97.200	89.229
73	15	97.400	97.400
74	8	97.400	88.783
75*	7	18.000	33.596
76	14	18.000	27.998
77	8	20.000	18.986
78	14	30.000	42.311
79*	14	39.000	46.482
80	3	50.000	39.155

Tablo 5.1'in devamı

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
81	13	60.000	50.958
82	13	65.000	47.576
83	26	80.000	60.574
84*	5	89.000	84.486
85	9	92.000	66.064
86*	15	95.000	67.608
87*	12	96.000	81.693
88	10	2.000	1.962
Not: "*" ile simgelenen bileşikler test bileşikleridir.			

Tablo 5.2'de Penisilin serisine ait optimizasyon sonucunda bulunan en iyi 12 parametre, referans bileşiğin 1. konformeri için bu parametrelerin sayısal değerleri ve kj değerleri verilmiştir. 73. bileşiğin 1. konformeri için sayısal değerler sadece örnek olması açısından tabloda yer almıştır. Hesaplamalarda bütün konformerler için parametrelerin sayısal değerleri kullanılmaktadır. Tablo 5.2'deki moleküler parametrelere bakıldığında 12 parametre içerisinde elektronik özelliklerin daha baskın olduğu görülmektedir. Tabloda verilen parametreler penisilin serisinde 88 adet bileşiğin bütün konformerleri için ortaktır. Her bir konformer için bu parametrelerin sayısal değerleri faklıdır. Bütün moleküllerin teorik aktivite hesaplamaları yapılırken ilgili bileşiğin bütün konformerleri için bu parametrelerin kendisine karşılık gelen değerleri kullanılır. p(1,1,1) ifadesindeki numaralar sırasıyla bileşik numarası, konformer numarası ve parametre numarasına karşılık gelmektedir.

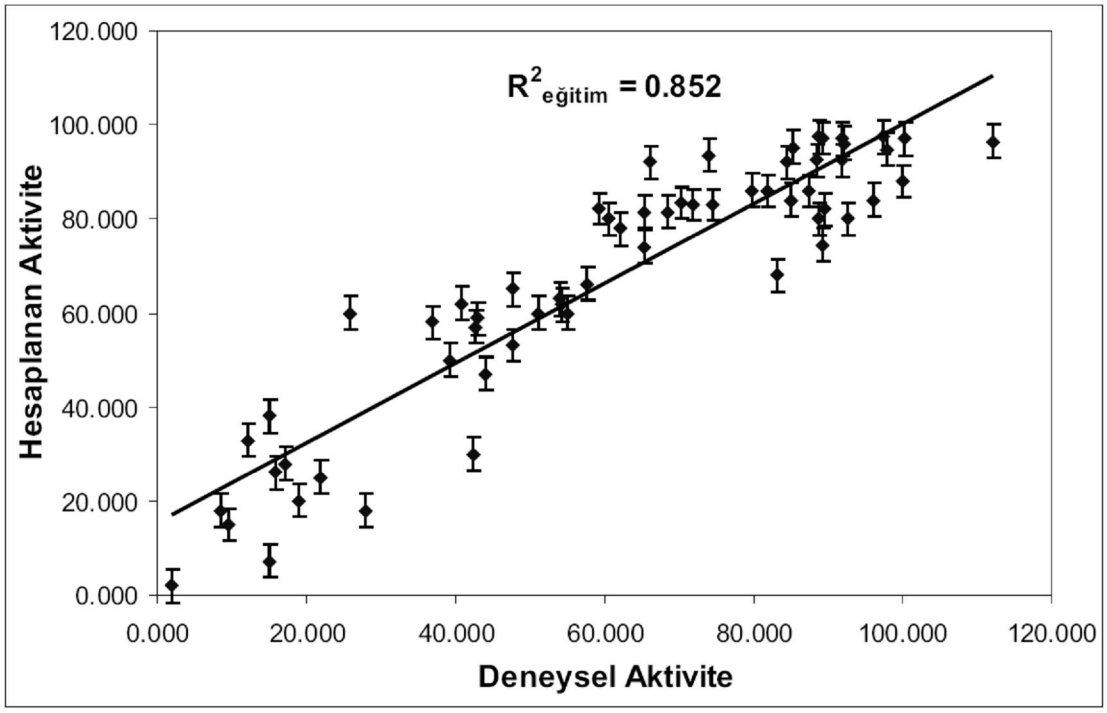
Tablo 5.2. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için en iyi 12 moleküler parametre, 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri ve Kappa (kj) değerileri.

Parametre No		Moleküler Parametreler	Psa	Kj değeri
$a^{(1)}$	**2**	Elektronegatiflik	4.898	0.520
$a^{(2)}$	**17**	Solvatasyon Enerjisi	-59.145	-0.004
$a^{(3)}$	**24**	Rotasyonel Entropi	36.886	-0.304
$a^{(4)}$	**40**	Polarlanabilirlik XZ (alpha)	-18.296	0.012
$a^{(5)}$	**43**	Polarlanabilirlik Average	192.667	-0.004
$a^{(6)}$	**81**	C6-O3-O1 Düzlemine C4 Atomunun Dik Uzaklığı	0.152	-6.048

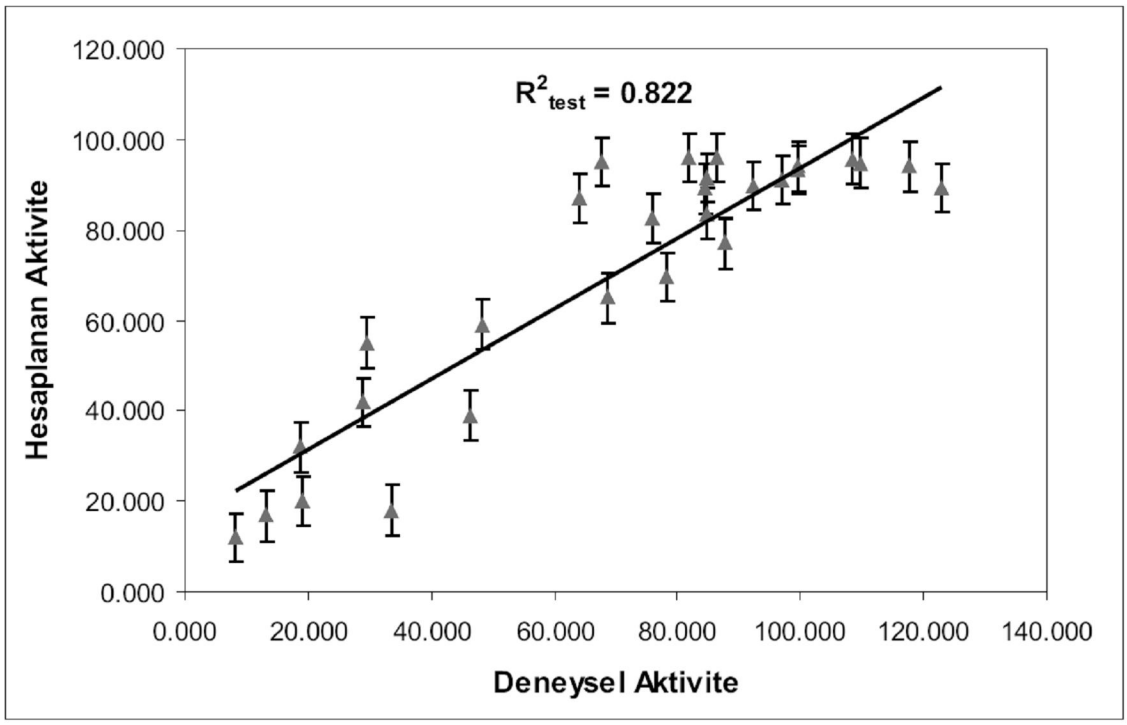
Tablo 5.2'nin devamı

Parametre No		Moleküler Parametreler	Psa	Kj değeri
$a^{(7)}$	**375**	C3-O1-O2 Düzlemine ile N1-C1 Doğrusu Arasındaki Açı	0.732	6.166
$a^{(8)}$	**442**	N2 Atomunun Nükleofilik Reaktivite İndeksi	0.001	-9.536
$a^{(9)}$	**446**	S1 Atomunun Nükleofilik Reaktivite İndeksi	1.654	-0.508
$a^{(10)}$	**451**	O3 Atomunun Nükleofilik Reaktivite İndeksi	0.000*	100.000
$a^{(11)}$	**653**	CFD (POS)	1.000	0.584
$a^{(12)}$	**704**	O3 Atomunun Mulliken Yükü	-0.2471	24.981
Psa :73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri *Gerçek Değeri 0.000072				

Penisilin serisi 60 bileşik eğitim ve 28 bileşik test seti için en iyi 12 moleküler parametre ile hesaplanan sonuçlarına göre R^2 değerleri, q^2 değerleri ve parametre sayısı göz önünde bulundurulduğunda, R^2, q^2 değeri yüksek ve parametre sayısı az olan 12 parametre için elde edilen değerlerin en iyi sonuç olduğuna karar verilmiştir. Aşağıda Şekil 5.2 ve 5.3'de 12 parametre için R^2 grafikleri sırasıyla eğitim (60 Bileşik) ve test (28 Bileşik) seti için verilmiştir.



Şekil 5.2. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin eğitim seti için çizilmiş R^2 grafiği.



Şekil 5.3. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin test seti için çizilmiş R^2 grafiği.

Eşitlik 3.2 dikkate alındığında parametrenin sayısal değeri ile kj değerinin çarpımı eğer pozitif ise bu parametre aktiviteyi artırıcı yönde negatif ise aktiviteyi azaltıcı yönde etki yapmaktadır. Tablo 5.2'e bakıldığında 2, 17 375, 451 ve 653 nolu parametreler aktiviteyi artırıcı (AG) yönde bir etkiye sahip iken 24, 40, 43, 81, 442, 446 ve 704 nolu parametreler aktiviteyi azaltıcı yönde (APS) bir etkiye sahiptir.

Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için aktivite değeri logaritma cinsinden olan hesaplama sonuçlarında tablo 4.7-4.8 ve şekil 4.5'de verilen hesaplamalara bakıldığında 12 parametre ile verilen sonuç en uygun parametre sayısı olarak belirlenmiştir. Tablo 5.3'de 60 bileşik eğitim ve 28 bileşik test seti için aktivite değeri logaritma cinsinden olan hesaplamalarda 12 parametreye göre ilgili bileşiklerin konformer sayıları, deneysel ve hesaplanan aktivite değerleri verilmiştir.

Tablo 5.3. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için aktivite değeri logaritma cinsinden olan hesaplamalarda 12 parametreye göre deneysel ve hesaplanan aktivite değerleri.

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
1	14	0.857	1.231
2*	13	1.079	1.270
3	8	1.176	1.047
4*	20	1.225	1.344

Tablo 5.3'ün devamı

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
5	15	1.255	0.953
6*	13	1.301	1.384
7	15	1.398	1.434
8	22	1.415	1.422
9	9	1.447	1.329
10*	12	1.505	1.563
11	11	1.519	1.330
12	18	1.580	1.588
13*	18	1.623	1.525
14	15	1.672	1.744
15	10	1.726	1.826
16*	14	1.740	1.844
17	24	1.756	1.760
18	25	1.763	1.711
19	19	1.769	1.669
20*	16	1.771	1.850
21	11	1.778	1.714
22	17	1.778	1.681
23	14	1.790	1.734
24	8	1.792	1.761
25	14	1.799	1.818
26*	14	1.813	1.865
27	9	1.821	1.820
28	17	1.833	1.941
29*	21	1.843	1.831
30	17	1.869	1.851
31	14	1.872	1.877
32*	13	1.887	1.857
33	10	1.892	1.871
34	10	1.903	1.947
35	11	1.903	1.898
36	13	1.911	1.863
37	16	1.911	1.920
38	11	1.914	1.951
39	17	1.915	1.920
40*	11	1.917	1.938
41	25	1.919	1.906
42	32	1.919	2.023
43	24	1.922	1.855
44*	12	1.922	2.021
45	12	1.924	1.922
46	14	1.924	1.880
47	26	1.935	1.998
48	9	1.935	1.877
49	13	1.935	1.944

Tablo 5.3'ün devamı

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
50*	8	1.940	1.744
51	9	1.945	1.994
52*	8	1.951	1.979
53*	11	1.953	2.109
54*	9	1.959	1.990
55*	10	1.961	1.971
56	10	1.964	1.993
57	13	1.966	1.898
58	12	1.966	1.926
59*	15	1.970	2.077
60	20	1.971	1.983
61*	24	1.973	2.050
62*	19	1.973	2.097
63*	9	1.976	2.031
64	8	1.977	1.906
65	12	1.979	2.042
66*	13	1.981	1.897
67*	11	1.982	2.054
68	12	1.982	2.062
69	14	1.985	2.039
70	13	1.987	1.950
71	21	1.987	1.927
72	9	1.988	1.855
73	15	1.989	1.989
74	8	1.989	1.951
75*	7	1.255	1.477
76	14	1.255	1.437
77	8	1.301	1.302
78	14	1.477	1.535
79*	14	1.591	1.697
80	3	1.699	1.682
81	13	1.778	1.696
82	13	1.813	1.638
83	26	1.903	1.683
84*	5	1.949	1.828
85	9	1.964	1.720
86*	15	1.978	1.945
87*	12	1.982	1.939
88	10	0.301	0.300
Not: "*" ile simgelenen bileşikler test bileşikleridir.			

Tablo 5.4'de Aktivite değeri logaritma cinsinden olan penisilin serisine ait optimizasyon sonucunda bulunan en iyi 12 parametre, referans bileşiğin 1. konformeri için bu

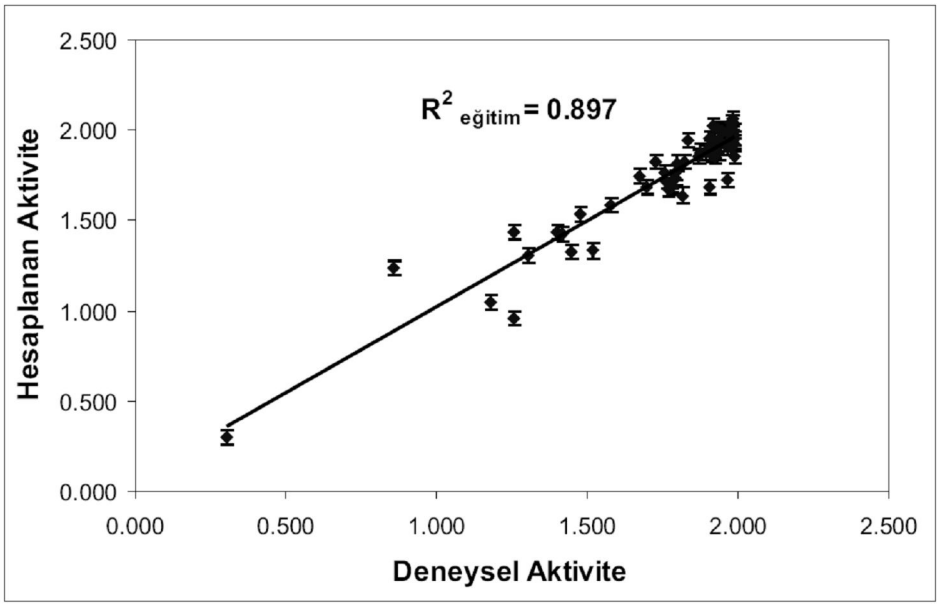
parametrelerin sayısal değerleri ve kj değerleri verilmiştir. 73. bileşiğin 1. konformeri için sayısal değerler sadece örnek olması açısından tabloda yer almıştır. Tablo 5.4'deki moleküler parametrelere bakıldığında 12 parametre içerisinde geometrik özelliklerin daha baskın olduğu görülmektedir. Özellikle bileşikler, enzimlerle ya da hedef proteinlerle etkileşirken konformasyonel yapısı değişmesi sebebiyle moleküllerin geometrik yapısının oldukça ön plana çıktığı görülmektedir.

Tablo 5.4. Penisilin serisinde 60 bileşik eğitim ve 28 bileşik test seti için en iyi 12 moleküler parametre, 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri ve Kappa (kj) değerileri.

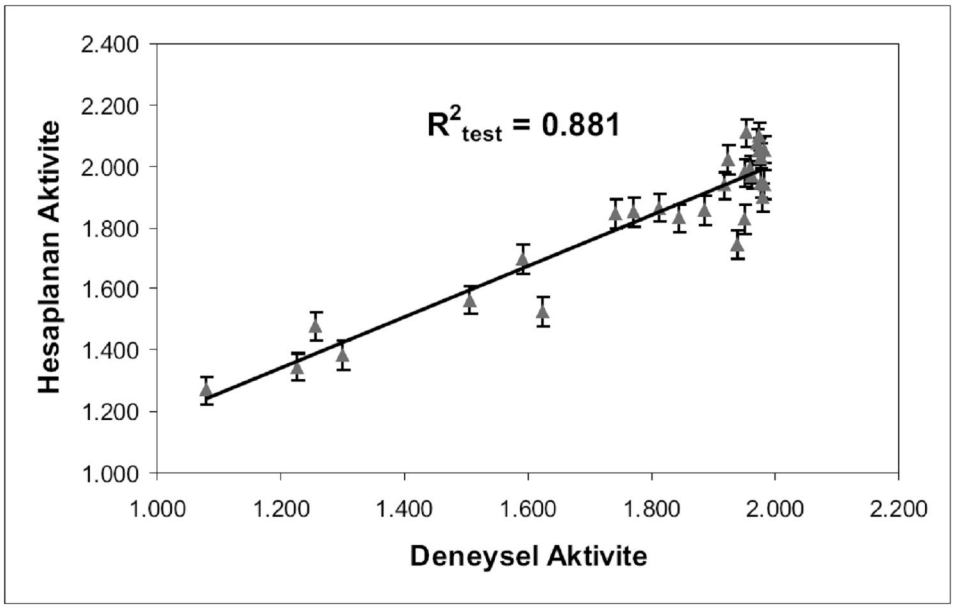
Parametre No		**Moleküler Parametreler**	**Psa**	**Kj değeri**
$a^{(1)}$	**15**	Moleküler Ağırlık	433.310	0.010
$a^{(2)}$	**23**	Ötelenme Entropisi	44.080	-1.468
$a^{(3)}$	**56**	Polarlanabilirlik XXYY (gamma)	16734.399	0.000*
$a^{(4)}$	**344**	C6-O3-O1 Düzlemine ile C5-O4 Doğrusu Arasındaki Açı	1.307	-0.299
$a^{(5)}$	**369**	C6-C3-O1 Düzlemine ile N1-C11 Doğrusu Arasındaki Açı	0.136	0.264
$a^{(6)}$	**380**	C3-O1-O2 Düzlemine ile C8-N1 Doğrusu Arasındaki Açı	0.879	1.752
$a^{(7)}$	**446**	S1 Atomunun Nükleofilik Reaktivite İndeksi	1.654	-0.076
$a^{(8)}$	**464**	C7 Atomunun Elektrofilik Reaktivite İndeksi	0.022	-2.984
$a^{(9)}$	**466**	S1 Atomunun Elektrofilik Reaktivite İndeksi	1.786	0.070
$a^{(10)}$	**634**	H1-O1-O2 Atomları Arasındaki Açı	0.435	0.546
$a^{(11)}$	**656**	PSA (Polar Yüzey Alanı)	83.561	0.004
$a^{(12)}$	**709**	C3-O1 Atomları Arasındaki Bağ Dercesi	1.833	1.189
Psa: 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri				
* Gerçek değeri : 0.000004				

Penisilin serisi 60 bileşik eğitim ve 28 bileşik test seti için aktivite değeri logaritma cinsinden olan hesaplamalarda en iyi 12 moleküler parametre ile hesaplanan sonuçlarına göre R^2 değerleri, q^2 değerleri ve parametre sayısı göz önünde bulundurulduğunda, R^2, q^2 değeri yüksek ve parametre sayısı az olan 12 parametre için elde edilen değerlerin en

iyi sonuç olduğuna karar verilmiştir. Aşağıda şekil 5.4 ve 5.5'de 12 parametre için R^2 grafikleri sırasıyla eğitim (60 Bileşik) ve test (28 Bileşik) seti için verilmiştir.



Şekil 5.4. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin eğitim seti için çizilmiş R^2 grafiği.



Şekil 5.5. Penisilin serisinde 12 parametre ile hesaplanan teorik aktivitenin test seti için çizilmiş R^2 grafiği.

Parametrenin sayısal değeri ile kj değerinin çarpımı eğer pozitif ise bu parametre aktiviteyi artırıcı yönde negatif ise aktiviteyi azaltıcı yönde etki yapmaktadır. Tablo 5.4'e bakıldığında 15, 56, 369, 380, 446, 634, 656 ve 709 nolu parametreler aktiviteyi

artırıcı (AG) yönde bir etkiye sahip iken 23, 344, 446 ve 464 nolu parametreler aktiviteyi azaltıcı yönde (APS) bir etkiye sahiptir.

Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti olarak aktivite hesaplama sonuçlarında tablo 4.11-4.12 ve şekil 4.6'da verilen hesaplamalara bakıldığında 14 parametre ile verilen sonuç en uygun parametre sayısı olarak belirlenmiştir. Tablo 5.5'de 74 bileşik eğitim ve 13 bileşik test seti için 14 parametreye göre ilgili bileşiklerin konformer sayıları, deneysel ve hesaplanan aktivite değerleri verilmiştir.

Tablo 5.5. Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti için 14 parametreye göre deneysel ve hesaplanan aktivite değerleri.

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
1	14	7.200	13.837
2	13	12.000	15.697
3	8	15.000	12.268
4	20	16.800	21.033
5	15	18.000	11.401
6	13	20.000	19.650
7	15	25.000	22.478
8	22	26.000	19.584
9	9	28.000	15.626
10	12	32.000	22.847
11	11	33.000	14.137
12	18	38.000	21.814
13	18	42.000	49.653
14	15	47.000	50.974
15	10	53.200	62.964
16	14	55.000	48.802
17	24	57.000	53.317
18	25	58.000	36.118
19	19	58.800	43.573
20	16	59.000	45.168
21	11	60.000	49.962
22	17	60.000	60.033
23	14	61.700	47.143
24	8	62.000	36.203
25	14	63.000	56.544
26	14	65.000	55.728
27	9	66.200	53.322
28	17	68.000	75.315

Tablo 5.5'in devamı

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
29	21	69.700	69.150
30	17	74.000	60.509
31	14	74.500	92.876
32	13	77.000	59.572
33	10	78.000	67.019
34	10	80.000	91.914
35	11	80.000	84.098
36	13	81.500	71.457
37	16	81.500	75.745
38	11	82.100	96.427
39	17	82.200	62.075
40	11	82.500	78.125
41	25	83.000	83.321
42	32	83.000	75.098
43	24	83.500	66.680
44	12	83.600	67.942
45	12	84.000	87.767
46	14	84.000	83.948
47	26	86.000	83.581
48	9	86.000	77.416
49	13	86.100	75.394
50	8	87.000	66.539
51	9	88.000	99.643
52	8	89.300	87.603
53	11	89.700	101.463
54	9	91.000	99.049
55	10	91.500	109.369
56	10	92.000	105.506
57	13	92.400	90.699
58	12	92.500	103.063
59	15	93.300	110.885
60	20	93.600	86.457
61	24	94.000	100.796
62	19	94.000	100.173
63	9	94.700	82.301
64	8	94.800	73.366
65	12	95.200	98.039
66	13	95.600	91.252
67	11	96.000	85.576
68	12	96.000	96.000
69	14	96.500	104.508
70	13	97.000	81.511
71	21	97.000	112.872
72	9	97.200	80.793

Tablo 5.5'in devamı

Bileşikler	Konformer Sayısı	$PAB_{deneysel}$	$PAB_{hesaplanan}$
74	8	97.400	83.854
75*	7	18.000	42.261
76*	14	18.000	34.721
77*	8	20.000	35.240
78*	14	30.000	41.781
79*	14	39.000	41.835
80*	3	50.000	43.460
81*	13	60.000	59.447
82*	13	65.000	63.257
83*	26	80.000	57.440
84*	5	89.000	77.711
85*	9	92.000	72.572
86*	15	95.000	76.360
87*	12	96.000	79.000
Not: "*" ile simgelenen bileşikler test bileşikleridir.			

Tablo 5.6'da Penisilin serisine ait optimizasyon sonucunda bulunan en iyi 14 parametre, referans bileşiğin 1. konformeri için bu parametrelerin sayısal değerleri ve kj değerleri verilmiştir. 73. bileşiğin 1. konformeri için sayısal değerler sadece örnek olması açısından tabloda yer almıştır. Tablo 5.6'daki moleküler parametrelere bakıldığında 14 parametre içerisinde elektronik özelliklerin daha baskın olduğu görülmektedir.

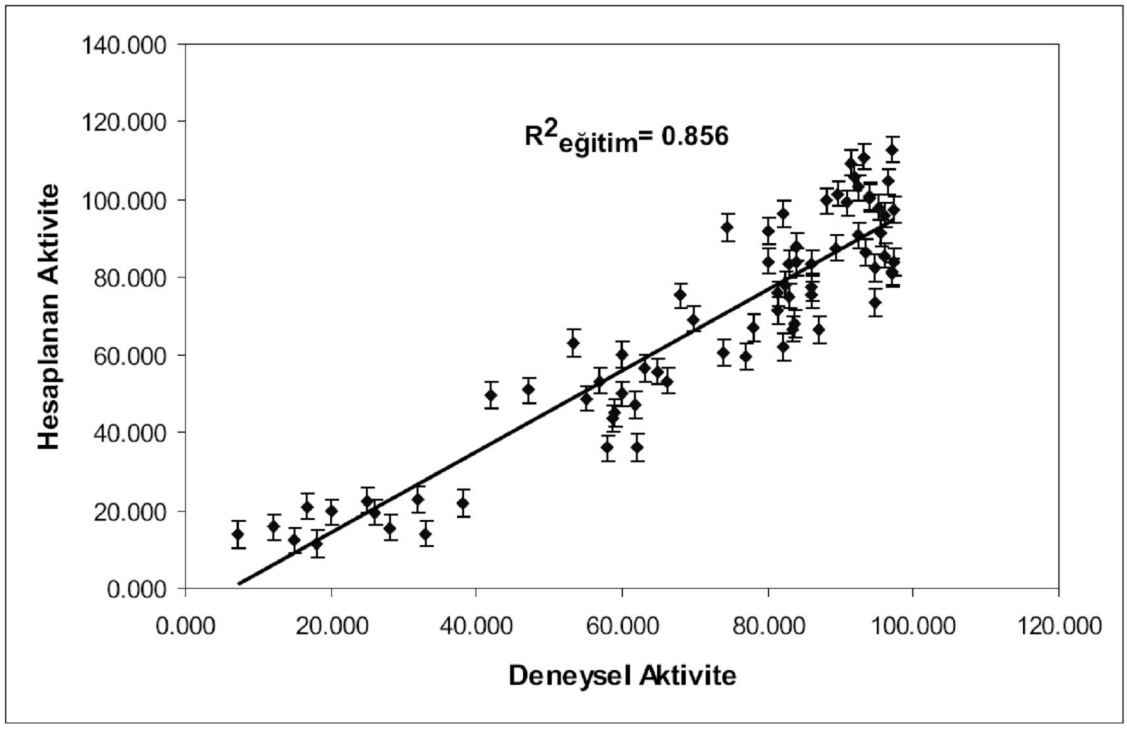
Tablo 5.6. Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti için en iyi 14 moleküler parametre, 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri ve Kappa (kj) değerileri.

Parametre No		Moleküler Parametreler	Psa	Kj değeri
$a^{(1)}$	**1**	Softness	0.230	16.773
$a^{(2)}$	**2**	Elektronegatiflik	4.898	1.978
$a^{(3)}$	**9**	1-Nükleofilik	0.370	14.305
$a^{(4)}$	**24**	Rotasyonel Entropi	6.886	-0.273
$a^{(5}$	**31**	LogP	2.090	-0.115
$a^{(6)}$	**38**	Polarlanabilirlik YY (alpha)	219.263	-0.001
$a^{(7)}$	**378**	C3-O1-O2 Düzlemine ile C2-C8 Doğrusu Arasındaki Açı	0.140	6.421
$a^{(8)}$	**429**	S1-C2 Atomları Arasındaki Bağ Derecesi	0.954	44.432
$a^{(9)}$	**446**	S1 Atomunun Nükleofilik Reaktivite İndeksi	1.654	-0.402

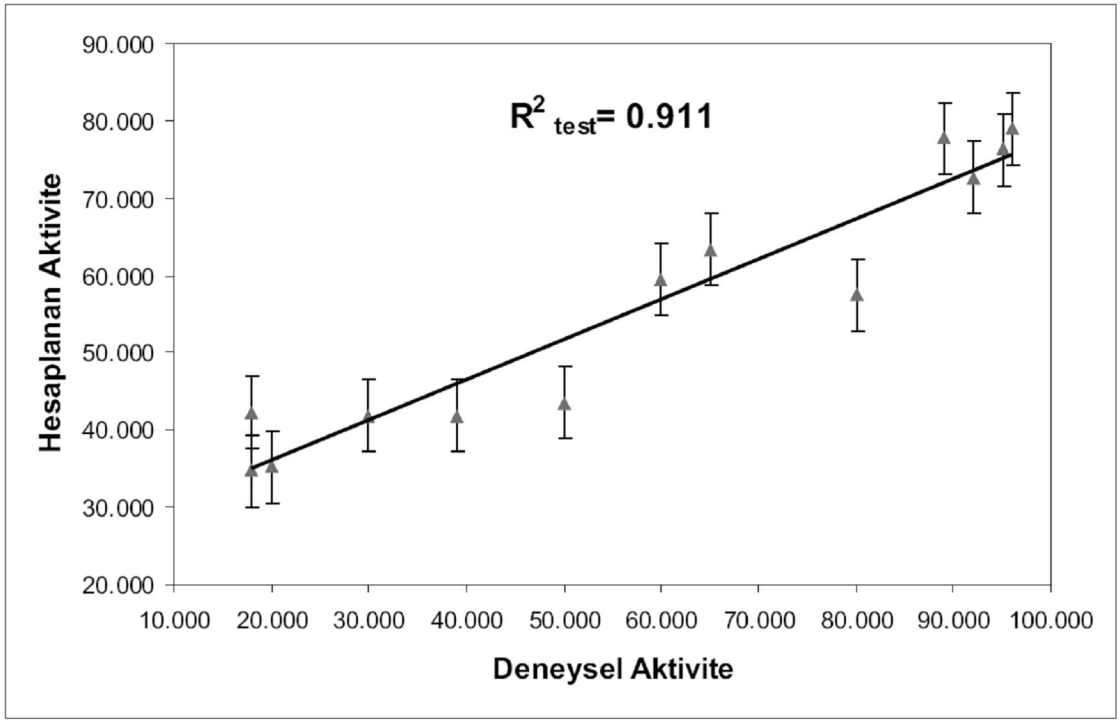
Tablo 5.6'nın devamı

Parametre No		Moleküler Parametreler	Psa	Kj değeri
$a^{(10)}$	**462**	N2 Atomunun Elektrofilik Reaktivite İndeksi	0.005	-18.674
$a^{(11)}$	**470**	C6 Atomunun Elektrofilik Reaktivite İndeksi	0.002	-49.144
$a^{(12)}$	**562**	C6-C11 Atomları Arasındaki Mesafe	8.784	-0.225
$a^{(13)}$	**563**	O3-C11 Atomları Arasındaki Mesafe	8.292	0.235
$a^{(14)}$	**653**	CFD (POS)	1.000	0.207
Psa: 73. Bileşiğin 1. Konformeri için parametrenin sayısal değeri				

Penisilin serisi 74 bileşik eğitim ve 13 bileşik test seti için en iyi 14 moleküler parametre ile hesaplanan sonuçlarına göre R^2 değerleri, q^2 değerleri ve parametre sayısı göz önünde bulundurulduğunda, R^2, q^2 değeri yüksek ve parametre sayısı az olan 14 parametre için elde edilen değerlerin en iyi sonuç olduğuna karar verilmiştir. Aşağıda şekil 5.6 ve 5.7'de 14 parametre için R^2 grafikleri sırasıyla eğitim (74 Bileşik) ve test (13 Bileşik) seti için verilmiştir.



Şekil 5.6. Penisilin serisinde 14 parametre ile hesaplanan teorik aktivitenin eğitim seti için çizilmiş R^2 grafiği.



Şekil 5.7. Penisilin serisinde 14 parametre ile hesaplanan teorik aktivitenin test seti için çizilmiş R^2 grafiği.

Parametrenin sayısal değeri ile kj değerinin çarpımı eğer pozitif ise bu parametre aktiviteyi artırıcı yönde negatif ise aktiviteyi azaltıcı yönde etki yapmaktadır. Tablo 5.6'ya bakıldığında 1, 2, 9, 378, 429, 563 ve 653 nolu parametreler aktiviteyi artırıcı (AG) yönde bir etkiye sahip iken 24, 31, 38, 446, 462, 470 ve 562 nolu parametreler aktiviteyi azaltıcı yönde (APS) bir etkiye sahiptir. Bu hesaplama sonuçları bileşik serisini aldığımız makaledeki sonuçlar ile karşılaştırıldığında sonuçlar aşağıda tabloda 5.7'de verilmiştir. Bizim çalışma grubumuz tarafından geliştirilen EC-GA metodu ile hesaplama sonuçları sırasıyla $R^2_{eğitim}$: 0.856, R^2_{test}:0.911, q^2: 0.675, q^2_{ext}: 0.783 bulunur iken MDL-QSAR ile yapılan hesaplamalarda ise $R^2_{eğitim}$: 0.800, R^2_{test}: 0.840, q^2: 0.760, q^2_{ext}: 0.840 bulunmuştur. Bizim çalışmamızdaki R^2 değerleri makaledeki sonuçlara göre daha iyi, fakat makaledeki sonuçlara göre q^2 değerlerinin biraz düşük olduğu görülmektedir. Çalışmamızda bileşiklerin konformerlerini dikkate alarak hesaplamalar yaptık. Fakat MDL-QSAR da bileşiklerin konformerlerini dikkate alarak hesap yapılamamaktadır.

Tablo 5.7. Penisilin serisinde 74 bileşik eğitim ve 13 bileşik test seti için hesaplama sonuçlarının makaledeki hesaplama sonuçları ile karşılaştırılması.

İstatistiksel Hesaplama	Bizim Hesapladığımız Değerler	Makaledeki Hesaplanan Değerler
$R^2_{eğitim}$	0.856	0.800
R^2_{test}	0.911	0.840
q^2	0.675	0.760
q^2_{ext}	0.783	0.840

Penisilin serisinde aktiviteye etki eden önemli parametreler ve bunlara karşılık gelen κj değerleri bulunmuştur. İstenilen bileşiğin bu parametre değerleri aktivite formülünde yerine koyularak çözüldüğünde aktivitesi bilinmeyen ve üzerinde deneysel çalışması yapılmayan maddelerin aktivite hesapları başarılı bir şekilde yapılabileceği q^2 değerlerinden ve test seti için hesaplanan R^2 ve standart hata değerlerinden anlaşılmaktadır. Penisilin serisi için semi emprik yöntem kullanılarak hesaplamalar yapılmıştır. Kullanılan bilgisayarların hızları yeterli ise geliştirlen bu programlar ile Hartree Fock, Density Functional Theory ve Møller Plesset hesaplamalarının sonuçlarıda aktivite hesaplamalarında kullanılabilmektedir. Çalışmalarımız sonucu geliştirilen yazılımlar yurt dışından yüksek maliyetlerle temin edilen programlara düşük maliyetli yerli bir alternatif olarak ortaya çıkmaktadır.

KAYNAKLAR

1. Hall, L.M., Hall, L.H., Kier, L.B., QSAR modeling of β-lactam binding to human serum proteins, Journal of Computer-Aided Molecular Design, 17, 103–118, 2003.
2. Roberto, G., et al., Computer-Assisted Drug Development (CADD): An Emerging Technology for Designing First-time-in-man and Proof-of-Concept Studies from Preclinical Experiments, J. Pharm. Sci., 13, 261–270, 2001.
3. Tosun, M., Farmakolojiye Giriş Ders Notları, Ege Üniversitesi, İzmir, 2007.
4. Bersuker, I.B., Dimoglo, A.S., Gorbachov, M.Yu., The Electron-Topologic Approach to The QSAR Problem Illustrated by Inhibitor Activity for Thymidine Phosphorylase and α-Chymo-Trypsin, In QSAR in Drug Design and Toxicology, Vol.10, pp. 43-48, Hadzi, D., Jerman-Blazic, B. (eds.), Elsevier, Amsterdam, 1987.
5. Hopfinger, A.J., Tokarski, J.S., Three-Dimensional Quantitative Structure-Activity Relationships. In Practical Application of Computer-Aided Drug Design, pp. 105–164, Charifson, P.S. (Ed.), Marcel Dekker, New York, USA, 1997.
6. Oprea, T.I., Waller, C.L., Theoretical and Practical Aspects of Three-Dimensional Quantitative Structure-Activity Relationships. In Reviews in Computational Chemistry, v.11, pp. 127-182, Lipkowitz, K.B., Boyd, D.B. (Eds.), Wiley-VCH, New York, 1997.
7. Mccarthy, D.J., Computational Approaches to Structure-Based Ligand Design, Pharmacology & Therapeutics, 84, 179–191, 1999.
8. Richon, A.B., An Introduction to Molecular Modeling, Mathematech, 1, 83, 1994.
9. Karacan, N., Inorganik Lab. Ders. Not, Molekül modelleme, HyperChem kullanarak MO diagramlarını oluşturma, Gazi Üniversitesi, Ankara, http://w3.gazi.edu.tr/web/nkaracan/inorglab/mm.pdf, 27 May, 2008.
10. Tosun, M., Antimikrobik İlaçlar, Farmasötik Mikrobiyoloji Ders Notları, Ege Üniversitesi, İzmir.

11. Hansch, C., et al., Correlation of biological activity of phenoxyacetic acids with hammett substituent constants and partition coefficients, Nature, 194, 178-180, 1962.

12. Topliss, J.G. (ed), Quantitative structure-activity relationships of drugs, Academic Press, New York, 1983.

13. Nakagawa, Y., et al., In Classical and Three-Dimensional QSAR in Agrochemistry, Vol. 606, Hansch, C., Fujita, T. (eds.), American Chemical Society, Washington, DC, 1995.

14. Yamagami, C., Takao, N., Fujita, T., Analysis and prediction of 1-octanol/water partition coefficients of substituted diazines with substituent and structural parameters. In QSAR and Drug Design-New Development and Applications, pp. 153–183, Fujita, T. (ed.), Elsevier, Amsterdam, 1995.

15. Devillers, J. (ed.), Comparative QSAR, Taylor and Francis, Washington, D.C., 1997.

16. Parrill, A.L., Reddy, M.R., Rational drug design novel methods and practical applications, American Chemical Society, Washington, DC., 1999.

17. Debnath, A.K., Quantitative structure-activity relationship (QSAR)—a versatile tool in drug design, Combinatorial library design and evaluation: principles, software tools, and applications, pp. 73-129, Ghose, A.K., Viswanadhan, V.N. (eds.), Mercel Dekker, New York, 2001.

18. Hansch, C., Leo, A., Exploring QSAR:Fundamentals and applications in chemistry and biology, American Chemical Society, Washington, D.C., 1995.

19. Debnath, A.K., Quantitative Structure-Activity Relationship (QSAR) Paradigm-Hansch Era to New Millennium, Mini-Rev. Med. Chem., 1, 187–195, 2001.

20. Richet, C., On the relationship between the toxicity and the physical properties of substances, Compt. Rend. Soc. Biol. 45, 775-776, 1893. Apud: Kubinyi, H., QSAR: Hansch Analysis and Related Approaches, VCH, New York, 1993.

21. Meyer, H., On the theory of alcohol narcosis I. Which property of anesthetics gives them their narcotic activity, Arch. Exp. Pathol. Pharmakol., 42, 109-118, 1899.

22. Overton, E., Studien uber die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie, G. Fischer, Jena, 1901; English translation by Lipnick, R.L.

(ed.), Studies on Narcosis, Charles Ernest Overton, Chapman and Hall, London, 1991.

23. Fühner, H., Neubauer, E., Hámolyse durch Substanzen homologen Reihen, Arch. Exp. Path. Pharm., 56, 333-345, 1907.

24. Bruice, T.C., Kharasch, N., Winzler, R.J., A correlation of thyroxine-like activity and chemical structure, Arch. Biochem. Biophys., 62, 305 -317, 1956.

25. Zahradnik, R., Chvapil, M., Study of the Relationships Between the Magnitude of Biological Activity and the Structure of Aliphatic Compounds, Experientia, 16, 511-512, 1960.

26. Zahradnik, R., Influence of the structure of aliphatic substituents on the magnitude of the biological effect of substances, Arch. Int. Pharmacodyn. Ther., 135, 311-329, 1962.

27. Zahradnik, R., Correlation of the Biological Activity of Organic Compounds by Means of the Linear Free Energy Relationships, Experientia, 18, 534-536, 1962.

28. Hammett, L.P., Physical Organic Chemistry. Reaction Rates, Equilibria and Mechanism, 2nd Ed., McGraw-Hill, New York, 1970.

29. Hansen, O.R., Hammett Series With Biological Activity, Acta Chem. Scand., 16, 1593-1600, 1962.

30. Hansch, C., Fujita, T., Rho-sigma-pi analysis. Method for correlation of biological activity and chemical structure, J. Am. Chem. Soc., 86, 1616–1626, 1964.

31. Free, Jr., S.M., Wilson, J.W., A mathematical contribution to structure activity studies, J. Med. Chem., 7, 395-399, 1964.

32. Karelson, M., Quantum chemical descriptors in QSAR. In Computational Medicinal Chemistry for Drug Discovery, pp. 641-667, Bultinck, P., De Winter, H., Langenaeker, W., Tollenaere, JP. (eds.), Dekker Inc., New York, 2004.

33. Karelson, M., Quantum chemical descriptors in QSAR. In Computational Medicinal Chemistry for Drug Discovery, pp. 1-32, Bultinck, P., De Winter, H., Langenaeker, W., Tollenaere, JP. (eds.), Dekker Inc., New York, Basel, 2004.

34. Todeschini, R., Consonni, V., Handbook of Molecular Descriptors, v.11, Mannhold, R., Kubinyi, H., Timmerman, H. (eds.), Wiley-VCH, Weinheim, 2000.

35. Stuper, A.J., Jurs, P.C., ADAPT: A Computer System for Automated Data Analysis Using Pattern Recognition Techniques, J. Chem. Inf. Comput. Sci., 16(2), 99-105, 1976.

36. Mekenyan, O., Bonchev, D., OASIS Method for Predicting Biological Activity of Chemical Compounds, Acta. Pharm., Yugosl, 36, 225-237, 1986.

37. Katritzky, A.R., Lobanov, V.S., CODESSA, Version 5.3, University of Florida, Gainesville, 1994.

38. MolConnZ, Ver. 4.05, 2003, Hall Ass. Consult., Quincy, MA.

39. Todeschini, R., Consonni, V., Mauri, A., Pavan, M., DRAGON—Software for the calculation of molecular descriptors, Ver. 5.4 for Windows, 2006, Talete srl, Milan, Italy.

40. Spartan, Version 06; Wavefunction, Inc., 2006.

41. Perkins, R., Fang, H., Tong, W., Welsh, W.J., Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology, Environmental Toxicology and Chemistry, 22(8), 1666-1679, 2003.

42. Lowis, D.R., HQSAR: A New, Highly Predictive QSAR Technique, Tripos Technical Notes, 1(5), 1-17, 1997.

43. Kubinyi, H., From Narcosis to Hyperspace: The History of QSAR, Quant. Struct.-Act. Relat., 21, 348-356, 2002.

44. Kubinyi, H., Folkers, G., Martin, Y.C., 3D-QSAR Drug Design, Vol. 3, Recent Advances, Kluwer-Escom, Boston, 1998.

45. Free, S.M., Wilson, J.W., A mathematical contribution to structure–activity studies. J. Med. Chem., 7, 395–399, 1964.

46. Cramer, R.D., III, Patterson, D.E., Bunce, J.D., Comparative molecular field analysis (CoMFA). 1. Effect of shape on binding of steroids to carrier proteins, J. Am. Chem. Soc., 110, 5959–5967, 1988.

47. Klebe, G., Comparative molecular similarity indices analysis: CoMSIA, 3D QSAR in Drug Design, Vol. 3, pp. 87–104, Kubinyi, H., Folkers, G., Martin, Y.C. (eds.), Recent. Advances, Kluwer Academic Publishers,, Dordrecht, 1998.

48. Perkins, R., et al., Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology, Environmental Toxicology and Chemistry, 22(8), 1666-1679, 2003.

49. Vedani, A., et al., Multiple-Conformation and Protonation-State Representation in 4D-QSAR: The Neurokinin-1 Receptor System, J. Med. Chem., 43, 4416-4427, 2000.

50. Lill, M.A., Multi-Dimensional QSAR in Drug Discovery, Drug Discovery Today, 12, 1013-1017, 2007.

51. Chitty, M., In Silico & molecular drug modeling glossary & taxonomy, Cambridge Healthtech Institute, Needham, USA, http://www.genomicglossaries.com/content/molecular_modeling_gloss.asp, Jan, 2010.

52. Smith, C., Molecular Modeling-Seeing the Whole Picture with Modeling Software Packages, Scientist, 12(17), 31, 1998.

53. Gund, T., Molecular Modelling of Small Molecules. In Guidebook on Molecular Modeling in Drug Design, Cohen, N.C. (ed.), pp. 55-92, Academic Press, New York, 1996.

54. Sarıpınar, E., et al., Pharmacophore Identification and Bioactivity Prediction for Triaminotriazine Derivatives by Electron Conformational-Genetic Algorithm QSAR Method, European Journal of Medicinal Chemistry Article in Press, accepted 6 June 2010, doi:10.1016/j.ejmech.2010.06.007.

55. Sahin, K., et al., Quantitative Bioactivity Prediction and Pharmacophore Identification for Benzotriazines Derivatives by Electron Conformational-Genetic Algorithm QSAR Method, SAR QSAR Environ. Res., Submitted manuscript.

56. Dimoglo, A.S., The Electron-Topological Approach in the Search and Design of Biologically Active Compounds, Thesis of Doctor of Sciences, Rostow-Don University, Russia, 1987.

57. Bersuker, I.B., et al., Origin of Musk Fragrance Activity: The Electron-Topological Approach, New J. Chem., 15, 307-320, 1991.

58. Bersuker, I.B., et al., An Electron- Conformational Method of Identification of Pharmacophore and Anti- Pharmacophore Shielding: Application to Rice Blast Activity, J. Comput-Aided Mol. Design, 13, 419–434, 1999.

59. Bersuker, I.B., et al., A Novel Electron-Conformational Approach to Molecular Modeling for QSAR by Identification of Pharmacophore and Anti-Pharmacophore Shielding, SAR and QSAR Environ. Res., 10, 157–173, 1999.

60. Bersuker, I.B., Bahçeci, S., Boggs, J.E., Improved Electron-Conformational Method of Pharmacophore Identification and Bioactivity Prediction. Application to Angiotensin Converting Enzyme Inhibitors, J. Chem. Inf. Comput. Sci., 40(6), 1363-1376, 2000.

61. Bersuker, I.B., Pharmacophore Identification and Quantitative Bioactivity Prediction Using the Electron-Conformational Method, Curr. Pharm. Design, 9, 1575–1606, 2003.

62. Bersuker, I.B., Dimoglo, A.S., The Electron-Topological Approach to the QSAR Problem, Reviews in Computational Chemistry, Vol. 2, pp. 423–460, Lipkowitz, K.B., Boyd, D.B. (eds.), VCH, New York, 1991.

63. Ehrlich, P., Ueber den jetzigen Stand der Chemotherapie, Dtsch. Chem. Ges. 42, 17–47, 1909.

64. Mason, J.S., Good, A.C., Martin, E.J., 3-D Pharmacophores in Drug Discovery, Current Pharmaceutical Design, 7, 567-597, 2001.

65. Marshall, G.R., Molecular Modeling in Drug Design. In Burger's Medicinal Chemistry and Drug Discovery, Principles and Practice, Vol. 1, pp. 573-659, 5th ed., Wolff, M.E. (ed.), John Wiley and Sons Inc., New York, 1995.

66. Ghose, A.K., Wendoloski, J.J., Pharmacophore Modelling, Methods, Experimental Verification and Applications. In 3D QSAR in Drug Design, Ligand-Protein Interactions and Molecular Similarity, Vol.2, pp. 253-271, Kubinyi, H., Folkers, G., Martin, Y.C. (eds.), Kluwer-Escom, Boston, 1998.

67. Wermuth, C.G., Langer, T., Pharmacophore Identification. In 3D QSAR in Drug Design, Theory, Methods and Applications, Vol. 1., pp. 117–136, Kubinyi, H. (ed.), Escom, Leiden, The Netherlands, 1993.

68. Glonder, V.E., Vorpagel, E.R., Computer-Assisted Pharmacophore Identification. In 3D QSAR in Drug Design, Theory, Methods and Applications, Vol. 1., pp. 137–149, Kubinyi, H. (ed.), Escom, Leiden, The Netherlands, 1993.

69. Doucet, J.P., Weber, J., Computer-Aided Molecular Design: Theory and Applications, pp. 364–404, Academic Press, San Diego, 1996.

70. Dimoglo, A.S., et al., Electronic-Topological Investigations of the Relationships Between Chemical Structure and Ambergris Odor, New Journal of Chemistry 19, 1217-1226, 1995.

71. Bersuker, I.B., Bahceci, S., Boggs, J.E., The Electron-Conformational Method of Identification of Pharmacophore and Anti-Pharmacophore Shielding, Pharmacophore Perception, Development, and Use in Drug Design, pp. 457–474, Guner, O.F. (ed.), International University Line, La Jolla, 2000.

72. Rosines, E., Bersuker, I.B., Boggs, J.E., Pharmacophere Identification and Bioactivity Prediction for Group I Metabotropic Glutamate Receptor Agonists by the Electron-Conformational QSAR Method, Quant. Struct.-Act. Relat., 20, 327-334, 2001.

73. Marenich, A.V., et al., Quantitative Antidiabetic Activity Prediction for the Class of Guanidino- and Aminoguanidinopropionic Acid Analogs Based on Electron-Conformational Studies, J. Chem. Inf. Model., 48, 556-568, 2008.

74. Parthasarathi, R., et al., Electrophilicty index as a possible descriptor of biological activity, Bioorganic & Medicinal Chemical, 12, 5533-5543, 2004.

75. Klunzinger, P., What are the units of polarizabilty, Wavefunction Support, USA, http://www.wavefun.com/support/sp_compfaq/Properties_FAQ.html, 3 Jun, 2006.

76. Franke, R., Theoretical Drug Design Methods, Elsevier, Amsterdam, 1984.

77. Fan, F., et al., Novel Dimer Based Descriptors with Solvational Computation for QSAR Study of Oxadiazoylbenzoyl-ureas as Novel Insect-growth Regulators, Journal of Computational Chemistry. 31, 586–591, 2009.

78. Chapra, S.C., Canale, R., Numerical Methods for Engineers: With Software and Programming Applications, McGraw-Hill Science, 4 edition,16 July 2001.

79. Goldberg, D.E., Genetic Algorithms in Search, Optimization and Machine Learning, pp. 1–7, Addison-Wesley Longman Publishing Co. Inc., Boston, MA, USA, 1989.

80. Yeniay, Ö., An Overview of Genetic Algorithms, Anadolu Üniversitesi Bilim ve Teknoloji Dergisi, 2(1), 37-49, 2001.

81. Engin, O., Akış Tipi Çizelgeleme Problemlerinin Genetik Algoritma ile Çözüm Performansının Arttırılmasında Parametre Optimizasyonu, Yayınlanmamış Doktora Tezi, İTÜ, İstanbul, 2001.

82. Jang, J.S.R., Derivative-Free Optimization, In:Neuro-Fuzzy and Soft Computing: A Computational Approach To Learning and Machine Intelligence, pp. 173–196, Chapter 7, Prentice-Hall, USA, 1997.

83. Yeo, M.F., Agyel, E.O., Optimising Engineering Problems Using Genetic Algorithms, Engineering Computations, 15(2), 268-280, 1996.

84. Fığlalı, A., Engin, O., Genetik Algoritmalarla Akış Tipi Çizelgelemede Üreme Yöntemi Optimizasyonu, İTÜ Dergisi, s.1-6, 2002.

85. Braysy, O., Local Search and Variable Neighborhood Search Algorithms for The Vehicle Routing Problem With Time Windows, Ph.D. Thesis, University of Vaasa, Finland, 2001.

86. Fung, R.Y.K., Tang, J., Wang, D., Extension Of A Hybrid Genetic Algorithm For Nonlinear Programming Problems With Equality And Inequality Constraints, Computers & Operations Research, 29(3), 261-274, 2001.

87. Altıparmak, F., Dengiz, B., Smith, A.E., An Evolutionary Approach For Reliability Optimization in Fixed Topology Computer Networks, Transactions On Operational Research, 12, 57-75, 2000.

88. Gonzales, E.L., Fernandez, M.A.R., Genetic Optimisation of A Fuzzy Distribution Model, International Journal of Physical Distribution and Logistics Management, 30, 681-696, 2000.

89. Gramatica, P., Pilutti, P., Papa, E., Validated QSAR Prediction of OH Tropospheric Degradation of VOCs: Splitting intoTraining-Test Sets and Consensus Modeling, J. Chem. Inf. Comput. Sci., 44, 1794-1802, 2004.

90. Schuurmann, G., et al., External Validation and Prediction Employing the Predictive Squared Correlation Coefficients Test Set Activity Mean vs Training Set Activity Mean, J. Chem. Inf. Model., 48, 2140–2145, 2008.

91. Consonni, V., Ballabio, D.,Todeschini, R., Comments on the Definition of the Q2 Parameter for QSAR Validation, J. Chem. Inf. Model. 49 (2009), pp. 1669–1678.

92. Waller, C.L., A Comparative QSAR Study Using CoMFA, HQSAR, and FRED/ SKEYS Paradigms for Estrogen Receptor Binding Affinities of Structurally Diverse Compounds, J. Chem. Inf. Comput. Sci., 44, 758-765, 2004.

ÖZGEÇMİŞ

Ersin YANMAZ 1984 yılında Kayseri' de doğdu. İlk, orta ve lise öğrenimini Kayseri'de tamamladı. 2003 yılında İnönü Üniversitesi Fen-Edebiyat Fakültesi Kimya Bölümünü kazandı. 2004 yılında Erciyes Üniversitesi Fen-Edebiyat Fakültesi Kimya Bölümüne yatay geçiş yaptırarak Erciyes Üniversitesi Kimya Bölümünden 2007 yılında mezun oldu. Aynı yıl Erciyes Üniversitesi Fen Bilimleri Enstitüsü Kimya Ana Bilim Dalında lisansüstü eğitimine başladı. Halen Erciyes Üniversitesi Fen Bilimleri Enstitüsü Kimya Ana Bilim Dalında lisansüstü eğitimine devam etmektedir.

Adres: Selçuklu Mah. Ahmet-Ayşe
İnci Cad. Alkış Sk. No:28
Melikgazi/KAYSERİ
Tel : 05547322234-05377082708
e-posta : ersinyanmaz@hotmail.com