

T.C.
ERCIYES ÜNİVERSİTESİ
BİLİMSEL ARAŞTIRMA PROJELERİ
KOORDİNASYON BİRİMİ



**BENZOTİYENOPRİMİDİN BİLEŞİK SERİSİNİN
ELEKTRON KONFORMASYON-GENETİK ALGORİTMA (EC-GA)
YÖNTEMİYLE QSAR İNCELEMESİ**

Proje No: FBY-11-3425

SONUÇ RAPORU

Proje Yürütücüsü:
Prof. Dr. Emin SARIPINAR
Fen Fakültesi/Kimya

Harun KESKİN-Sevtap ÇAĞLAR-
Semiha AYDIN-Sevinç ÇATALKAYA
Fen Fakültesi/Kimya

Kasım 2013

KAYSERİ

TEŞEKKÜR

Bu proje çalışmasına maddi destek veren Erciyes Üniversitesi Bilimsel Araştırma Projeleri Birimi'ne (Proje No: FBY-11-3425) teşekkür ederiz.

**BENZOTİYENOPİRİMİDİN BİLEŞİK SERİSİNİN ELEKTRON
KONFORMASYON-GENETİK ALGORİTMA (EC-GA) YÖNTEMİYLE QSAR
İNCELEMESİ**

ÖZET

Bu çalışmada benzotiyenopirimidin türevleri için elektron konformasyonel genetik algoritma (EC-GA) metodu kullanılarak biyoaktiviteden sorumlu farmakofor grupların belirlenmiş ve biyoaktivite tahmini yapılmıştır. Bu amaçla kapsamlı bir 4D-QSAR yazılım paketi (EMRE, ECSP ve Matlab kodları) kullanılmıştır. Öncelikle her bir serideki bileşiklerin semiempirik-PM3/HF metodu ile kuantum kimyasal hesaplamaları ve konformasyonel analizleri gerçekleştirilmiştir. Bu hesaplamalardan elde edilen veriler kullanılarak EMRE programı ile elektronik ve yapısal özelliklerle temsil edilen elektron konformasyonel uygunluk matrisleri (ECMC) her bir bileşiğin her bir konformeri için hazırlanmıştır. En aktif bileşiğin en düşük enerjili konformeri referans seçilmiş ve ECSP programı ile diğer bileşiklerin ECMC'leri referans bileşiğin ECMC'si ile belirli tolerans aralığında karşılaştırılarak aktiviteden sorumlu elektron konformasyonel alt matris (ECSA) elde edilmiştir. EMRE programı ile benzotiyenopirimidin serisi için 1594 parametre hazırlanmıştır. Her bir seride aktiviteye en fazla etkisi olan alt parametre setini seçmek ve teorik aktivite değerlerini hesaplamak için genetik algoritma ve doğrusal olmayan en küçük kareler (lsqnonlin) yöntemi kullanılmıştır.

Anahtar Kelimeler: QSAR, farmakofor, elektron konformasyonel metot, genetik algoritma, benzotiyenopirimidin.

QSAR STUDIES WITH ELECTRON CONFORMATION-GENETIC ALGORITHM (EC-GA) METHOD ON BENZOTHIENOPYRIMIDINE COMPOUNDS

ABSTRACT

We present the results of pharmacophore identification and bioactivity prediction for benzotienopyrimidin derivatives using the electron conformational genetic algorithm (EC-GA) method. 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. The three dimensional structures of compounds were geometry-optimized and then subjected to conformational analysis by means of the semiempirical PM3/HF method. Using the data obtained from quantum chemical calculations, electron-conformational matrices of congruity (ECMC) which are representation of structural and electronic features were constructed by EMRE software. The ECMC of lowest energy conformer of compound with the highest activity was chosen as a template and compared with the ECMCs of conformers with lowest energy of other compounds within given tolerances to reveal electron conformational submatrix of activity (ECSA) by ECSP software. A descriptor pool was generated for each compound series taking into account the revealed pharmacophore with EMRE software. To predict the theoretical activity and select the best subset of variables affecting bioactivities, nonlinear least square regression method and genetic algorithm were performed. Through the studies, a comprehensive 4D-QSAR software package (EMRE, ECSP and Matlab codes) developed by us is used.

Keywords: QSAR, pharmacophore, electron conformational method, genetic algorithm, benzothienopyrimidine.

İÇİNDEKİLER

TEŞEKKÜR.....	III
ÖZET.....	IV
ABSTRACT.....	V
İÇİNDEKİLER.....	VI
GİRİŞ	1

1. BÖLÜM TEMEL KAVRAMLAR

1.1. Giriş	1
1.2. İlaç	4
1.3. Bilgisayar Destekli İlaç Tasarımı.....	4
1.4. Hesaplamalı Kimya	5

2. BÖLÜM QSAR

2. 1. QSAR’a Genel Bakış.....	6
2.2. QSAR’ın Temel İlkeleri.....	6
2.3. Hansch Analizi.....	9
2.4. Free-Wilson Analizi	10
2.5. QSAR Teorisi.....	10
2.6. QSAR Modelinin Geliştirilmesi	10
2.7. 2D-QSAR Analizi	11
2.8. 3D-QSAR	12
2.9. 4D-QSAR	13
2.10. 5D QSAR.....	13
2.11. 6D QSAR.....	13

2.12. Moleküler Modelleme ve QSAR	14
---	----

3. BÖLÜM

MATERYAL VE METOT

3.1. Materyal.....	16
3.2. Metot.....	17
3.3. Elektron Topolojik Metot.....	17
3.4. Elektron Konformasyonel Metot	18
3.4.1. ECMC Matrislerinin Oluşturulması	19
3.4.2. Farmakofor Tanımlama	19
3.4.3. APS ve AG Grupları	22
3.4.4. Aktivitenin Genel Formülü	23
3.5. Genetik Algoritma.....	28
3.5.1. Genetik Algoritmaların Temel Teoremi ve Basit Genetik Algoritma.....	28
3.5.2. Çözümlerin Kodlanması.....	30
3.5.3. İlk Populasyonun Oluşturulması	30
3.5.4. Uygunluk Değerinin Hesaplanması	31
3.5.5. Çoğalma İşleminin Uygulanması	31
3.5.6. Çaprazlama İşleminin Uygulanması	31
3.5.7. Mutasyon İşleminin Uygulanması	32
3.5.8. Yeni Kuşağın Oluşması ve Döngünün Durdurulması	32
3.5.9. Genetik Algoritmalarda Parametre Seçimi	32
3.5.10. Genetik Algoritmaların Uygulama Alanları.....	33
3.6. Elektron Konformasyon-Genetik Algoritma (EC-GA) Metodu.....	33
3.6.1. EMRE Yazılım Sisteminin Geliştirilmesi.....	37
3.6.2. Adım 1.....	39
3.6.3. Adım 2.....	48
3.6.4. Adım 3.....	53
3.6.5. Adım 4.....	55
3.6.6. Adım 5.....	60
3.6.6.1 Moleküler Parametreler.....	62

3.6.7. Adım 6.....	66
3.6.7.1. Optimizasyon.....	67
3.7. İstatistiksel Analiz.....	72

4. BÖLÜM

BULGULAR

4.1. ECMC Matrislerinin Hazırlanması ve Farmakofor Bulunması.....	77
4.2. Aktivite Hesaplamasında Kullanılan Parametreler.....	84
4.3. Benzotiyenopirimidin Bileşik Serisinin Aktivite Hesaplamaları.....	97
4.3.1. Benzotiyenopirimidin Bileşik Serisinin Eğitim ve Test Seti Olarak Hesaplanması.....	97

Hata! Yer işareti tanımlanmamış.

5. BÖLÜM

TARTIŞMA VE SONUÇ.....	109
KAYNAKLAR	Hata! Yer işareti tanımlanmamış.

1. BÖLÜM

TEMEL KAVRAMLAR

1.1. Giriş

İlaç keşfi ve geliştirilmesinde yeni ilaçların tasarlanması deneme-yanılma ile ilerleyen disiplinler arası, zahmetli, zaman alıcı ve son derece pahalı bir süreçtir [1]. Bu nedenle ilaç adaylarının aktiviteyle ilgili özelliklerinin önceden tahmin edilmesi bu alanda önemli bir problem olup, yapısal bilgilerden yararlanarak biyolojik aktivitenin tahmin edilmesi için daha etkili ve daha pratik metotların kullanılması kaçınılmazdır. Nicel yapı-aktivite ilişkisi (Quantitative Structure-Activity Relationship, QSAR) çalışmaları biyoaktif moleküllerin davranışlarını anlamak ve olası yeni ilaçların tanımlanmasına imkân sağlamak amacıyla ilaç keşfi alanında yaygın olarak uygulanmaktadır.

Molekülün yapısı ve biyolojik aktivitesi arasında bir ilişki olduğu temeline dayanan QSAR analizi, bu ilişkiyi çeşitli yapısal özelliklerin bir fonksiyonu olarak matematiksel bir eşitlikle ifade etmeye çalışarak genellikle aynı sınıfına ait olan bileşikler için biyolojik aktivite tahmini yapar. Aktivitenin kimyasal yapıdaki değişikliklere nasıl bağlı olduğunu açıklar [2]. QSAR çalışmaları temel olarak deneysel veri toplama, farmakor tanımlama, moleküler parametre oluşturulması ve seçilmesi, model kurulması ve modelin değerlendirilmesi basamaklarını içermektedir.

Kimyasal etkileşimler üç boyutta gerçekleşen olaylar olduğu için QSAR çalışmaları çoğu zaman ilgili bileşikler için oluşturulan üç boyutlu moleküler modellere dayanır.

Çoğu biyolojik özellikler, molekülün üç boyutlu konformerlerine dayanmasına rağmen 3D-QSAR metotlarında her ligand molekülü tek bir biyoaktif konformasyon ve yönelme ile temsil edilir. Diğer konformerler 3D-QSAR tekniklerinde analiz edilmez ve modellenmez. Genellikle en düşük enerjili konformer model oluşturulmasında kullanılır. Bileşiklerin yapıları ve biyolojik aktiviteleri arasındaki ilişkinin daha iyi anlaşılması için çok boyutlu QSAR teknikleri kullanılması daha uygundur. Hopfinger ve arkadaşları tarafından geliştirilen [3] ve 3D-QSAR'ın bir uzantısı olarak düşünülen 4D-QSAR analizinde her bir ligand molekülü tek bir konformer yerine konformerler grubu ile temsil edilir. Böylece biyoaktif konformasyonun belirlenmesindeki önyargının da önüne geçilmiş olur.

İlaç tasarımında önemli bir yere sahip olan farmakofor kavramı bir bileşiğin yapısındaki bazı kısımların belirli biyoalcılarla güçlü bir şekilde etkileşirken, diğer kısımlarının etkileşimdeki rolünün daha az olduğu düşüncesine dayanır. Buna göre farmakofor, belli bir biyolojik aktivite gösteren ligand grubunun tüm aktif konformasyonlarında ortak olan biyolojik özelliklerin üç boyutlu bir haritası olarak tanımlanabilir [4]. Biyolojik aktivite için gerekli olan farmakofor ve buna ait özelliklerin kullanımı yeni ve daha iyi ilaçların geliştirilmesine yardımcı olabilir. Farmakofor tanımlama ve uygulamaları üzerine birçok metot geliştirilmiştir [5]. Elektron topolojik (ET) metot [6,7,8] ve elektron konformasyonel (EC) metot [9-12] farmakofor tanımlaması için kullanılan QSAR metotlarından iki tanesidir. EC metot ile sadece farmakofor tanımlama değil aynı zamanda doğrusal olmayan matematiksel modellere dayanarak biyoaktivite tahmini de mümkündür.

QSAR'da karşılaşılan temel problemlerden birisi de moleküler özellikleri karakterize edebilen uygun parametreleri kullanarak moleküler bilginin elde edilmesi ve matematiksel ilişki kurmak için bu parametrelere uygulanacak olan istatistiksel metodun seçilmesidir. Son yıllarda fiziksel ve kimyasal özellikleri belirten çok sayıda farklı moleküler parametreler tanımlanmıştır [13,14]. Moleküllerin fizikokimyasal ve yapısal özelliklerinin temsil edilmesinde kullanılan parametreler QSAR analizlerinin bir parçası olarak kuantum kimyasal hesaplamalardan elde edilirler ve molekülün aktivitesi bir şekilde bu parametrelerin değerleri ile ilişkilidir. Fakat bir bileşik için oluşturulan

parametrelerin hepsi aynı derecede öneme sahip değildir. Bu nedenle çok sayıda parametre içerisinde daha önemli olanları seçmek ve tahmin gücü yüksek modeller oluşturmak için birçok parametre seçim metodu mevcuttur. Temel bileşen analizi (principal-component analysis, PCA) [15], kısmi en küçük kareler (partial least-squares, PLS) [16], sinir ağları (networks, NNs) [17] ve genetik algoritma (genetic algorithm, GA) [18,19] QSAR çalışmalarında parametre seçimi için sıklıkla kullanılan yöntemlerden bazılarıdır.

Bu çalışmada benzotiyenopirimidin bileşik serisi için farmakofor tanımlama ve biyoaktivite tahmini için yeni bir 4D-QSAR yaklaşımı olarak geliştirilen elektron konformasyonel-genetik algoritma metodu kullanılmıştır. Bu bileşik serisinde Pim-1 ve Pim-2 inhibitörleri için farmakofor grup belirlenmiş ve aktivite hesapları yapılmıştır. Pim kinazlar serin/trionin ailesinin üyesidir. Pim-1, Pim-2, Pim3 ilk defa retroviral insersiyonal mutasyon çalışmalarında tanımlanmıştır ve genetik bölge olarak adlandırılmıştır [20].

Pim-1, Pim-2, Pim3 Serine/threonine kinaz reseptörleri hücre çoğalması, programlı hücre ölümü (apoptosis) hücre farklılaşması ve embriyonik gelişim düzenlenmesinde rol oynar [21].

Genellikle moloney kemirgen lösemi virüsü ile aktivite odağı haline gelen Pim-1 onkogeni ilk olarak fare T hücreesindeki lenfoma ile ilişkilendirilmiştir. Daha sonra, prostat kanseri, akut miyeloid lösemi ve diğer hematopoetik (kan üreten organ) tümörler dahil olmak üzere insanlarda görülen birçok kanser çeşidinde rol oynadığı ortaya çıkarılmıştır.

Pim-1 esas olarak hücre döngüsünün ilerlemesi, programlı hücre bölünmesi ve transkripsiyonel aktivasyon ve sinyal transdüksiyon yolunu ilgilendirmektedir [22]. Pim kinazların küçük moleküllü inhibitörlerinin kanser terapisinde terapatik öneme sahip olabileceği düşünülmektedir [20].

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. 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 yollarda rol oynayan enzimler (dihidrofolat redüktaz, asetilkolinesteraz gibi), transport işlevi olan proteinler ($\text{Na}^+\text{K}^+-\text{ATPaz}$) 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 [23].

1.3. Bilgisayar Destekli İlaç Tasarımı

Bilgisayarlar, ilaç tasarımı yöntemlerinin önemli bir kısmını oluşturmaktadır ve yapı analizleri, yapı karşılaştırmaları, öncü bileşik tasarımı, aktif konformasyon ve farmakofor (Pharmacophore, Pha) belirlenmesi, protein ve bağlanma alanı 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 bileşiklerin fizikokimyasal özelliklerine göre aktiviteyi 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. Pratik bakımdan oldukça uzun ve zor olan böyle bir problemin çözümü; deneysel olarak belirli bir aktivite gösterdiği bilinen bileşikleri birbiri ile karşılaştırarak aktivitelerine neden olan ortak yapıdaki sorumlu fragmenti bulmakla olur. Böylece molekülün yapı-aktivite ilişkisinde, aktiviteden sorumlu olan temel üç boyutlu elektronik özelliği belirlenmiş olur [24-26].

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), yarıdenel (semiempirik) yöntem ve ab initio yöntemidir. [27].

2. BÖLÜM

QSAR

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 [28] tarafından ortaya konuldu.

Bu kavram zaman içerisinde gelişerek 2D-QSAR’dan 3D-QSAR’a ve son zamanlarda da 4D-QSAR [29] 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:

$$\Phi = f(C) \quad (2.1)$$

Richet [30], bileşiklerin toksisitesinin sudaki çözünürlüklerinin tersi bir yol takip ettiğini keşfetti. Böyle bir ilişki Eşitlik 2.1'e karşılık gelmektedir. Burada $\Delta\Phi$, biyolojik aktiviteler arasındaki farktır. Bu fark kimyasal ve özellikle de fizikokimyasal özelliklerdeki değişikliklerden ΔC kaynaklanmaktadır. ΔC , kimyasal yapıdaki değişiklikleri ifade etmektedir.

$$\Delta\Phi = f(\Delta C) \quad (2.2)$$

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

Meyer [31] ve Overton [32] birbirinden bağımsız olarak yaptıkları çalışmalarında yağ-su dağılma katsayısı olarak ifade edilen lipofilite ile narkotik aktiviteler arasında doğrusal ilişki olduğunu gözlemlediler. Fühner ve Neubauer [33], 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 [34], ilk Free-Wilson tipi analizi olarak ele alınabilecek çalışmalarında tiroid hormonunun bir serisi için grup katkılarının biyolojik aktivite değerlerine etkisini formülize ettiler.

Zahradnik [35], organik bileşiklerin reaktivitesini nicel olarak tanımlamak için kullanılan Hammett eşitliği (eşitlik 2.3) [36] kavramını biyolojik aktivite verilerine (eşitlik 2.4) de uygulamaya çalıştı.

$$\log k_{R-X} - \log k_{R-H} = Q\sigma \quad (2.3)$$

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

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”[37] Modern QSAR teknolojisi 1964’de Hansch – Fujita [38] ve Free-Wilson’nın [39] birbirinden bağımsız olarak ortaya çıkardıkları yayınları ile başlamıştır. Bu iki çalışma iki yeni QSAR metodunun gelişimine yol açtı. Bu çalışmalar daha sonra sırasıyla Hansch analizi 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 Corwin 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.

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 [40] 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. ED₅₀ ya da ID₅₀) 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. pED₅₀ ya da pIC₅₀), $\log K$ (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:

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

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. Hansch Analizi

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. 1962'de Hansch

ve arkadaşları bitki gelişimini düzenleyicilerinin yapı aktivite ilişkisi ve bunların Hammett sabiti ve hidrofobisiteye bağımlılığına dair çalışmalarını yayınladılar [28].

2.4. Free-Wilson Analizi

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 [41].

2.5. QSAR Teorisi

Bir bileşiğin farklı yapısal ve kimyasal özelliklerinin, kendi biyolojik aktivitesine doğrusal toplam katkısının göz önüne alınmasına dayanır. Burada biyolojik aktivitenin belirli fizikokimyasal özelliklere dayanarak taşınma ve bağlanma ile lineer ilişki içerisinde olduğu ispatlanmıştır.

Bu durum, bazı araştırmalar, Örneğin; de novo ilaç dizaynı programı LUDI (2.13), tarafından ispat edilmiştir. Buna ek olarak, Free Wilson ve Hansch analizleri birleştirilerek yapılan çalışmaların sonucunda bu görüş desteklenmiştir [42].

2.6. 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.
- 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,

sübstitünlere ya da tüm moleküle ait olabilir. Bunlar tipik yapısal parametrelerdir. İki boyutlu yapısal bilgilerden geliştirilebilir veya molekülün üç boyutlu konformasyonlarından hesaplanabilir.

ler ise yük ve enerji değerlerini, bağlanma özellikleri de biyolojik makromolekülleri içerir.

iii) Yapı-aktivite ilişkisinin uygun korelasyonunu 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 [43,44].

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

2.7. 2D-QSAR Analizi

Aktivite ve toksisite de fizikokimyasal özelliklerin etkisinin araştırılması 19. yy a dayanır. 1935 de Hammett sabiti σ , organik asit ve bazların reaktifliği ile denge sabitleri arasındaki bağıntıyı açıklamak için kullanılmıştır. 1969'da Hansch tarafından yapılan yarı deneysel çalışmada ise QSAR için yeni bir dönem başlamıştır. QSAR da çeşitli biyolojik aktiviteyi tahmin etmek için oktanol-su dağılma katsayısı (logP) hidrofobik özellik olarak tanımlanmıştır. logP veya hidrofobik özelliğin diğer ölçümleri toksikolojide ve ilaç araştırmalarında halen geniş çapta kullanılmaktadır. Aktivite ile fizikokimyasal özellikler arasındaki ilişkiyi kuran Hansch tipi yaklaşım çok değişkenli regresyon metodunu kullanarak enzim inhibisyonu, metabolizma, ligand reseptör bağlanması, toksiklik gibi problemlere çözüm getirmek için uygulanır. QSAR daki temel çalışmalar Hansch tarafından yapılmıştır.

2D-QSAR analizi kantitatif ilişkileri elde etmek için lineer regresyon tekniklerini kullanır. 2D-QSAR tekniklerinin çoğunun istatistiksel verileri şüphelidir. Ligand ve reseptör arasındaki başlangıç etkileşimi üzerinde durur. 2D-QSAR, 3D-QSAR kadar açık değildir. Genellikle yapı temelli farmakofor tanımlaması yapar ve reaksiyon bilgileri, yapısal özelliklerin tahmini ile kimyasal bilgiler üzerinde durur [43].

2.8. 3D-QSAR

İlk olarak Cramer 1988 de üç boyutlu moleküler parametrelerle tanışmış ve daha sonra 3D-QSAR metodunu geliştirmiştir. Uzaydaki moleküllerin üç boyutlu yapısını hesaba katan QSAR metotları 3D-QSAR olarak adlandırılır. Karşılaştırmalı moleküler alan analizi (Comparative molecular field analysis, CoMFA), en önemli 3D-QSAR metotlarından biridir. Biyoaktif datalarda 3D-QSAR modellerini incelemek için geliştirilen bu metot genellikle reseptörün üç boyutlu yapısı bilinmediği zaman kullanılır. Bu metotta ligandın 3 boyutlu yapısının ve deneysel aktivitelerinin bilinmesi gerekir. CoMFA metodu molekülün geometrik, sterik ve elektrostatik özellikleri ile biyolojik aktiviteleri arasında ilişki kurmak için geliştirilmiş bir QSAR metodudur. CoMFA genellikle en düşük enerjili konformerini aktif konformasyon olarak düşünür. Temel bileşen analizi (Principal Component Analysis, PCA) ve kısmi en küçük kareler (partial least squares, PLS) genellikle CoMFA da model geliştirmek için kullanılan metotlardır [45]. 3D-QSAR da diğer bir önemli metot, karşılaştırmalı moleküler benzerlik indisleri analizi (comparative molecular similarity indices analysis, CoMSIA) dır. Bu metot benzerlik indisleri (similarity indices) ve CoMFA metotlarının birleşimidir. CoMSIA'da hem sterik hem de elektrostatik özellikler, hidrojen bağ donörü, hidrojen bağ akseptörü ve hidrofobik alanlar dikkate alınır. Karşılaştırmalı Moleküler Şekil Analizi olarak da adlandırılan ve Gerard Klebe tarafından geliştirilen CoMSIA, CoMFA ile benzerlik gösterir. Fakat CoMFA da kulon elektrostatik etkileşim potansiyeli kullanılırken CoMSIA, hidrojen bağı, hidrofobisite, elektrostatik ve sterik özellikleri tayin etmek için daha çok Gaussian fonksiyonunu kullanır [46].

Ayrıca 3D-QSAR da kullanılan diğer önemli metotlarda tarihsel sıralamasına göre DYLOMMS (dynamic lattice oriented molecular modeling system), Molecular shape analysis (MSA), GRID – determining energetically favorable binding sites on Macromolecules, DG (distance geometry), CoMFA(comparative molecular field analysis), CoMSIA(comparative molecular similarity index analysis), Receptor surface model, CoMMA(comparative molecular moment analysis) şeklinde sıralanmıştır. 3D QSAR çalışmalarında Adapt, Admet Predictor, Codessa, Dragon, Grid, Molconn-Z, Sybyl, Discovery Studio gibi programlar ticari olarak kullanılmaktadır.

2.9. 4D-QSAR

Hopfinger 1997 yılında 3D-QSAR modeline dördüncü boyutu kazandırmıştır ve 4D-QSAR analizi terimini ortaya çıkarmıştır. Bu metodun 3D-QSAR dan başlıca farkı, konformerleride hesaba katarak biyoaktivite hesaplamaları yapmasıdır. Bu metoda göre çoğu 3D-QSAR metodlarından farklı olarak, aktif konformasyon minimum enerjili konformer değildir. Belirli değerler içinde kalan minimum enerjili konformer çalışılan bileşik serisindeki her bir bileşiğin en aktif konformeri olarak tanımlanır ve aktivite hesaplamalarında bu konformerlerin tamamı hesaba katılır. 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 [47].

2.10. 5D QSAR

Vedani tarafından geliştirilen 5D QSAR yöntemi ile QSAR tekniklerine daha güçlü ve yeni bir boyut kazandırılmıştır. 5D QSAR, birçok ligandın proteine bağlanması süresince meydana gelen değişikliklerin de incelendiği bir yöntemdir. Ligantların proteinlere bağlanması ile oluşan konformasyonel değişikliklerin belirlenmesi QSAR teknikleri için en önemli ilerlemelerden biri olmuştur. Aslında konformasyonel değişikliklerin mekanizması ve derecesinin belirlenmesi proteinin yapısı bilinmediği durumlarda mümkün değildir. Ancak 5D QSAR tekniğinde farklı konformasyonel değişiklikler alternatif bağlanma modelleri sunularak açıklanabilir [48].

2.11. 6D QSAR

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 [47,48]. 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

2.12. Moleküler Modelleme ve QSAR

Moleküler modelleme, bilgisayarlı kimya ve grafiksel tasarım teknikleri kullanılarak moleküler yapı ve özellikleri araştırmak için kullanılan bir tekniktir [49]. 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 [50].

Biyoalıcı ile ilaç molekülün etkileşim mekanizmasında 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.

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. Molekülün hangi

tür özelliklerinin biyolojik aktivite formülünde kullanılacağı belirli olmadığından incelenecek girdilerin kesinliği azdır.

BÖLÜM 3

MATERYAL VE METOT

3.1. Materyal

Bu çalışmada:

- 1) Spartan 02 ve Spartan 08 paket programlarında kuantum kimyasal hesaplamalar yapılmıştır. Bu programda; semiempirik hesaplamalar 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 V2.0 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 6.0 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, 06 ve 08 programı kullanılmıştır. Kuantum kimyasal hesaplamalar içerisinde semi-empirik, HF, DFT, moller plessset (MP2) olmak üzere çeşitli hesaplama yöntemleri bulunmaktadır. Bu metotlar arasında en hızlı olan semiempirik 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. Ayrıca bu çalışmada HF hesaplama yöntemi de kullanılmıştır. Teorik aktivite hesaplamaları için de Elektron Konformasyon-Genetik Algoritma (EC-GA) metodu kullanılmıştır.

3.3. Elektron Topolojik Metot

A.S. Dimoglo ve çalışma arkadaşları tarafından geliştirilen elektron topolojik metodun [51] 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 (TDETM) 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 değer 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 de olsa aktivite göstermektedirler ve aktivitelerindeki bu düşüş moleküle bağlı olan süstituentlerin farklı olmasından kaynaklanmaktadır. 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 konformerini 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 konformerini 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 hemde aktivite hesaplamaları yapılırken bütün konformerler dikkate alınmıştır.

3.4. Elektron Konformasyonel Metot

EC metotta [12,52-54] 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. Ayrıca Bersuker farmakoforlara bağlı olarak aktivitenin hesaplanmasında kullanılmak üzere eşitlik 3.5’te 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 [53].

EC metodunda kullanılan ECMC nin diyagonal elemanları, ilgili atomun etkileşim indeksini (interaction index) gösterir, nondiyagonal elemanlar ise birbirine bağlı atomlar ise bağ derecesi, bağlı değilse bağ mesafesini gösterir. ETM metodunda atom yükleri yaygın olarak kullanılmıştır [55]. Bu çalışmada ise mulliken yükleri kullanılmıştır.

3.4.2. Farmakofor Tanımlama

Pha çoğunlukla, kesin bir biyolojik aktivite sağlamak zorunda olan bileşiğin, kimyasal, yapısal ve fiziksel özelliklerinin kombinasyonlarını kapsayan bir grup özelliği olarak ifade edilir [54]. 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 [56].

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 kimyacılar yardım eder. Farmakofor kavramı Bilgisayar Destekli İlaç Tasarımında reseptör veya enzim ile etkileşen çeşitli moleküllerdeki genel üç boyutlu özelliklerin araştırılmasında kullanılır [57].

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

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.

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.

Pha aktivitenin tanımlanmasında gereklidir fakat tek başına yeterli değildir. Pha'nın olmasına rağmen aktivite Pha dışındaki gruplar tarafından azaltılabilir (APS, anti pharmacophore shielding) ya da artırılabilir (AG, auxiliary group). Bu gruplar, Pha'nın biyoaktif ile uygun etkileşimini engelleyen APS ve moleküldeki hidrofobisite gibi özellikleri sağlayan AG gruplarıdır [56].

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 [56].

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 yapı 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 aktivitelere 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 konformasyonel 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 konformasyonel 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 yada 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 karşılaştırılır.

ECMC'nin ETMC ye göre diğer üstünlüğü de AG [56] ve APS grupları [6,54] tanımlayarak hesaba katmasıdır. Bu grupların aktiviteye katkıları, Pha ile birlikte hesaplanmasıdır. Aslında Pha'ı 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 hidrojen 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 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 [56].

3.4.4. Aktivitenin Genel Formülü

APS grupları aktiviteyi kısmen veya tamamen azaltan ve sterik engelle 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 yada azaltan S fonksiyonunu kullanırız. S fonksiyonu AG ve APS'yi temsil eden bir grup parametreye bağlıdır.

$$A \sim e^{-S} \quad (3.1)$$

Molekül-biyoalıcı etkileşiminde AG ve APS grupları $\exp(-E'n/kT)$ nin bir faktörü olarak E' miktarı kadar aktiviteyi artırır yada 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.

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

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 k_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. k_j değerleri ilgili parametrenin birimine bağlı olarak farklı birimlerde değerler alır ve o parametre ile k_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 [56].

Konformasyonel analiz, bilgisayar destekli ilaç tasarımıyla 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:

$$A_n = A_0 \frac{\sum_{i=1}^{m_e} \delta_{ni} [Pha] e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_e} e^{-E_{ni}/kT}} \quad (3.3)$$

A_0 aşağıda açıklanacağı üzere sabittir ve m_n n. molekülün konformer sayısıdır. Bu formülde δ , Kronecker δ fonksiyonunun bir çeşididir

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

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 değer katkısı 0.01 den daha azdır.) [56].

Bu formülde A_0 sabittir ve bu değeri belirlemek için, çalışılan molekül setinden (training set) aktivitesi bilinen bir referans molekül (I) seçeriz ve bu molekül için aktiviteyi (A_I) aşağıdaki eşitlik ile hesaplarız.

$$A_l = A_o \frac{\sum_{i=1}^{m_l} d_{li} [Pha] e^{-S_{li}} e^{-E_{li}/kT}}{\sum_{i=1}^{m_l} e^{-E_{li}/kT}} \quad (3.4)$$

A_n ve A_l için yazılan bu iki formül birleştirilirse A_0 lar birbirini götürüleceği için bilinmesine gerek kalmadan aşağıdaki formül elde edilir.

$$A_n = A_l \frac{\sum_{i=1}^{m_l} e^{-E_{li}/kT} \sum_{i=1}^{m_n} d_{ni} [Pha] e^{-S_{ni}} e^{-E_{ni}/kT}}{\sum_{i=1}^{m_n} e^{-E_{ni}/kT} \sum_{i=1}^{m_l} d_{li} [Pha] e^{-S_{li}} e^{-E_{li}/kT}} \quad (3.5)$$

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 [56].

Pha'ya sahip olan konformer P ile gösterilirse enerjisi E_{np} olur. (en düşük enerjili konformeri $E_{n0}=0$ olarak gösterilir.) Pha sahip birçok konformer varsa p biyoalıcıya en iyi bağlanandır ve genellikle en düşük E_{np} değerindedir. Reseptöre bağlandıktan sonra bağ yapmış durumdaki bu konformerin enerjisi $E_{np} + E'_{np} - \Delta$ olur. Burada E'_{np} APS ve AG grupların katkısı ve Δ molekül-biyoalıcı bağlanma enerjisidir. $\Delta > \delta$ olursa bu bağlanma en düşük enerjilidir. (Bkz. Şekil.3.1.) Termal dönüşümden sonra diğer tüm konformerlerin enerjileri E_{np} ve dengede bağlanma enerjisi $E_{ni} - [\Delta - (E_{np} + E'_{np})]$ olur.

$$A_n = A_0 \frac{\sum_i \exp\left[\left(-E_{ni} + \Delta - E_{np} - E'_{np}\right)/kT\right]}{\sum_i \exp\left(-E_{ni}/kT\right)} = A_0 \exp\left[\left(\Delta - E_{np} - E'_{np}\right)/kT\right] \quad (3.6)$$

$\exp\Delta$ sabittir ve A_0 a dahil edilebilir. Yukarıdaki formülde $S_{np} = E'_{np} / kT$ eşitliği kullanılarak

$$A_n = A_0 e^{-E_{np}/kT} e^{-S_{np}} \quad (3.7)$$

eşitliğini yazabiliriz. Aktivitesi A_l olarak bilinen, l referans bileşiği ile karşılaştırılarak A_0 belirlenir. Böylece eşitlik:

$$A_n = A_l e^{-\left(E_{np} - E_{lp}\right)/kT} e^{-\left(S_{np} - S_{lp}\right)} \quad (3.8)$$

haline dönüşür. Bu formül sadece, Pha ya sahip en düşük enerjili konformerin biyoaktivitesini tahmin etmek için kullanılır. Elde edilen bu son formül eşitlik 3.5'in basitleştirilmiş halidir. Çünkü eksponensiyal faktörden dolayı, $\exp(-E'_{np} / kT)$, yüksek enerjili konformerler ihmal edilebilir. (oda sıcaklığında E_{np} değeri yaklaşık olarak 1.5 kcal/mol olduğu için $\exp(-E'_{np} / kT)$ ifadesinin sayısal değeri yaklaşık 0.08 dir.) çoğu durumda Pha temel hal koformasyonunda bulunur. Bersuker Eşitlik 3.5'de gösterilen formül kullanarak bir bileşiğin tüm konformerlerini hesaba katarak aktivite hesaplaması için algoritma geliştiremediğinden dolayı sadeleştirilmiş formül olan ve en düşük enerjili bir konformerini hesaba katan eşitlik 3.8'i kullanmıştır. Diğer konformerlerin inaktif pozisyonda olduğunu kabul etmiştir. Bu çalışmanın en zor kısmını κ_j değerlerini hesaplama kısmı oluşturmaktadır. Bersuker çalışmalarında cross-validation yöntemlerinden olan leave one out metodunu kullanarak en uygun kappa sayısını ve değerini hesaplama yoluna gitmiştir [56].

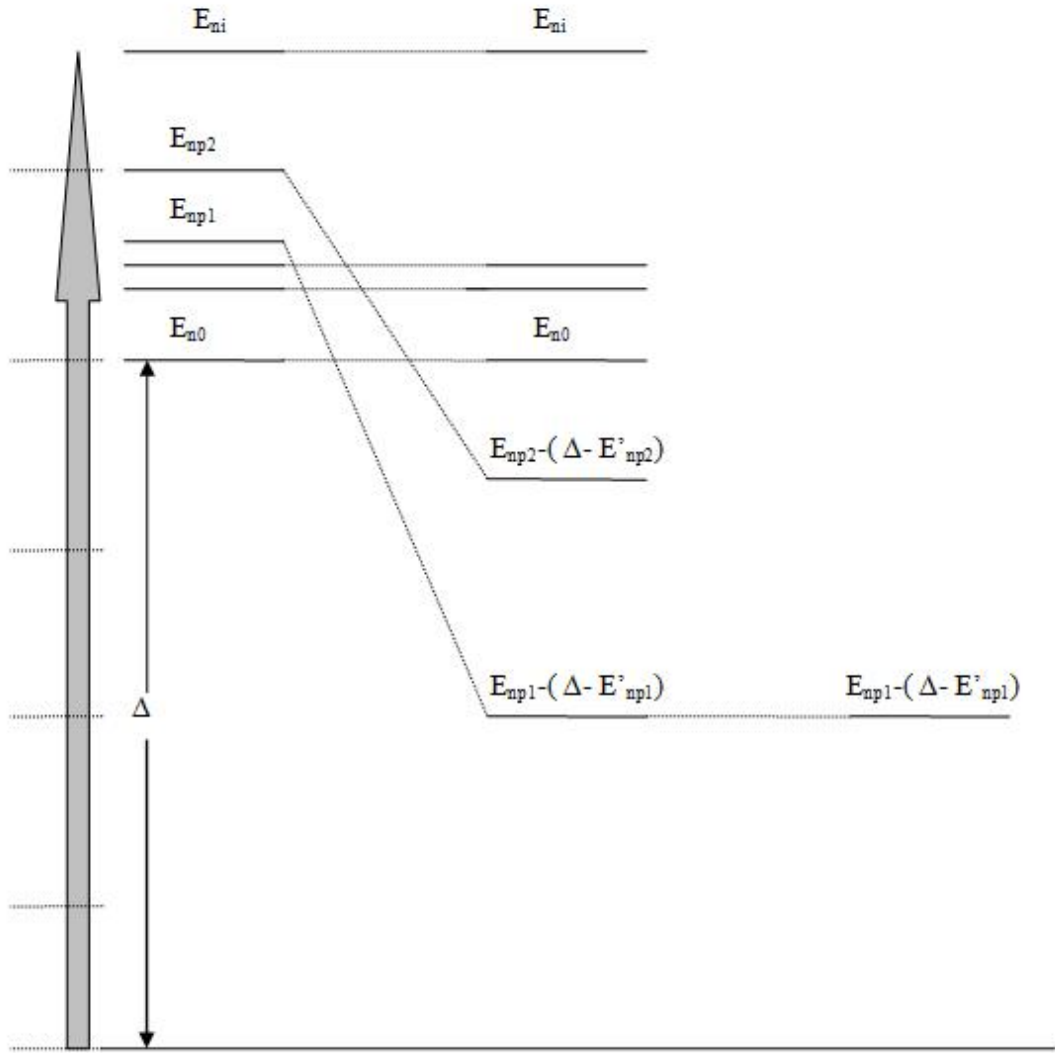
Molekül setindeki bileşiklerin bilinen deneysel aktiviteleri kullanılarak yukarıdaki formül çözülür ve hesaplanan aktivite değeri, A_n^{hes} , deneysel aktivite değeri, A_n^{den} , ve N

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. Bağlı durumda olmayan yani serbest haldeki ilaç molekülünün

(substrat) konformerlerinin ve bağ yapmış ilaç-biyoalıcı yapısının enerji karşılaştırmaları Şekil 3.1’de verilmiştir.



Şekil 3.1. Serbest haldeki ilaç molekülünün konformerlerinin ve bağ yapmış ilaç-biyoalıcı yapısının enerji karşılaştırmaları.

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 [56].

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

3.5. 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 [58]. Diğer bir önemli üstünlükleri ise çözümlerden oluşan popülasyonu 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” [59]. 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.

3.5.1. 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 [58].

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 $\delta(H)$ 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 [59] 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 imkanı ve uygulamaları olan genetik algoritmaların işleme adımları şöyle açıklanabilir [60].

Arama uzayındaki tüm mümkün çözümler dizi olarak kodlanır. Genellikle rastsal bir çözüm kümesi seçilir ve başlangıç popülasyonu 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 popülasyon olarak noktalar kümesini muhafaza eder. Her kuşakta, genetik algoritma, çaprazlama ve mutasyon gibi genetik operatörleri kullanarak yeni bir popülasyon oluşturur. Birkaç kuşak sonunda, popülasyon 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 [61].

3.5.2. 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 [61].

3.5.3. İlk Popülasyonun Oluşturulması

Olası çözümlerin kodlandığı bir çözüm grubu oluşturulur. Çözüm grubu popülasyon, çözümlerin kodları da kromozom olarak adlandırılır. İkili alfabenin kullanıldığı kromozomların gösteriminde, ilk popülasyonun oluşturulması için rastsal sayı üreticileri kullanılabilir. Rastsal sayı üreticisi çağrılır ve değer 0,5'den küçükse konum 0'a değilse 1 değerine ayarlanır [62].

3.5.4. Uygunluk Değerinin Hesaplanması

Bir kuşak oluşturulduktan sonraki ilk adım, popülasyondaki her üyenin uygunluk değerini hesaplama adımıdır. Örneğin, bir maksimizasyon problemi için i . üyenin uygunluk değeri $f(i)$, genellikle o noktadaki amaç fonksiyonunun değeridir [61]. 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 [59].

3.5.5. 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. Uygunluk değerinin hesaplanması adımından sonra mevcut kuşaktan yeni bir popülasyon 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 tanımadır. Bir dizinin kopyalanma şansı, uygunluk fonksiyonuyla hesaplanan dizinin uygunluk değerine bağlıdır [61]. Seçim yöntemlerine rulet tekerleği seçimi, turnuva seçimi ve sıralama seçimi gibi seçim yöntemleri örnek verilebilir.

3.5.6. 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 [61]. Genetik algoritmanın performansını etkileyen önemli parametrelerden biri olan çaprazlama operatörü doğal popülasyonlardaki çaprazlamaya karşılık gelmektedir. Çoğalma işlemi sonucunda elde edilen yeni popülasyondan rastsal olarak iki kromozom seçilmekte ve karşılıklı

çaprazlama işlemine tabi tutulmaktadır. Çaprazlama işleminde dizi uzunluğu L olmak üzere, $1 \leq k \leq L-1$ aralığında k tamsayısı seçilmektedir. Bu tamsayı değerine göre dizi çaprazlamaya uğratılır. En basit çaprazlama yöntemi tek noktalı çaprazlama yöntemidir.

3.5.7 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 [58]. 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 [63].

3.5.8. Yeni Kuşağın Oluşması ve Döngünün Durdurulması

Yeni kuşak çoğalma, çaprazlama ve mutasyon işlemlerinden sonra tanımlanmakta ve bir sonraki kuşağın ebeveynleri olmaktadır. Süreç yeni kuşakla çoğalma için belirlenen uygunluk ile devam eder. Bu süreç, önceden belirlenen kuşak sayısı kadar veya bir hedefe ulaşıncaya 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 [64].

3.5.9. Genetik Algoritmalarda Parametre Seçimi

Parametreler, genetik algoritma performansı üzerinde önemli etkiye sahiptir. Optimal kontrol parametreleri bulmak için bir çok çalışma yapılmıştır fakat tüm problemler için genel olarak kullanılabilecek parametreler bulunamamıştır [65]. 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 [59].

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 populyasyondaki 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

3.5.10. Genetik Algoritmaların Uygulama Alanları

Karmaşık problemleri hızlı ve optimale yakın olarak çözeabilen 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 [66]. Genetik algoritmalar özellikle çözüm uzayının geniş, süreksiz ve karmaşık olduğu problem tiplerinde başarılı sonuçlar vermektedir.

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

EC metot da farmakofor grubu belirlenirken serideki bileşiklerin bütün konformerleri dikkate alınırken aktivite hesaplamalarında hesaplama zorluğu ve uygun bir yöntemin bulunmamasından dolayı farmakofora sahip en düşük enerjili tek konformer dikkate alınmıştır. Bu metodun geçerliliğini kontrol etmek için bileşiklerden birini sırasıyla dışarıda bırakarak çapraz doğrulama (leave-one-out cross validation) yöntemi ile teorik aktivite hesaplamaları yapılmaktadır. Aktiviteye etki eden en iyi parametrelerin

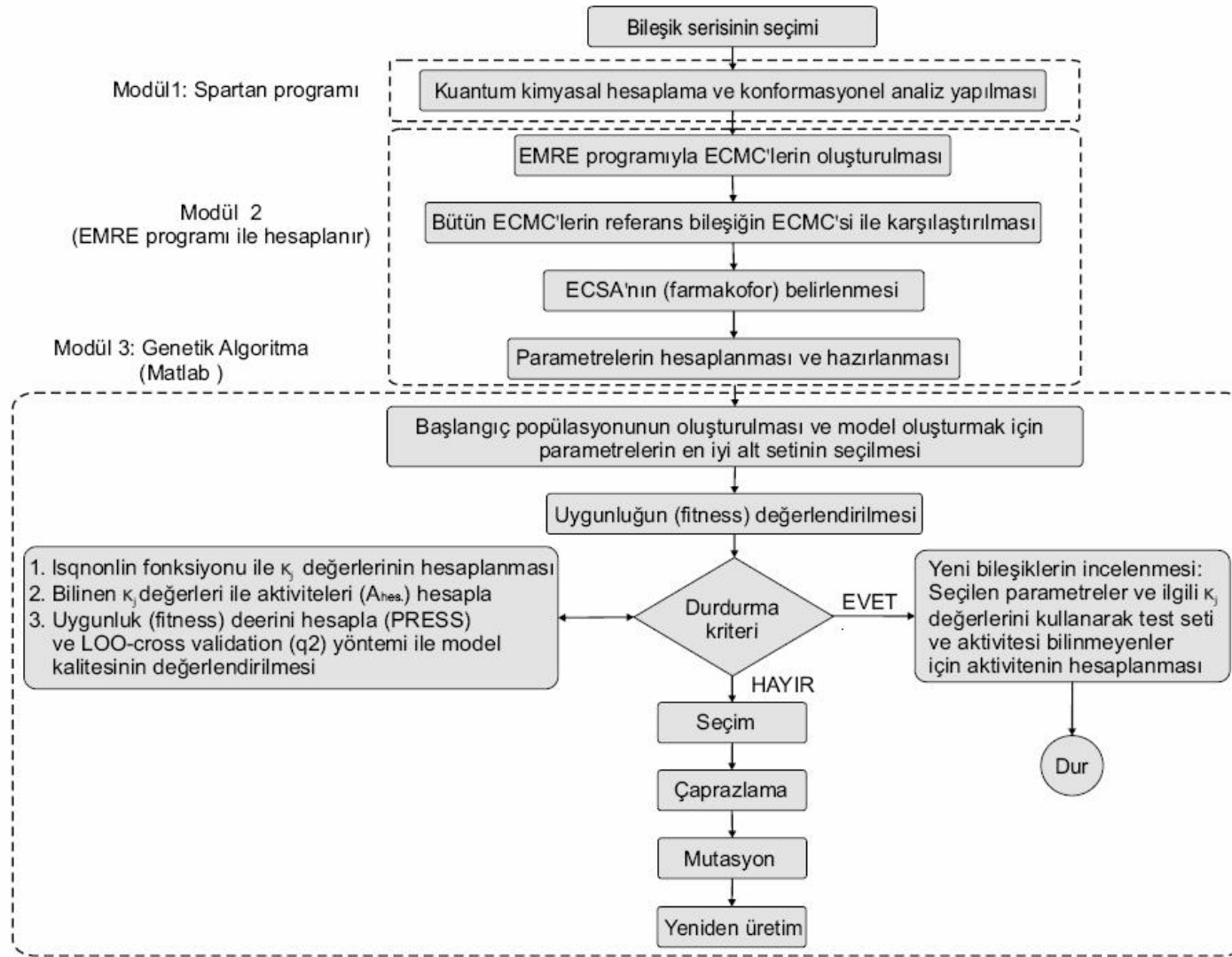
seiminde hangi yntemin kullanıldıđı net olarak anlařılamamıřtır. Sadece sonu kısmında seilen parametreler verilmiřtir. Parametrelerin geerliliđini kontrol etmek iin E istatistik yapılmıřtır [12,54-55]. Bu alıřmada tarafımızdan geliřtirilen paket programlarda ise (EMRE, ECSP programları ve Matlab 7.0 iin yazılan kodlar) btn konformerler aktivite hesaplamalarına dhil edilmiřtir. Pha grubu, aktiviteyi artıran (AG) ve aktiviteyi azaltan (APS) grupları tanımlamak ve biyolojik aktiviteyi hesaplamak iin elektron konformasyonel genetik algoritma (EC-GA) hibrit metodu geliřtirilmiřtir. EC-GA metodunda farmakofor, AG ve APS gruplarının belirlenmesi iin EC metod ile en iyi parametrenin seimi iin de genetik algoritma optimizasyon tekniđi birleřtirilmiřtir. Aktivite hesaplamalarında kullanılmak zere binlerce parametre hazırlanabilmektedir. Hazırlanan bu parametrelerin iinden aktivite zerinde en etkili olanların seimi rneđin 100 parametre ierisinden en iyi 10 parametrenin seimi 1.73×10^{13} gibi bir iřlem gerektirmektedir ve bu da uzun 10-20 yıl gibi sreler gerekmektedir. Bizim alıřmalarımızda bir seri iin yaklařık 600-700 parametre hazırlandıđı dřnlrse bu sayı daha da artmaktadır ve sre daha da uzamaktadır. EC-GA Metodunda aktiviteye etki eden en iyi parametreleri belirlemek ve kısa zamanda sonuca ulařmak iin genetik algoritma optimizasyon tekniđi kullanılmıřtır. Bu metotta molekl ve konformerlerini tanımlayacak yzlerce parametre hazırlamak ve bunların iinden hem eđitim hem de test seti iin en uygun olan parametreleri semek mmkndr. nk bazı parametreler eđitim seti iin uygun olurken test seti iin kt olabilmektedir. Aktivite hesaplamaları eřitlik 3.5 kullanılarak lineer olmayan en kk kareler yntemine gre yapılmıřtır. Aktivite hesaplamaları bileřiklerden biri her defasında dıřarıda bırakılarak apraz dođrulama (leave-one-out cross validation) yntemi ile test edilir. İleride ayrıntılı olarak aıklanacađı gibi geliřtirilen bu yazılımların tamamı řekil 3.2’de gsterilen ila tasarımında EMRE yazılım sistemi řeklinde isimlendirilmiřtir. řekilden grldđ gibi EC-GA metodu ierisinde Spartan paket programı, alıřma grubumuz tarafından geliřtirilen ECSP programı ve son olarakta yine alıřma grubumuz tarafından geliřtirilen, aktiviteyi ve κ_j deđerlerini hesaplamak iin Matlab 7.0 paket programında genetik algoritma optimizasyon tekniđini kullanarak hesaplamaları gerekleřtiren yeni bir yazılım olmak zere  farklı program kullanılmaktadır. Geliřtirilen bu yazılımlar alıřmanın ileriki kısımlarında ayrıntılı olarak aıklanacaktır. řekil 3.2’de modl 1 kısmında spartan paket programı kullanılarak kuantum kimyasal hesaplamalar ve konformasyonel analiz yapılmaktadır.

Modül 2 kısmında EMRE programı ile bileşik serilerindeki her bir bileşiğin bütün konformerleri için ECMC oluşturulmakta daha sonra oluşturulan bu ECMC ler referans bileşiğin ECMC si ile karşılaştırılarak ECSA (farmakofor) belirlenmektedir. Daha sonra yine EMRE programı kullanılarak aktivite hesaplamalarında kullanılmak üzere bütün bileşiklerin bütün konformerleri için parametreler hazırlanır. Modül 3 kısmında hazırlanan bu parametreler içerisinde aktivite üzerinde en etkili olanları bulmak için genetik algoritma optimizasyon tekniği kullanılır ve aktivite hesaplamalarına geçilir.

EC-GA metodu kapsamında daha önce gurubumuz tarafından farklı seriler üzerinde çalışmlr yapılmış bunlarda bir kısmı uluslar arası dergilerde yayınlanmıştır [67-70]. Yapılan lisansüstü tezler de çok sayıda bileşik için çalışmalar yapılmış bunlardan bazıları aşağıda verilmiştir.

Her bir seride ECMC lerin karşılaştırılması ile farmakofor gruplar belirlenmiş ve aktiviteye en fazla etkisi olan alt parametre setini seçmek ve teorik aktivite değerlerini hesaplamak için genetik algoritma ve doğrusal olmayan en küçük kareler (lsqnonlin) yöntemi kullanılmıştır. Nifedipin analogu 1,4-dihidropiridin, nitroimidazolil analogu 1,4-dihidropiridin, artemisinin ve triazolpiridin oksazol serileri için çalışmalar yapılmış olup sonuçlar tezde verilmiştir [71]. Pirazol, benzotriazin, tetrahidrodibenzazosin, 4-(3-bromoanilino)-6,7-dimetoksikinazolin serileri için çalışmalar yapılmış olup sonuçlar tezde verilmiştir [72]. Penisilin türevleri için çalışmalar yapılmış olup sonuçlar tezde verilmiştir [73]. Kinolin serisi için çalışmalar yapılmış olup sonuçlar tezde verilmiştir [74]. Bu çalışmalarda geliştirilen 4D QSAR metodu ile oluşturulan modellerin benzer yapılara sahip yeni bileşiklerin tasarlanmasında kullanılabileceği görülmüştür.

1-[(2-hidroksietoksi)metil]-6-(feniltiyo)timin (HEPT), tetrahidro-imidazo[4,5,1-jk][1,4]-benzodiazepin-2(1H)-on ve -tion (TIBO), dihidro-alkil (yada siklo-alkil)tiyo-benzil-oksopirimidin (DABO), ve 2,3-Diaril-1,3-tiyazolidin-4-on serileri için çalışmalar yapılmış olup sonuçlar tezde verilmiştir [75].



Şekil 3.2. İlaç tasarımında EMRE yazılım sisteminin basamakları.

3.6.1. 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 kazandırabilmektir. 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 programcılarının 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 ulaşip 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.. Bu mantık çerçevesi içinde bu yazılımlar EMT, ECM ve EC-GA metotlarını da kapsamaktadır. Bu yazılımlar kullanılarak bütün bileşik serileri için ECMC matrisleri, geometrik, elektronik ve fizikokimyasal parametreler hazırlanabilir ve teorik aktivite hesaplamaları yapılabilir. Şekil 3.3’de ise EC-GA metodunda takip edilen aşamalar gösterilmektedir.

Adım 1: Aktivitesi deneysel olarak bilinen bileşik serisi seçilir.
Adım 2: Bileşiklerin konformasyonel analizi yapılır. Yüksek enerjili ve sanal frekansa sahip olan konformerler elimine edilir
Adım 3: ECSP veya EMRE V2.0 programı kullanılarak serideki bileşiklerin ECMC’leri hazırlanır.
Adım 4: ECSP programını kullanılarak farmokofor grubu temsil eden alt matris (ECSA) belirlenir.
Adım 5. Aktiviteyi artıran (AG) ve azaltan (APS) gruplar olduğu tahmin edilen değişkenler hazırlanır.
Adım 6: Genetik algoritma optimizasyon tekniği kullanılarak MATLAB 7.04 programında teorik aktivite hesabı yapılır ve istatistiksel analizlere geçilir. Seri eğitim ve test seti olmak üzere ikiye bölünür ve aktivite hesaplamaları yapılır.

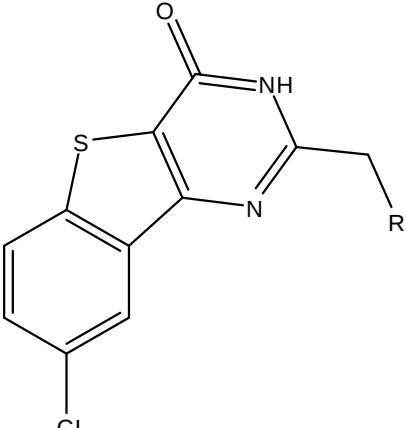
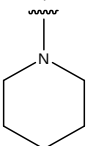
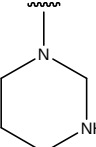
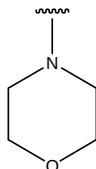
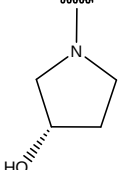
Şekil 3.3. EC-GA metodunda takip edilen aşamalar

3.6.2. Adım 1

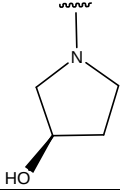
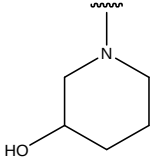
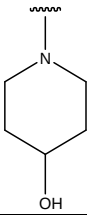
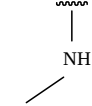
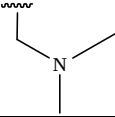
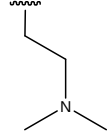
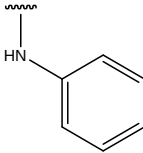
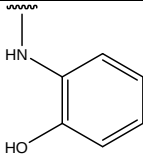
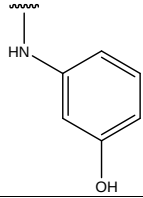
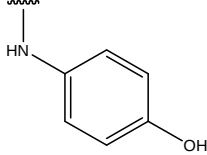
QSAR modelinin geliştirilmesindeki en önemli etken deneysel verilerin doğru ve güvenilir olmasıdır. Deneysel veriler doğru ve kesin olmaması durumunda anlamlı bir model elde etmek ve geliştirmek imkânsızdır. Uygun bir QSAR modelinin geliştirilmesi için bu verilerin benzer kimyasal bileşik serilerinden oluşması gerekmektedir. Bu nedenle bu çalışmada doğruluğuna ve sonuçların güvenilir olduğuna inandığımız benzotiyenopirimidin serisi literatürden alınmıştır [20].

Bu çalışmada ele alınan bileşik serisi aşağıda Tablo 3.1’de verilmiştir. Benzotiyenopirimidin serisi için temel iskelete bağlı olan süstitüentler, hesaplamalarımız sonucu belirlenen konformer sayıları ve deneysel aktivite değeri yer almaktadır.

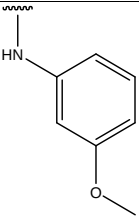
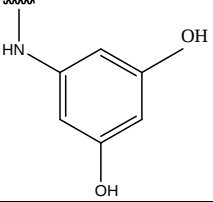
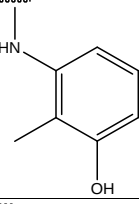
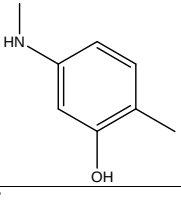
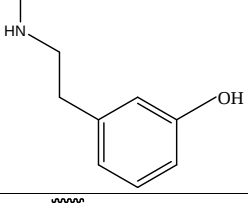
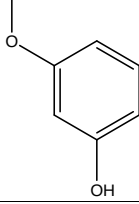
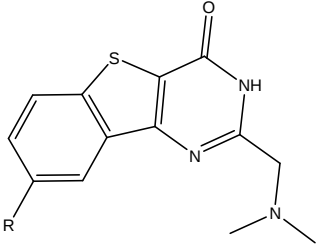
Tablo 3.1. Benzotiyenopirimidin Serisine Ait Temel İskelet, Süstitüentler ve Konformer Sayıları

		
No	R	Konformer Sayısı
1	H	1
2		2
3		4
4		6
5		4
6		13

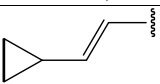
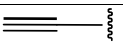
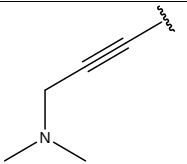
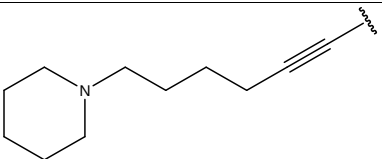
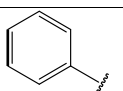
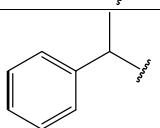
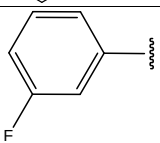
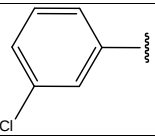
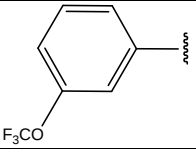
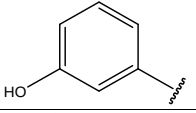
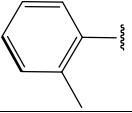
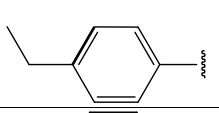
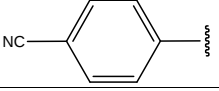
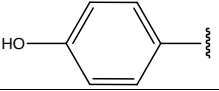
Tablo 3.1'in Devamı

7		14
8		23
9		18
10		5
11		5
12		12
13		2
14		12
15		7
16		6

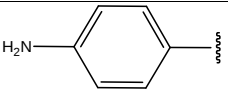
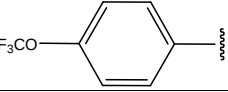
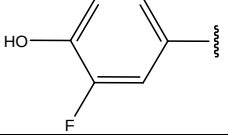
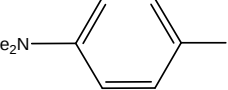
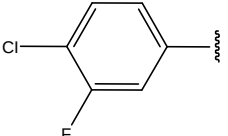
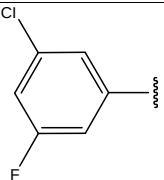
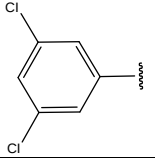
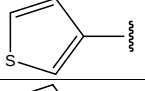
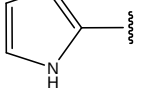
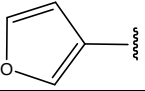
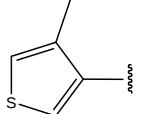
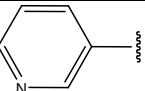
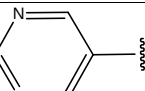
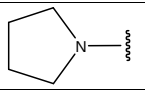
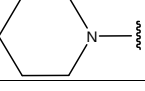
Tablo 3.1'in Devamı

17		7
18		6
19		7
20		5
21		3
22		8
		
No	R	Konformer Sayısı
23	Cl	2
24	Br	2
25	H	2
26	Me	4
27	CF ₃	4
28	Et	4

Tablo 3.1'in Devamı

29		5
30		9
31		23
32		2
33		15
34		24
35		4
36		14
37		8
38		8
39		20
40		16
41		6
42		8
43		4
44		8

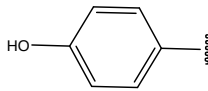
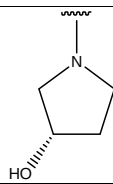
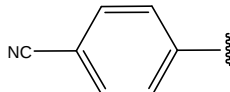
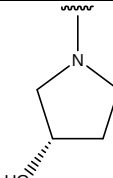
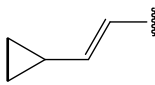
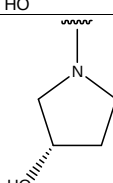
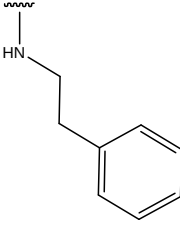
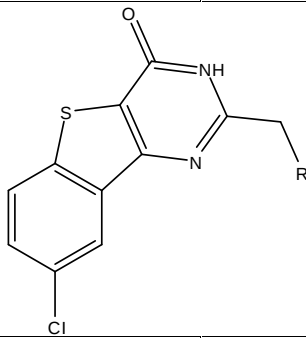
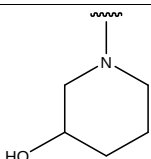
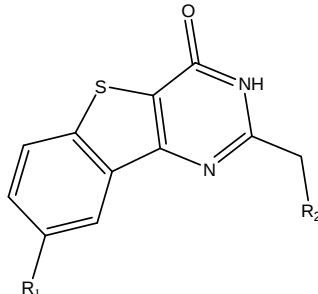
Tablo 3.1'in Devamı

45		8
46		12
47		12
48		9
49		8
50		10
51		4
52		8
53		6
54		8
55		4
56		8
57		4
58		5
59		15

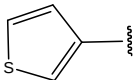
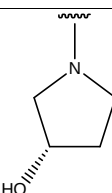
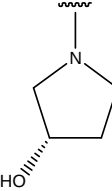
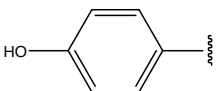
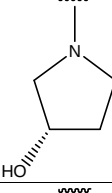
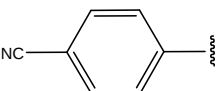
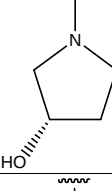
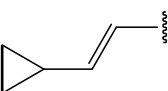
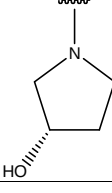
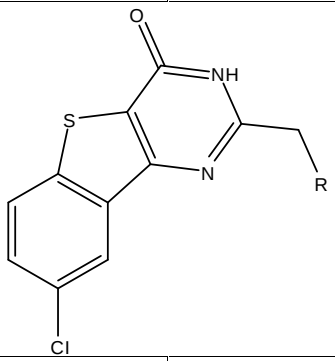
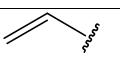
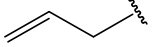
Tablo 3.1'in Devamı

60		20	
61		24	
No	R1	R2	Konformer Sayısı
62	Br		7
63			16
64	Ph		11
65			25
66	Ph		17

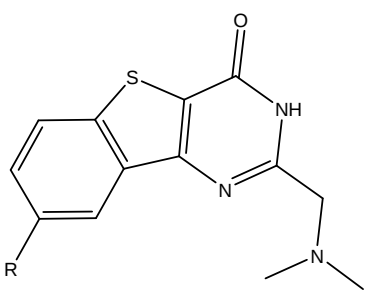
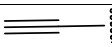
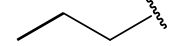
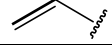
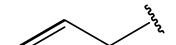
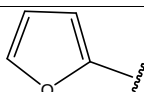
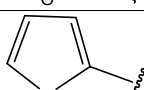
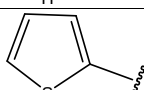
Tablo 3.1'in Devamı

67			26
68			17
69			26
70	Br		26
			
No	R	Konformer Sayısı	
71		20	
			
No	R1	R2	Konformer Sayısı

Tablo 3.1'in Devamı

72			5
73	Ph		5
74			4
75			3
76			17
			
No	R	Konformer Sayısı	
77	Me	2	
78	Et	3	
79	Ph	1	
80	Ph-CH ₂	5	
81		2	
82		9	

Tablo 3.1'in Devamı

		
No	R	Konformer Sayısı
83	NO ₂	1
84	OH	3
85	NH ₂	2
86		1
87	NO ₂	3
88	OH	4
89		14
90	CN	2
91		11
92		13
93	NH ₂	4
94		8
95		8
96		8

3.6.3. Adım 2

Çalışmanın bu basamağında Spartan paket programında molekül iskeletlerinin çizimi yapıp, 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 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.

3.6.3.1. Boltzman Dağılımı

Boltzman dağılımı; farklı enerjili hallerin popülasyonlarını hesaplamak için ortaya atılan bir formüldür. Boltzmann dağılımının bilinmesi gereken en önemli özelliklerinden biri popülasyon 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

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

şeklinde tanımlar. Boltzmann dağılımının yorumlanmasına ilişkin önemli bir nokta, bu ifadenin seviyelerin değil enerji hallerinin popülasyonlarına işaret ettiğidir [9]. 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 3.3’de benzotiyenopirimidin serisi için konformerler, her bir konformerin relatif enerjisi ve bu enerjilere karşılık gelen Boltzmann değerleri verilmiştir. Tabloda da görüldüğü gibi konformerlerin bağıl enerjileri arttıkça Boltzmann dağılımı değerleri azalmaktadır.

Tablo 3.1.1 Benzotiyenopirimidin serisinden 61. Bileşiğin 24 konformeri için bağıl enerji ve boltzmann değerleri

Konformer	rel. E (kcal/mol)	Boltzmann Dağılımı
h61_1	0	0.1085
h61_2	0.02	0.1052
h61_3	0.05	0.0995
h61_4	0.06	0.0978
h61_5	0.07	0.0969
h61_6	0.07	0.0959

h61_7	0.16	0.0826
h61_8	0.73	0.0318
h61_9	0.77	0.0298
h61_10	0.82	0.027
h61_11	0.84	0.0264
h61_12	0.84	0.0261
h61_13	0.85	0.0257
h61_14	0.85	0.0257
h61_15	0.87	0.0251
h61_16	0.93	0.0228
h61_17	1.17	0.015
h61_18	1.3	0.0122
h61_19	1.35	0.011
h61_20	1.39	0.0104
h61_21	1.48	0.009
h61_22	1.58	0.0075
h61_23	1.92	0.0043
h61_24	1.98	0.0038

Tablo 3.1.2 Benzotiyenopirimidin serisinden 69. bileşiğin 27 konformeri için bağlı enerji ve boltzmann değerleri

Konformer	rel. E (kcal/mol)	Boltzmann Dağılımı
h69_1	0	0.2032
h69_2	0.31	0.1207
h69_3	0.69	0.0635
h69_4	0.71	0.0618
h69_5	0.72	0.06
h69_6	0.73	0.059
h69_7	0.78	0.0548
h69_8	0.95	0.0411
h69_9	0.99	0.0385
h69_10	1.09	0.0323
h69_11	1.11	0.031
h69_12	1.13	0.0302
h69_13	1.31	0.0223
h69_14	1.33	0.0217
h69_15	1.34	0.0212
h69_16	1.4	0.0191
h69_17	1.48	0.0167
h69_18	1.55	0.0148
h69_19	1.57	0.0142
h69_20	1.61	0.0134
h69_21	1.72	0.0111
h69_22	1.73	0.0109
h69_23	1.76	0.0104

h69_24	1.79	0.0099
h69_25	1.79	0.0098
h69_26	1.90	0.0082

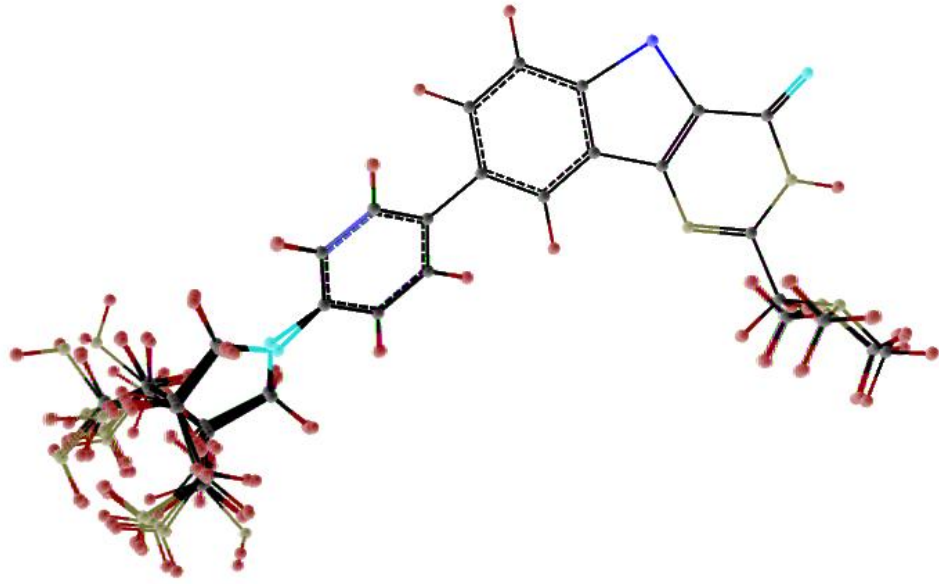
Burada konformerler elimine edilirken boltzmann değerleri dikkate alınmıştır. Virgülden sonra 4. hanesi sıfır olanlar silindi. Aktivite formülü üstel bir fonksiyon olduğu için Boltzmann değeri küçüldükçe aktiviteye katkısı çok az olmaktadır. Kalan konformerler için kendi aralarında bazı atomlar üzerinden 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(yada Pha) bulunması için kullanıldı.

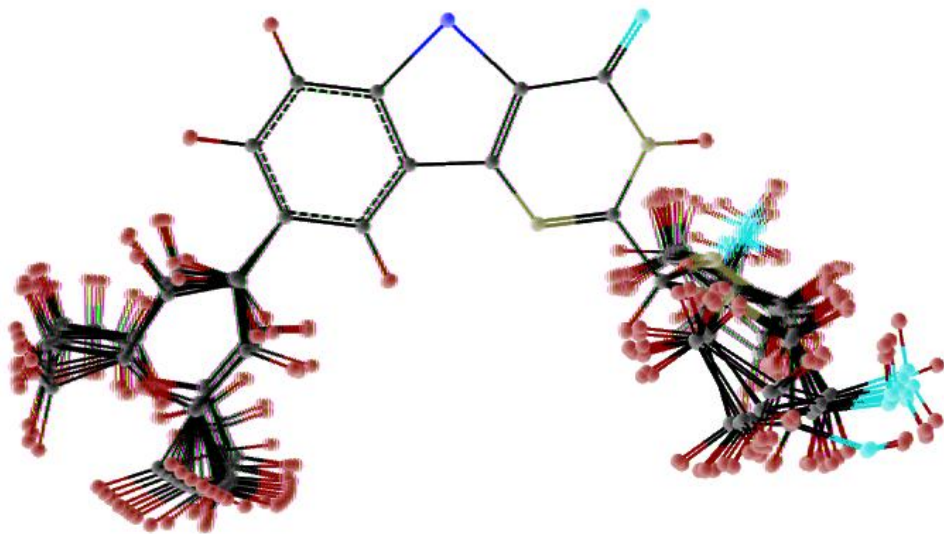
Moleküler özelliklerin (sterik durum, elektrostatik potansiyel ve hidrofobisite gibi) üç boyutta göz önüne alınması, aktif ve inaktif bileşikler arasındaki benzerlikler ve farklılıkları belirlemek için ilk adımdır. Bir bileşik serisi için genel olan üç boyutlu yapısal özellikleri ortaya çıkarmak için 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. 96 bileşikli benzotiyenopirimidin 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

olacađından benzotiyenopirimidin serisine ait 61. ve 69. bileřikleri iin sırasıyla řekil 3.5-3.6'da, konformerlerinin st ste akıřtırılma iřlemi gsterilmiřtir. Burada 61. bileřik Pim-1 iin aktivitesi en yksek bileřiktir. Pim-2 iin ise aktivitesi en yksek olan bileřik 69. bileřiktir.



řekil 3.5. Benzotiyenopirimidin serisinden h61 bileřiđine ait 24 konformerin st ste akıřtırılması

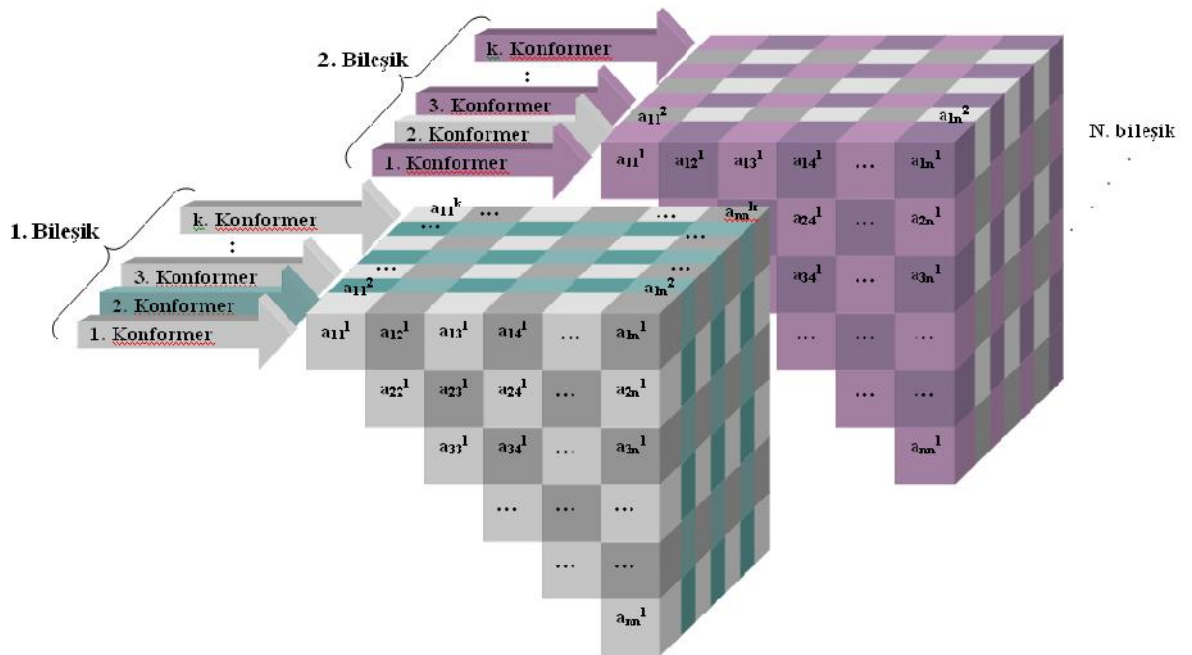


řekil 3.6 Benzotiyenopirimidin serisinden h69 bileřiđine ait 27 konformerin st ste akıřtırılması

3.6.4. Adım 3

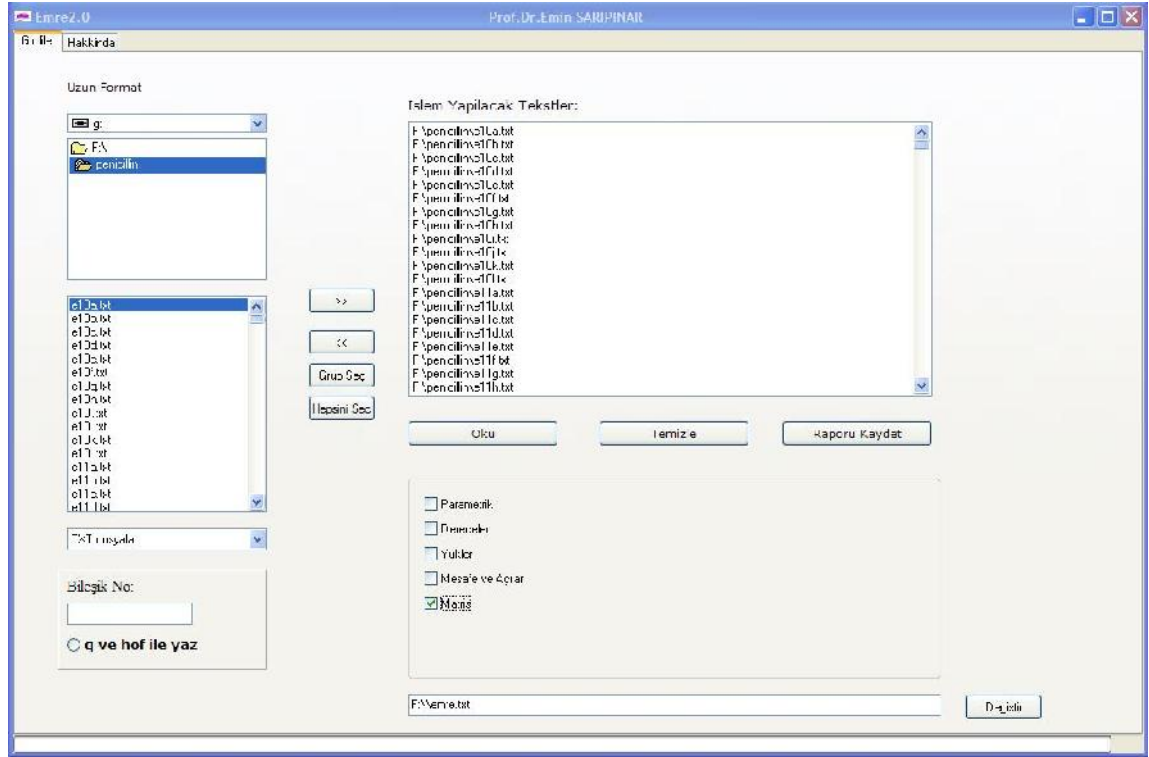
Referans bileşiğin (aktivitesi en yüksek olan bileşiğin en düşük enerjili konformeri referans olarak seçilir) fonksiyonel atomlarına göre serideki bileşiklerin her bir konformeri için ECMC matrisleri hazırlanmıştır. Diyagonal elemanlarına göre simetrik olan bu kare matrisin sadece üst kısmını içeren 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 3.7’de 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.



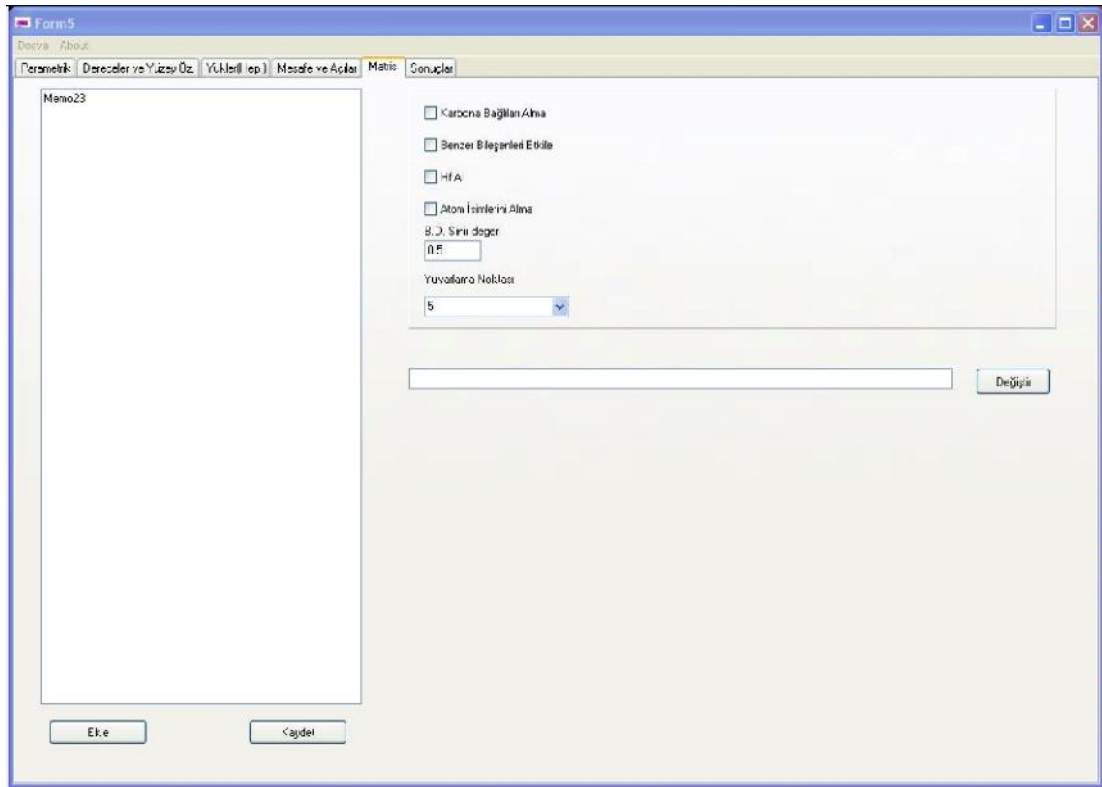
Şekil 3.7. Üç boyutlu ECMC'nin gösterimi

Şekil 3.8’de EMRE V2.0 programına ait girdiler sekmesi gösterilmektedir.



Şekil 3.8. EMRE V2.0 programına ait girdiler sekmesi

Şekil 3.9’da EMRE V2.0 programına ait matris sekmesi gösterilmektedir.



Şekil 3.9. EMRE V2.0 programına ait matris sekmesi

3.6.5. 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 [76]. 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 yapılabilir. I.B.Bersuker ise bileşiklerin konformerlerinin de hesaba katarak üç boyutlu matrislerini kullanarak farmakofor grubunu belirlemiştir. Bunun için referans bileşiğine göre diğer tüm konformerler % bağıl sapma olarak hesaplanmıştır. Grubumuz tarafından geliştirilen ECSP programı ise bileşiklerin en düşük enerjili tek konformerini hesaba kattığı gibi bütün konformerlerinin de 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 yapıya göre oluşturulmuştur. Şekil 3.11'de 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 karşılaştırılarak yazılmıştır. Bu dosya alt matris hesaplaması yapılırken atomların karşılaştırılmasında programın çalışma hızını artırmakta ve çalışma grubumuza oldukça büyük bir kolaylık sağlamaktadır.

h1	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h2	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h3	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h4	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h5	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h6	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h7	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h8	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h9	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h10	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h11	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h12	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h13	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h14	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h15	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h16	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1

h70	C2	C3	C1	C5	C6	C4	S1	C8	C7	N2	C12	C9	N1	H1	C13	O1
h71	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h72	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C11	N1	H7	C7	O1
h73	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C21	N1	H7	C7	O1
h74	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C21	N1	H7	C7	O1
h75	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C22	N1	H7	C7	O1
h76	C1	C2	C3	C4	C5	C6	S1	C9	C8	N2	C10	C22	N1	H7	C7	O1
h77	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h78	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h79	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h80	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h81	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h82	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h83	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h84	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h85	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h86	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h87	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h88	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h89	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h90	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h91	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h92	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h93	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h94	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h95	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1
h96	C1	C2	C3	C4	C5	C6	S1	C9	C8	N1	C10	C11	N2	H5	C7	O1

Şekil 3.11. Benzotiyenopirimidin 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 3.13’de verilmektedir.

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 etmek 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 [77] 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 bileşikler aktif (yüksek aktiviteli) ve inaktif (düşük aktiviteli) olmak üzere ikiye bölünür. Aktivitesi yüksek olan moleküllerin belirli tolerans değerleri içinde veya en küçük bağl 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 $P(21,3) \times P(20,3) + P(21,4) \times P(20,4) + \dots + P(21,20) \times P(20,20) = 1.98 \times 10^{38}$ 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.

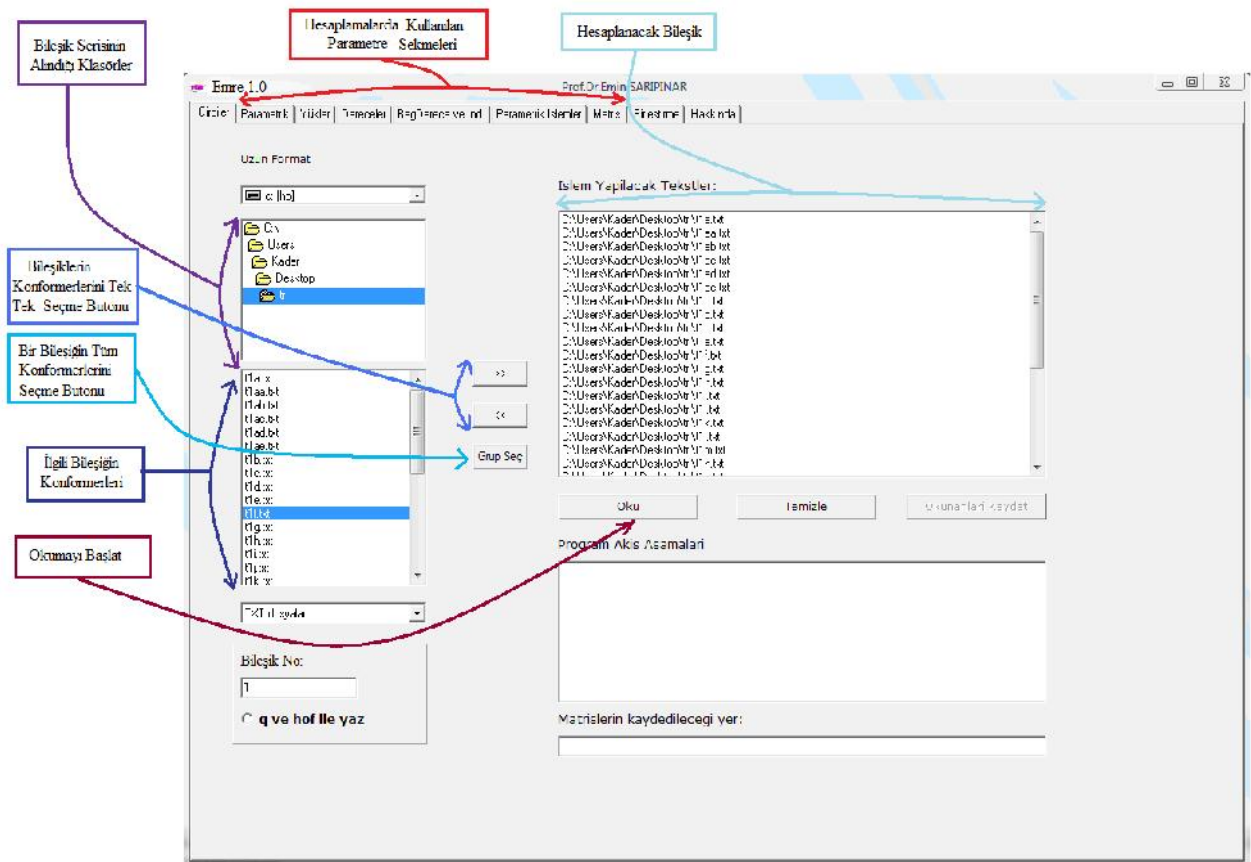
Bu işlemden sonra diğer aktif moleküllerin en düşük enerjili konformerlerinin ECMC’ lerinden bir alt matris (Pha) belirlemek için referans bileşiğin en düşük enerjili konformerine ait matris ile karşılaştırmalar yapılır. Karşılaştırma sonucu ortaya çıkan olası alt matrisleri elimine etmek için ETM’ de olduğu gibi farmakoforun bulunabilme ihtimalini gösterebilmek için p_a ve α_a tanımları yapılmıştır, genellikle 0.6’ dan büyük değerleri yeterli olarak kabul edilmiştir. $P_a = (n+1)/(n_1+n_3+2)$, $\alpha_a = (n_1 * n_4 -$

$n_2 \cdot n_3 / (N_1 \cdot N_2 \cdot N_3 \cdot N_4)^{1/2}$ burada n_1 ve n_2 sırasıyla aktif bileşikler içinde aktif fragmanre sahip olan ve olmayan moleküllerin sayısı n_3 ve n_4 ise inaktif bileşikler içinde aktif fragmente sahip olan ve olmayan moleküllerin sayısıdır. N_1 ve N_2 sırasıyla aktif ve inaktif bileşiklerin sayısını verir. $N_3 = n_1 + n_3$ $N_4 = n_2 + n_4$

Aşağıdaki matrisde her bir bileşik için en düşük enerjili bir konformer dikkate alındığında ortaya çıkan alt matris elemanları görülmektedir. Dikkat edildiğinde inaktif olarak kabul edilen bileşiklerde yükler ve uzaklık için tolerans değerlerinin aktif olan bileşiklerde daha büyük olduğu görülmektedir. Literatürdeki çalışmalara göre inaktif moleküllerde tolerans değeri daha büyük olduğu görülmektedir [78]. Bunun için hesaplanan p_a ve α_a değerleri matrisin yanında verilmiştir. En düşük enerjili konformerde matris elemanı bulunmaması durumunda daha yüksek enerjili konformerlerde tarama yapılabilir [77].

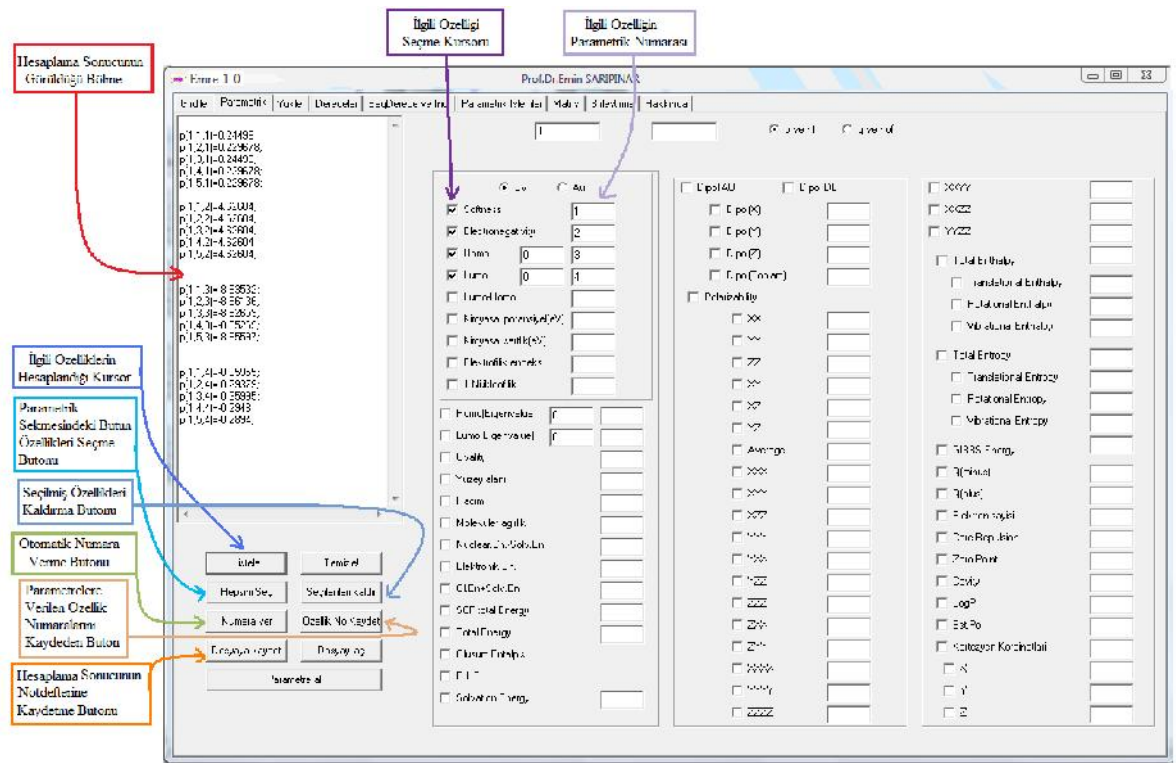
3.6.6. Adım 5

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, fizikokimyasal ve yük gibi özellikleri belirlemek için APS ve AG değerlerini kantitatif olarak karakterize eden S fonksiyonu eşitlik 3.2’de formüle 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 V2.0) ile aktivite üzerinde etkisi olduğu tahmin edilen bu değişkenler düzenlenir. Şekil 3.14’de EMRE V2.0 programına ait girdiler sekmesi gösterilmiştir. Aşağıdaki şekilde girdiler sekmesinin ilgili butonları tek tek gösterilmiştir.



Şekil 3.14. Emre programına ait girdiler sekmesi

Şekil 3.15’de EMRE V2.0 programına ait parametrik sekmesi gösterilmiştir. İlgili bileşik serisinin hesaplanmasında Parametrik sekmesinde hangi özellikler kullanılacaksa ilgili özellikler buradan seçilmektedir.



Şekil 3.15. Emre programına ait parametrik sekmesi

3.6.6.1. 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 [79].

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

$$\eta \approx \frac{1}{2}(I - A) \approx \frac{1}{2}(\epsilon_{\text{LUMO}} - \epsilon_{\text{HOMO}}) \quad (3.10)$$

$$I \approx -\epsilon_{\text{HOMO}} \quad \text{ve} \quad A \approx -\epsilon_{\text{LUMO}} \quad (3.11)$$

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 [79].

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 formülle hesaplanır fakat sonuç eV (electron volt) cinsinden bulunur.

Softness: Softness kimyasal reaktivitenin derecesini ölçen moleküllerin bir özelliğidir.

Ayrıca Hardness in tersidir. S şeklinde gösterilir.

$$S = \frac{1}{\eta} \quad (3.12)$$

Kimyasal potansiyel: μ şeklinde gösterilir ve aşağıdaki formülle hesaplanır [79].

$$\mu \approx -\frac{1}{2}(I + A) \approx \frac{1}{2}(\varepsilon_{\text{HOMO}} + \varepsilon_{\text{LUMO}}) \quad (3.13)$$

Elektronegatiflik: Bir molekülde ki atomun elektronları kendine çekme kapasitesidir.

Kimyasal potansiyelin zıt işaretlisidir. χ şeklinde gösterilir.

$$\chi = -\mu \quad \chi = \frac{I + A}{2} \quad (3.14)$$

Elektrofilik indeks: Donör ve akseptör arasındaki elektron alışverişi sırasında oluşan enerji azalmasının bir ölçüsüdür. (ω) şeklinde gösterilir ve aşağıdaki formülle hesaplanır [79].

$$\omega = \frac{\mu^2}{2\eta} \quad (3.15)$$

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

Hacim ve alan oranına bakılarak aşağıdaki şekilde tanımlanır.

$$O = A / (4 * \pi * ((3 * V) / 4 * \pi) ^ {2/3}) \quad (3.16)$$

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:

$$\mu = \delta d \quad (3.17)$$

$+\delta - \delta^-$ 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 elektrostatik 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.

$$P = \alpha E \quad (3.18)$$

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.

$$\text{Bağ derecesi} = \frac{\text{Bağ MO deki } \epsilon \text{ sayısı toplamı} - \text{Karşıbağ MO } \epsilon \text{ sayısı}}{2} \quad (3.19)$$

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. Oktan-ol-su çözünme katsayısı iki çözücü olan su ve oktanoldaki diferansiyel çözünürlüğün bir ölçüsüdür. Bu ise bir maddenin bağıl hidrofobisite ve hidrofilisitesinin bir ölçüsüdür.

İlaç benzeri maddelerin içeriğindeki moleküler büyüklük ile birlikte, dönen bağların sayısı ve hidrojen bağlarına yatkınlık, LogP değeri; absorpsiyon, 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 V2.0 programı kullanılarak, S fonksiyonundaki formata uygun şekilde hazırlanır. EMRE V2.0 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 V2.0 programında

gerçekleştirilmiştir. Son olarak yine bu programda düzlem ile seçilen iki atom arasındaki açığı ölçmek için yine bu program kullanılmıştır.

Fukui Atomik Nükleofilik Reaktivite indeksi

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

Basit formül
$$N'_A = \sum_{i \in A} C_{iHOMO}^2$$

$$N_A = \sum_{i \in A} C_{iLUMO}^2 / (1 - e_{LUMO})$$

$$N'_A = \sum_{i \in A} C_{iLUMO}^2$$

3.6.7. 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 3.5 kullanılarak yapılmıştır.

Çalışmalarda n_j adedi, popülasyon sayısı ve n_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 3.20 kullanılır. Standart hatayı hesaplamak için eşitlik 3.21'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.

$$r^2 = \frac{\left(\sum xy - \frac{\sum x \sum y}{n} \right)^2}{\left(\sum x^2 - \frac{(\sum x)^2}{n} \right) \left(\sum y^2 - \frac{(\sum y)^2}{n} \right)} \quad (3.20)$$

$$se = \sqrt{\frac{1-r^2}{n-2}} \quad (3.21)$$

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

3.6.7.1. 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 3.16'da 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 3.16. 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 dahilinde 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.

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

1. Problemin hedefi içeren bir hedef fonksiyonu olacaktır.
2. Tasarım değişkenleri olmalıdır.
3. Problemdaki 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 ifadelerden denir

Lineer olmayan 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.

3.6.7.2. 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 geliştirilmiştir. Bu çalışmada aktivite denklemini çözmek ve aktiviteye en çok etki eden parametre grubunu bulmak için MATLAB 7.04 programında optimizasyon metodlarından 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 3.17'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 V2.0 programı kullanılarak hazırlanan parametre (özellik) listesinden rasgele seçilen özellikler ile bu matrisin elemanları oluşturulur. Matrisin her bir satırı (kapa seti) lsqnonlin fonksiyonuna gönderilerek k 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:

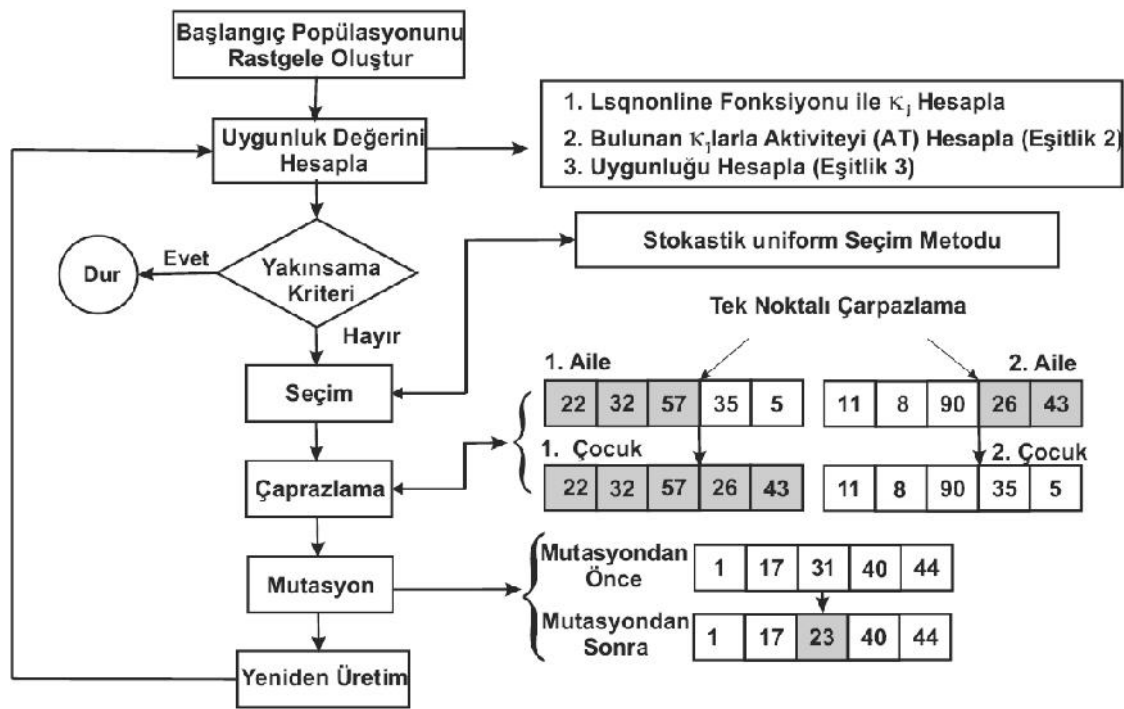
$$f = \text{PRESS}_N = \sum_{n=1}^N \left| A_n^{\text{den}} - A_n^{\text{hes}} \right|^2 \quad (3.22)$$

A_n^{den} : Deneysel aktivite değeri

A_n^{hes} : Hesaplanan aktivite değeri

n: Bileşik sayısı

Çalışmalarda κj adedi, popülasyon sayısı ve κj 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 3.17. Genetik algoritmanın aşamaları

3.7. İstatiksel Analiz

Çalışmalarda κ_j adedi, popülasyon sayısı ve κ_j limitleri değiştirilerek çeşitli denemeler yapılmış değişik bağıl hataya sahip birçok sonuç elde edilmiş ve bu sonuçların içinden en uygunlarını seçmek için istatistiksel olarak değerlendirme işlemine geçilmiştir. İstatistiksel değerlendirme işlemi regresyon analizi ve çapraz doprulama yapılarak gerçekleştirilmiştir. 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^2 kavramını hesaplayabilmek için ilk olarak eşitlik 3.26 kullanılır. Standart hatayı hesaplamak için eşitlik 3.27’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. Ayrıca kurulan modelin doğruluğunu test etmek için çapraz doğrulama katsayısı (q^2) eşitlik 3.28 .ile hesaplanmıştır. q^2 değerinin 0.35 den büyük değeri kabul edilebilir değerdir. Ayrıca q^2 nin 1’e yakınlığı kurulan modelin doğruluğunu gösterir [80].

$$r^2 = \frac{\left(\frac{\sum xy - \frac{\sum x \sum y}{n}}{\left(\sum x^2 - \frac{(\sum x)^2}{n} \right) \left(\sum y^2 - \frac{(\sum y)^2}{n} \right)} \right)^2}{\left(\sum x^2 - \frac{(\sum x)^2}{n} \right) \left(\sum y^2 - \frac{(\sum y)^2}{n} \right)} \quad (3.26)$$

Burada x hesaplanan aktivite değerleri ve y ise deneysel aktivite değerleridir.

$$se = \sqrt{\frac{1-r^2}{n-2}} \quad (3.27)$$

Burada r^2 regresyon katsayısı, n ise veri sayısıdır.

$$q^2 = 1 - \frac{\sum_{n=1}^N |A_n^{den} - A_n^{hes}|^2}{\sum_{n=1}^N |A_n^{den} - \bar{A}^{den}|^2} \equiv 1 - \frac{PRESS}{SSY} \quad (3.28)$$

Çapraz doğrulama tüm veriyi kullanmayı sağlayan bir metottur. İncelenen bileşik serisi rasgele 2 kısıma ayrılır. İlk veri setiyle model kurulur. Kurulan modelin doğruluğu birini sırasıyla dışarıda bırakarak çapraz doğrulama (leave one out croos validation) metoduyla test edilir. Model oluşturmada kullanılan veri setinde her bir bileşik sırasıyla bir defa dışarıda bırakılarak aktivite hesaplamaları yapılır.

QSAR modellerinin önemli karakteristiklerinden biri de tahmin gücüdür. Model geçerliliği dahili (internal) ve harici (external) doğrulama (validation) teknikleri ile test edilmiştir [81].

Dahili doğrulamada (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.

$$PRESS_N = \sum_{n=1}^N |A_n^{den} - A_n^{hes}|^2 \quad (3.30)$$

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 çapraz doğrulama katsayısı karesi (q^2) eşitlik 3.28 ile hesaplanmaktadır.

Harici Doğrulama (External Validation) metodunda ise 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 [82,83].

$$q_{ext1}^2 = 1 - \frac{\sum_{n=1}^N |A_{n_{test}}^{den} - A_{n_{test}}^{hes}|^2}{\sum_{n=1}^N |A_{n_{test}}^{den} - \bar{A}_{eğitim}^{den}|^2} \quad (3.31)$$

$$q_{ext2}^2 = 1 - \frac{\sum_{n=1}^N |A_{n_{test}}^{den} - A_{n_{test}}^{hes}|^2}{\sum_{n=1}^N |A_{n_{test}}^{den} - \bar{A}_{n_{test}}^{den}|^2} \quad (3.32)$$

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

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

Genetik algoritma optimizasyon tekniği kullanılarak belirlenen en iyi parametre setinde hangi parametrenin aktivite üzerinde ne kadar etkili olduğunu yani parametrenin fiziksel katkısını anlamak için E istatistik yapılmıştır. Böylece eşitlik 3.33 kullanılarak hangi parametrenin daha önemli olduğu ve hangi parametrenin de ihmal edilebileceği belirlenmiştir.

$$E = \frac{PRESS_N}{PRESS_{N-1}} \quad (3.33)$$

Modelin tahmin gücünü doğrulamak için Chirico ve Gramatica'nın kullandığı formülleri kullanarak da hesap yapılmıştır [84].

$$q_{ext3}^2 = 1 - \frac{\left[\sum_{n=1}^{N_{test}} |A_{n_{test}}^{den} - A_{n_{test}}^{hes}|^2 \right] / N_{test}}{\left[\sum_{n=1}^{N_{eğitim}} |A_{n_{eğitim}}^{den} - \bar{A}_{eğitim}^{den}|^2 \right] / N_{eğitim}} \quad (3.34)$$

$$CCC = \hat{P}_{eğitim} = \frac{2 \sum_{i=1}^{n_{eğitim}} (A_i^{hes} - \bar{A}^{hes})(A_i^{den} - \bar{A}^{den})}{\sum_{i=1}^{n_{eğitim}} (A_i^{hes} - \bar{A}^{hes})^2 + \sum_{i=1}^{n_{eğitim}} (A_i^{den} - \bar{A}^{den})^2 + n_{eğitim} (\bar{A}^{hes} - \bar{A}^{den})^2} \quad (3.35)$$

$$CCC = \hat{P}_{test} = \frac{2 \sum_{i=1}^{n_{test}} (A_i^{hes} - \bar{A}^{hes})(A_i^{den} - \bar{A}^{den})}{\sum_{i=1}^{n_{test}} (A_i^{hes} - \bar{A}^{hes})^2 + \sum_{i=1}^{n_{test}} (A_i^{den} - \bar{A}^{den})^2 + n_{test} (\bar{A}^{hes} - \bar{A}^{den})^2} \quad (3.36)$$

$$CCC = \hat{P}_{hepsi} = \frac{2 \sum_{i=1}^{n_{hepsi}} (A_i^{hes} - \bar{A}^{hes})(A_i^{den} - \bar{A}^{den})}{\sum_{i=1}^{n_{hepsi}} (A_i^{hes} - \bar{A}^{hes})^2 + \sum_{i=1}^{n_{hepsi}} (A_i^{den} - \bar{A}^{den})^2 + n_{hepsi} (\bar{A}^{hes} - \bar{A}^{den})^2} \quad (3.37)$$

Yukarıdaki formüllerde,

CCC : Concorde Correlation Coefficient (Uyum Korelasyon Katsayısı)

n_{egitim} : Eğitim bileşikleri sayısı

n_{test} : Test bileşikleri sayısı

n_{hepsi} : Bütün bileşikleri sayısı

A_i^{den} : Deneysel aktivite değerlerini,

A_i^{hes} : Hesaplanan aktivite değerlerini

$\bar{A}^{hes}$: Hesaplanan aktivite değerleri ortalaması,

$\bar{A}^{den}$: Deneysel aktivite değerleri ortalamasını göstermektedir.

Uyum korelasyon katsayısı, kavramsal ve basit istatistiksel bir parametre olup yeni kimyasalların aktivitelerini tahmin edildiği QSAR metodunun harici doğrulanmasında bir kriter olarak değerlendirilir. Uyum korelasyon katsayısının performansı eşik değerler kullanılarak literatürde diğer harici doğrulama ölçümleri ile kriter karşılaştırılır. Yapılan bu bana benzer çalışmalarda uyum korelasyon katsayısının bir modelin harici olarak tahmin edici olarak en sınırlayıcı kriter olduğu bulunmuştur. Ayrıca diğer doğrulama kriterleriyle karşılaştırıldığında uyum korelasyon katsayısının iyi bir kararlılığa sahip olduğu ve beklenildiği gibi data dağılımı arttıkça uyum korelasyon katsayısının azaldığı gösterilmiştir. Sonuç olarak uyum korelasyon katsayısı bir modelin tahmin edici olarak

kabul edilip edilmemesinde son kararın verilmesinde yardımcı olur. Chirico ve arkadaşları bir QSAR metodunun gerçek tahmin ediciliğinin değerlendirilmesinde uyum korelasyon katsayısının tek doğrulama ölçütü olduğunu ileri sürmüşlerdir [84].

BÖLÜM 4

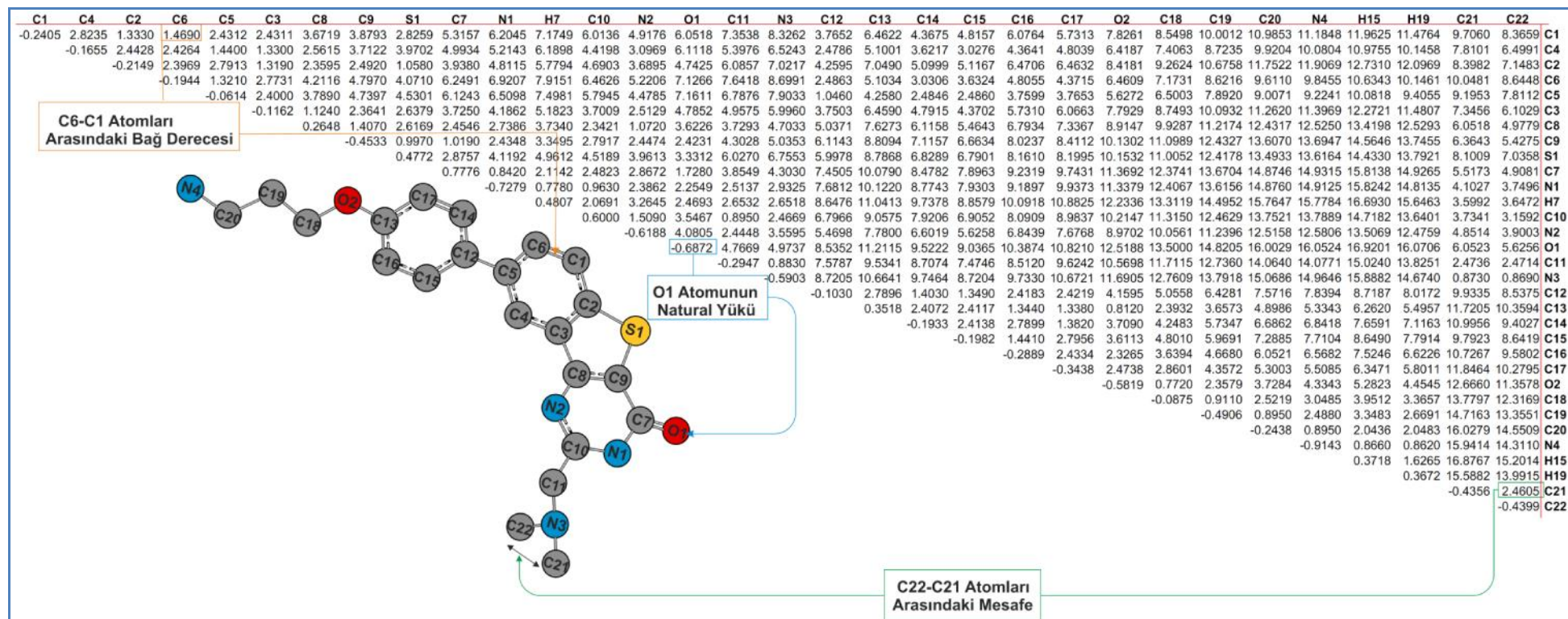
BULGULAR

4.1. ECMC Matrislerinin Hazırlanması ve Farmakofor Bulunması

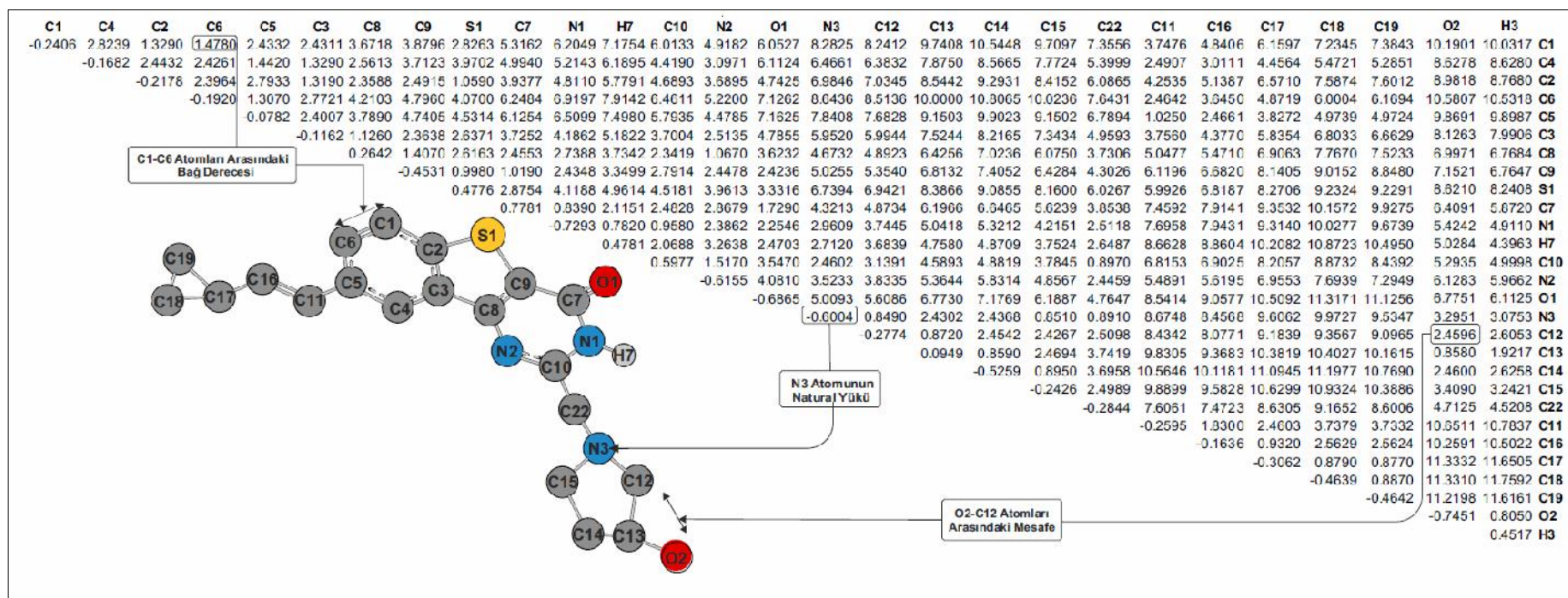
Bu çalışmada *.txt dosyalarındaki bilgiler içerisinde atomların yükleri, kartezyen koordinatları, bağ dereceleri ve atomlar arası mesafeler ile moleküle ait enerjiler kullanılarak 3. bölümde açıklandığı gibi ECMC oluşturulmuştur [24].

Bu çalışmada, matris hesaplamalarında txt dosyalarından natural yükleri kullanılarak matris oluşturulmuştur ve farmakoforu belirleme işlemi bütün matrisler ayrı ayrı dikkate alınarak gerçekleştirilmiştir. Şekil 4.1 ve 4.2’de benzotiyenopirimidin serisinde 61. ve 69. bileşiğe ait konformerler için natural yükleri kullanılarak oluşturulan farklı ECMC’ler gösterilmiştir.

Konformerler değiştikçe matris değişmektedir. Karbona bağlı hidrojen atomları reseptörle etkileşimde çok fazla etkiye sahip olmadığı için matris de dikkate alınmamıştır [24,54,56].



Şekil 4.1. Benzotienopirimidin serisinde 61. bileşiğinin 1. konformerine ait ECMC



Şekil 4.2. Benzotiyenopirimidin serisinde 69. bileşiğinin 1. konformerine ait ECMC

Şekil 4.1-4.2’de benzotiyenopirimidin serisindeki 61 ve 69 nolu bileşiklere ait 1. konformerleri için matrisler verilmiştir. Şekil 4.1’de 61. bileşikteki O1 atomuna ait natural yük, C6-C1 atomları arasındaki bağ derecesi, C22-C21 atomları arasındaki mesafe şekil üzerinde işaretlenerek gösterilmiştir. O1 atomunun natural yükü -0,6872 değerini göstermektedir. C6-C1 atomları arasındaki bağ derecesi 1,4690 değerini gösterirken, C22-C21 atomları arasındaki mesafe ise 2,4605 değerini göstermektedir. Şekil 4.2’de ise 69. Bileşikteki N3 atomuna ait natural yük, C1-C6 atomları arasındaki bağ derecesi, O2-C12 atomları arasındaki mesafe şekil üzerinde işaretlenerek gösterilmiştir. N3 atomunun natural yükü -0,6004 değerini göstermektedir. C1-C6 atomları arasındaki bağ derecesi 1,4780 değerini gösterirken, O2-C12 atomları arasındaki mesafe ise 2,4596 değerini göstermektedir. Konformerlerdeki atomların farklı yönelmeleri 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. Çünkü reseptörün aktiflenmiş durumda yapısının bilinmesi neredeyse imkânsızdır. Her bir konformerin katkısı aktivite formülünde görüldüğü gibi boltzman değerleri ile çarpılıp toplanmasıyla elde edilmektedir.

Düşük enerjili olan konformerin aktiviteye katkısı dolayısıyla resaptörle etkileşme ihtimali yüksek enerjili konformerden daha fazladır. 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 sonucu elde edilen kuantum kimyasal özelliklerin ve bilgisayarda elde edilen output dosyalarından matrislerin hazırlanması için EMRE V2.0 veya ECSP programı kullanılmıştır. Bu programlar Spartan 02, Spartan 04, Spartan 06 ve Spartan 08 programı ile elde edilen oldukça karışık çıktı dosyalarını düzenli birer matris dosyası haline dönüştürür ve bu program kullanılarak benzotiyenopirimidin serisine ait 96 bileşiğin 856 adet konformerinin hepsi için matrisler oluşturulmuştur. Şekil 4.1-4.2’de görülen matrisler tek konformer için üç boyutlu olarak hazırlanmıştır. Bileşik serilerine ait tüm konformerler için Şekil 3.8’de 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 V2.0 programı işimizi oldukça kolaylaştırmıştır. Hazırlanan ECMC matrisleri kullanılarak atomik yükler ve atomlar arası mesafelerin karşılaştırılması ile ECSP (Elektron Konformasyon Alt Matris Programı) programı ile farmakofor 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 içerisinde analiz yapılarak en uygun olan farmakofor grup belirlenir.

ECSP programı kullanılarak benzotienopirimidin 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 gruplar tekrar analiz edilerek en uygun grup bulunmaya çalışılır ve bulunan bu alt matris farmakofor grubu olarak tanımlanır.

İncelediğimiz benzotienopirimidin serisinde 96 adet bileşik ve bunlara ait toplam 856 konformer bulunmaktadır. Farmakofor grubunun tayininde ilk önce aktivite değeri en yüksek olan Pim-1 serisi için 61 nolu bileşik, Pim-2 için ise 69 nolu bileşik referans alınmıştır. Serideki bütün bileşikler referans bileşiğe göre belli tolerans değerlerinde karşılaştırılarak çıkan düşük tolerans değerine sahip atom grubu farmakofor atom grubu olarak belirlenmiştir. Pim-1 için farmakofor atomları C1, C5, S1, C7, N1, H7, N2, O1, C11 olarak bulunmuştur. Pim-2 için ise farmakofor atomları C1, C5, S1, C7, N1, H7, N2, O1, C22 olarak bulunmuştur. Tablo 4.1’de Pim-1 için bulunan sonuçlar verilmiştir.

Tablo 4.1. a) 61. bileşiğe (referans) ait alt matris (ECSA)

b) En düşük enerjili yüksek aktiviteli bileşikler için tolerans değerleri

c) En düşük enerjili düşük aktiviteli bileşikler için tolerans değerleri

d) 69 bileşiğin 672 konformeri için tolerans değerleri

a) 61. bileşiğe (referans) ait alt matris (ECSA)

C1	C5	S1	C7	N1	H7	N2	O1	C11
-0.2134	+2.4312	+2.8259	+5.3157	+6.2045	+7.1749	+4.9176	+6.0518	+7.3538
	-0.0496	+4.5301	+6.1243	+6.5098	+7.4981	+4.4785	+7.1611	+6.7876
		+0.4928	+2.8757	+4.1192	+4.9612	+3.9613	+3.3312	+6.0270
			+0.8792	+0.8420	+2.1142	+2.8672	+1.7280	+3.8549
				-0.9623	+0.7780	+2.3862	+2.2549	+2.5137
					+0.4320	+3.2645	+2.4693	+2.6532
						-0.8090	+4.0805	+2.4448
							-0.6796	+4.7669
								-0.2300

b) En düşük enerjili yüksek aktiviteli bileşikler için tolerans değerleri

C1	C5	S1	C7	N1	H7	N2	O1	C11
±0.0126	±0.0397	±0.0041	±0.0027	±0.0062	±0.0051	±0.0103	±0.0012	±0.0100
	±0.4091	±0.0560	±0.0421	±0.0332	±0.0320	±0.0211	±0.0469	±0.0178
		±0.0176	±0.0011	±0.0013	±0.0014	±0.0008	±0.0009	±0.0043
			±0.0025	±0.0060	±0.0014	±0.0011	±0.0100	±0.0099
				±0.0018	±0.0040	±0.0014	±0.0013	±0.0106
					±0.0021	±0.0016	±0.0039	±0.0191
						±0.0067	±0.0017	±0.0046
							±0.0081	±0.0112
								±0.0489

m_max (maksimum mesafe): ±0.0560 h24a

y_max (maksimum yük) : ±0.4091 h58a

c) En düşük enerjili düşük aktiviteli bileşikler için tolerans değerleri

C1	C5	S1	C7	N1	H7	N2	O1	C11
±0.0119	±0.0145	±0.0058	±0.0046	±0.0046	±0.0361	±0.0125	±0.0067	±0.0224
	±0.3299	±0.0224	±0.0203	±0.0173	±0.0158	±0.0119	±0.0194	±0.0529
		±0.0284	±0.0064	±0.0046	±0.0407	±0.0024	±0.0048	±0.0101
			±0.0335	±0.0320	±0.0429	±0.0088	±0.1290	±0.0293
				±0.0490	±0.1500	±0.0033	±0.0034	±0.0389
					±0.0570	±0.0201	±0.0517	±0.0761
						±0.0611	±0.0052	±0.0294
							±0.0266	±0.0389
								±0.3808

m_max: ±0.1500 h12a

y_max: ±0.3808 h1a

d) 69 bileşiğin 672 konformeri için tolerans değerleri

C1	C5	S1	C7	N1	H7	N2	O1	C11
±0.0132	±0.0397	±0.0058	±0.0046	±0.0085	±0.0361	±0.0125	±0.0059	±0.0403
	±0.4100	±0.0561	±0.0421	±0.0374	±0.0352	±0.0216	±0.0469	±0.0533
		±0.0317	±0.0064	±0.0089	±0.0407	±0.0024	±0.0059	±0.0167
			±0.0355	±0.0370	±0.0429	±0.0088	±0.1290	±0.0304
				±0.0490	±0.1500	±0.0078	±0.0034	±0.0395
					±0.0570	±0.0201	±0.0517	±0.0771

±0.0611 ±0.0053 ±0.0407
 ±0.0266 ±0.0400
 ±0.3808

m_max:±0.1500 h12a
 y_max:±0.4100 h58e

Tablo 4.2’de Pim-2 için sonuçlar verilmiştir.

Tablo 4.2. a) 69. bileşiğe (referans) ait alt matris (ECSA)

b) En düşük enerjili yüksek aktiviteli bileşikler için tolerans değerleri

c) En düşük enerjili düşük aktiviteli bileşikler için tolerans değerleri

d) 69 bileşiğin 672 konformerleri için tolerans değerleri

a) 69. bileşiğe (referans) ait alt matris (ECSA)

C1	C5	S1	C7	N1	H7	N2	O1	C22
-0.2406	+2.4332	+2.8263	+5.3162	+6.2049	+7.1754	+4.9182	+6.0527	+7.3556
	-0.0782	+4.5314	+6.1254	+6.5099	+7.4980	+4.4785	+7.1625	+6.7894
		+0.4776	+2.8754	+4.1188	+4.9614	+3.9613	+3.3316	+6.0267
			+0.7781	+0.8390	+2.1151	+2.8679	+1.7290	+3.8538
				-0.7293	+0.7820	+2.3862	+2.2546	+2.5118
					+0.4781	+3.2638	+2.4703	+2.6487
						-0.6155	+4.0810	+2.4459
							-0.6865	+4.7647
								-0.2844

b) En düşük enerjili yüksek aktiviteli bileşikler için tolerans değerleri

C1	C5	S1	C7	N1	H7	N2	O1	C22
±0.0222	±0.0417	±0.0044	±0.0032	±0.0050	±0.0046	±0.0096	±0.0045	±0.0082
	±0.2941	±0.0573	±0.0430	±0.0333	±0.0315	±0.0211	±0.0478	±0.0182
		±0.0235	±0.0014	±0.0042	±0.0048	±0.0010	±0.0022	±0.0043
			±0.0194	±0.0330	±0.0077	±0.0095	±0.1300	±0.0070
				±0.0031	±0.0430	±0.0028	±0.0037	±0.0213
					±0.0135	±0.0082	±0.0077	±0.0393
						±0.0080	±0.0046	±0.0073
							±0.0286	±0.0159
								±0.0133

m_max:±0.1300 h21a
 y_max:±0.2941 h58a

c) En düşük enerjili düşük aktiviteli bileşikler için tolerans değerleri

C1	C5	S1	C7	N1	H7	N2	O1	C22
±0.0116	±0.0415	±0.0062	±0.0051	±0.0058	±0.0356	±0.0119	±0.0076	±0.0242
	±0.2606	±0.0573	±0.0432	±0.0331	±0.0319	±0.0210	±0.0483	±0.0547
		±0.0164	±0.0061	±0.0035	±0.0405	±0.0024	±0.0052	±0.0098
			±0.0064	±0.0350	±0.0420	±0.0059	±0.0150	±0.0304
				±0.0251	±0.1540	±0.0033	±0.0016	±0.0408
					±0.0232	±0.0208	±0.0507	±0.0806
						±0.0355	±0.0057	±0.0305
							±0.0098	±0.0411
								±0.4535

m_max:±0.1540 h12a
 y_max:±0.4535 h1a

d) 69 bileşiğin 672 konformeri için tolerans değerleri

C1	C5	S1	C7	N1	H7	N2	O1	C22
±0.0237	±0.0417	±0.0062	±0.0051	±0.0089	±0.0356	±0.0119	±0.0076	±0.0385
	±0.2952	±0.0574	±0.0432	±0.0375	±0.0351	±0.0216	±0.0483	±0.0551
		±0.0262	±0.0061	±0.0085	±0.0405	±0.0024	±0.0059	±0.0170
			±0.0213	±0.0400	±0.0420	±0.0095	±0.1300	±0.0315
				±0.0251	±0.1540	±0.0078	±0.0037	±0.0414
					±0.0232	±0.0208	±0.0507	±0.0816
						±0.0355	±0.0058	±0.0396
							±0.0286	±0.0422
								±0.4535
m_max:±0.1540 h12a								
y_max:±0.4535 h1a								

Literatürdeki çalışmalara göre inaktif moleküllerde tolerans değerin daha büyük olması gerektiği verilmektedir [78]. Bulduğumuz farmakofor sonuçlarında mesafe ve yükün maksimum tolerans değerleri, aktivitesi büyük olan bileşiklerde daha düşük, düşük aktivitelilerde ise daha yüksek olduğu yukarıdaki matrisde görülmektedir. Bunun için hesaplanan P_α ve α_a değerleri sonra verilmiştir. En düşük enerjili konformerde matris elemanı bulunmaması durumunda daha yüksek enerjili konformerlerde tarama yapılabilir [77].

İncelediğimiz benzotiyenoprimidin serisinde 69 adet bileşik ve bunlara ait toplam 672 conformer bulunmaktadır. Yukarıda bahsedildiği gibi bu bileşikler 35 aktif ve 34 düşük aktiviteli olarak sınıflandırılmıştır. Aktif bileşiklerde bulunan bu farmakofor grubunun olma ihtimali yani P_α ve α_a benzotiyenoprimidin serisindeki Pim-1 ve Pim-2 için sırasıyla $P_\alpha=0,8049$ ve $\alpha_a=0,7144$ $P_\alpha=0,7619$ ve $\alpha_a=0,6290$ değerlerine sahiptir.

4.2. Aktivite Hesaplamasında Kullanılan Parametreler

Farmakofor grubu bulunduktan sonra seçilen bileşik serisinde aktivite hesaplamalarına geçilmiştir. Bunun için aktiviteye etki ettiği düşünülen 1594 adet parametre benzotiyenoprimidin serisindeki 96 adet bileşik için EMRE V2.0 programında aktivite formülünde (eşitlik 3.5) kullanılabilmesi için uygun formatta hazırlanmıştır. Bileşik ve konformer sayısı arttıkça üstel olarak parametre sayısı da artacağından, benzotiyenoprimidin serisinde bütün konformerler için 1594 adet parametre hazırlanmıştır. Benzotiyenoprimidin serisi için hesaplamalarda kullanılan parametreler ve 61. bileşiğinin en düşük enerjili konformerine ait sayısal değerler Tablo 4.3'te verilmiştir. Bütün bileşiklerin bütün konformerlerinin parametrelerini vermek çok yer

kaplayacağından sadece Pim-1 için referans bileşik olan 61 nolu bileşiğin en düşük enerjili konformer için parametreler tablo 4.3'te verilmiştir.

Tablo 4.3. Benzotiyenopirimidin serisi için hesaplamalarda kullanılan parametreler

NO	PARAMETRELER	NO	PARAMETRELER
	Mesafe (Å)		
1	S1 O1	14	C9 C12
2	N1 N2	15	C9 N4
3	N1 S1	16	S1 C11
4	S1 N2	17	S1 C22
5	C7 N2	18	O1 C11
6	S1 C12	19	O1 C22
7	S1 N4	20	N1 C11
8	O1 C12	21	N1 C22
9	O1 N4	22	N2 C11
10	N1 C12	23	N2 C22
11	N1 N4	24	C9 C11
12	N2 C12	25	C9 C22
13	N2 N4	26	N4 C21
	Mesafe + Van der Waals (Å)		
27	S1(S1) O1(O1)	40	C9(C9) C12(C12)
28	N1(H7) N2(N2)	41	C9(C9) N4(H19)
29	N1(H7) S1(S1)	42	S1(S1) C11(C11)
30	S1(S1) N2(N2)	43	S1(S1) C22(H5)
31	C7(O1) N2(N2)	44	O1(O1) C11(C11)
32	S1(S1) C12(C12)	45	O1(O1) C22(H5)
33	S1(S1) N4(H19)	46	N1(H7) C11(C11)
34	O1(O1) C12(C12)	47	N1(H7) C22(H5)
35	O1(O1) N4(H19)	48	N2(N2) C11(C11)
36	N1(H7) C12(C12)	49	N2(N2) C22(H5)
37	N1(H7) N4(H19)	50	C9(C9) C11(C11)
38	N2(N2) C12(C12)	51	C9(C9) C22(H5)
39	N2(N2) N4(H19)	52	N4(N4) C21(H21)
	C1 C5 C8 Düzlemine Yüzey Dik Uzaklığı (Å)		
53	S1	59	C10
54	C9	60	H7
55	C2	61	N4
56	O1	62	C12
57	N1	63	C11
58	N2	64	C22
	C1 C5 C8 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)		
65	S1(S1)	70	N4(H19)
66	O1(O1)	71	C12(C14)
67	N1(H7)	72	C11(N3)
68	N2(N2)	73	C22(H24)
69	H7(H7)		
	C8 S1 C7 Düzlemine Yüzey Dik Uzaklığı (Å)		
74	C1	80	C10
75	C9	81	H7
76	C2	82	N4
77	O1	83	C5
78	N1	84	C11

79	N2	85	C22
C8 S1 C7 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)			
86	C1(H1)	90	H7(H7)
87	O1(O1)	91	N4(H19)
88	N1(H7)	92	C5(C12)
89	N2(N2)	93	C11(N3)
		94	C22(H24)

S1 C7 N1 Düzlemine Yüzey Dik Uzaklığı (Å)			
95	C8	101	C10
96	C9	102	H7
97	C2	103	N4
98	O1	104	C5
99	C3	105	C11
100	N2	106	C22
S1 C7 N1 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)			
107	C8(C8)	112	N4(H19)
108	O1(O1)	113	C5(C12)
109	C3(C3)	114	C11(N3)
110	N2(N2)	115	C22(H24)
111	H7(H7)		
S1 C7 O1 Düzlemine Yüzey Dik Uzaklığı (Å)			
116	C8	122	C10
117	C9	123	H7
118	C2	124	N4
119	N1	125	C5
120	C3	126	C11
121	N2	127	C22
S1 C7 O1 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)			
128	C8(C8)	133	N4(H19)
129	N1(H7)	134	C5(C12)
130	C3(C3)	135	C11(N3)
131	N2(N2)	136	C22(H24)
132	H7(H7)		
C1 C7 O1 Düzlemine Yüzey Dik Uzaklığı (Å)			
137	S1	143	C10
138	C9	144	H7
139	C2	145	N4
140	N1	146	C5
141	C3	147	C11
142	N2	148	C22
C1 C7 O1 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)			
149	S1(S1)	154	N4(H19)
150	N1(H7)	155	C5(C12)
151	C3(C3)	156	C11(N3)

152	N2(N2)	157	C22(H24)
153	H7(H7)		
	O1 C7 N1 Düzlemine Yüzey Dik Uzaklığı (Å)		
158	S1	164	C10
159	C9	165	H7
160	C2	166	N4
161	C8	167	C5
162	C3	168	C11
163	N2	169	C22
	O1 C7 N1 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)		
170	S1(S1)	175	N4(H19)
171	C8(C8)	176	C5(C12)
172	C3(C3)	177	C11(N3)
173	N2(N2)	178	C22(H24)
174	H7(H7)		
	C7 N1 H7 Düzlemine Yüzey Dik Uzaklığı (Å)		
179	S1	185	C10
180	C9	186	C1
181	C2	187	N4
182	O1	188	C12
183	C3	189	C11
184	N2	190	C22
	C7 N1 H7 Düzlemine Yüzey Dik Uzaklığı + Van der Waals (Å)		
191	S1(S1)	196	N4(H19)
192	O1(O1)	197	C12(C14)
193	C3(C3)	198	C11(N3)
194	N2(N2)	199	C22(H24)
195	C1(H1)		
	C7 N1 N2 Düzlemine Yüzey Dik Uzaklığı (Å)		
200	S1	206	C10
201	C9	207	H7
202	C2	208	N4
203	O1	209	C12
204	C3	210	C11
205	C1	211	C22
	C7 N1 N2 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)		
212	S1(S1)	217	N4(H19)
213	O1(O1)	218	C12(C14)
214	C3(C3)	219	C11(N3)
215	C1(H1)	220	C22(H24)
216	H7(H7)		
	N2 N1 H7 Düzlemine Yüzey Dik Uzaklığı (Å)		
221	S1	227	C10
222	C9	228	C7

223	C2	229	N4
224	O1	230	C12
225	C3	231	C11
226	C1	232	C22
N2 N1 H7 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
233	S1(S1)	238	N4(H19)
234	O1(O1)	239	C12(C14)
235	C3(C3)	240	C11(N3)
236	C1(H1)	241	C22(H24)
237	C7(O1)		
N2 N1 C11 Düzlemine Yüzey Dik Uzaklığı (Å)			
242	S1	248	C10
243	C9	249	C7
244	C2	250	N4
245	O1	251	C12
246	C3	252	H7
247	C1	253	C22
N2 N1 C11 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
254	S1(S1)	259	N4(H19)
255	O1(O1)	260	C12(C14)
256	C3(C3)	261	H7(H7)
257	C1(H1)	262	C22(H24)
258	C7(O1)		
S1 N2 O1 Düzlemine Yüzey Dik Uzaklığı (Å)			
263	C7	269	C10
264	C9	270	H7
265	C2	271	N4
266	N1	272	C5
267	C3	273	C11
268	C1	274	C22
S1 N2 O1 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
275	C7(O1)	280	N4(H19)
276	N1(H7)	281	C5(C12)
277	C3(C3)	282	C11(N3)
278	C1(H1)	283	C22(H24)
279	H7(H7)		
S1 N2 N1 Düzlemine Yüzey Dik Uzaklığı (Å)			
284	C8	290	C10
285	C9	291	H7
286	C2	292	N4
287	C7	293	C5
288	C3	294	C11
289	O1	295	C22
Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
296	C8(C8)	301	N4(H19)
297	C7(O1)	302	C5(C12)
298	C3(C3)	303	C11(N3)
299	O1(O1)	304	C22(H24)

300	H7(H7)		
	C1 C8 O1 Düzlemine Yüzey Dik Uzaklığı (Å)		
305	S1	311	C10
306	C9	312	H7
307	C2	313	N4
308	C7	314	C5
309	N1	315	C11
310	N2	316	C22
	C1 C8 O1 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)		
317	S1(S1)	322	N4(H19)
318	C7(O1)	323	C5(C12)
319	N1(H7)	324	C11(N3)
320	N2(N2)	325	C22(H24)
321	H7(H7)		
	C1 C8 O1 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)		
317	S1(S1)	322	N4(H19)
318	C7(O1)	323	C5(C12)
319	N1(H7)	324	C11(N3)
320	N2(N2)	325	C22(H24)
321	H7(H7)		
	C1 C8 H7 Düzlemine Yüzey Dik Uzaklığı (Å)		
326	S1	332	C10
327	C9	333	O1
328	C2	334	N4
329	C7	335	C5
330	N1	336	C11
331	N2	337	C22
	C1 C8 H7 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)		
338	S1(S1)	343	N4(H19)
339	C7(O1)	344	C5(C12)
340	N1(H7)	345	C11(N3)
341	N2(N2)	346	C22(H24)
342	O1(O1)		
	C1 C8 C11 Düzlemine Yüzey Dik Uzaklığı (Å)		
347	S1	353	C10
348	C9	354	H7
349	C2	355	N4
350	O1	356	C12
351	N1	357	C7
352	N2	358	C22
	C1 C8 C11 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)		
359	S1(S1)	364	N4(H19)
360	O1(O1)	365	C12(C14)
361	N1(H7)	366	C7(O1)
362	N2(N2)	367	C22(H24)

363	H7(H7)		
C1 C5 C4 Düzlemine Yüzey Dik Uzaklığı (Å)			
368	S1	374	C10
369	C9	375	H7
370	C2	376	N4
371	O1	377	C12
372	N1	378	C7
373	N2	379	C22
C1 C5 C4 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
380	S1(S1)	385	N4(H19)
381	O1(O1)	386	C12(C14)
382	N1(H7)	387	C7(O1)
383	N2(N2)	388	C22(H24)
384	H7(H7)		

C4 C8 S1 Düzlemine Yüzey Dik Uzaklığı (Å)			
389	C3	395	C10
390	C9	396	H7
391	C2	397	N4
392	O1	398	C12
393	N1	399	C7
394	N2	400	C22
C4 C8 S1 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
401	C3(C3)	406	N4(H19)
402	O1(O1)	407	C12(C14)
403	N1(H7)	408	C7(O1)
404	N2(N2)	409	C22(H24)
405	H7(H7)		
C4 O1 C7 Düzlemine Yüzey Dik Uzaklığı (Å)			
410	S1	416	C10
411	C9	417	H7
412	C2	418	N4
413	C8	419	C12
414	N1	420	C3
415	N2	421	C22
C4 O1 C7 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
422	S1(S1)	427	N4(H19)
423	C8(C8)	428	C12(C14)
424	N1(H7)	429	C3(C3)
425	N2(N2)	430	C22(H24)
426	H7(H7)		
C4 C8 N2 Düzlemine Yüzey Dik Uzaklığı (Å)			
431	S1	437	C10
432	C9	438	H7

433	C2	439	N4
434	O1	440	C12
435	N1	441	C11
436	C7	442	C22
C4 C8 N2 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
443	S1(S1)	448	N4(H19)
444	O1(O1)	449	C12(C14)
445	N1(H7)	450	C11(N3)
446	C7(O1)	451	C22(H24)
447	H7(H7)		
C4 O1 H7 Düzlemine Yüzey Dik Uzaklığı (Å)			
452	S1	458	C10
453	C9	459	N2
454	C2	460	N4
455	C8	461	C12
456	N1	462	C11
457	C7	463	C22
C4 O1 H7 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
464	S1(S1)	469	N4(H19)
465	C8(C8)	470	C12(C14)
466	N1(H7)	471	C11(N3)
467	C7(O1)	472	C22(H24)
468	N2(N2)		
C4 N1 N2 Düzlemine Yüzey Dik Uzaklığı (Å)			
473	S1	479	C10
474	C9	480	H7
475	C2	481	N4
476	O1	482	C12
477	C3	483	C11
478	C1	484	C22
C4 N1 N2 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
485	S1(S1)	490	N4(H19)
486	O1(O1)	491	C12(C14)
487	C3(C3)	492	C11(N3)
488	C1(H1)	493	C22(H24)
489	H7(H7)		
H7 N1 C11 Düzlemine Yüzey Dik Uzaklığı (Å)			
494	S1	500	C10
495	C9	501	C3
496	C2	502	N4
497	O1	503	C12
498	C7	504	C8
499	N2	505	C22
H7 N1 C11 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			

506	S1(S1)	511	N4(H19)
507	O1(O1)	512	C12(C14)
508	C7(O1)	513	C8(C8)
509	N2(N2)	514	C22(H24)
510	C3(C3)		
C11 C7 O1 Düzlemine Yüzey Dik Uzaklığı (Å)			
515	S1	521	C10
516	C9	522	H7
517	C2	523	N4
518	C1	524	C12
519	N1	525	C8
520	N2	526	C22
C11 C7 O1 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
527	S1(S1)	532	N4(H19)
528	C1(H1)	533	C12(C14)
529	N1(H7)	534	C8(C8)
530	N2(N2)	535	C22(H24)
531	H7(H7)		
C11 O1 C8 Düzlemine Yüzey Dik Uzaklığı (Å)			
536	S1	542	C10
537	C9	543	H7
538	C2	544	N4
539	C7	545	C12
540	N1	546	C3
541	N2	547	C22
C11 O1 C8 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
548	S1(S1)	553	N4(H19)
549	C7(O1)	554	C12(C14)
550	N1(H7)	555	C3(C3)
551	N2(N2)	556	C22(H24)
552	H7(H7)		
C11 C1 C4 Düzlemine Yüzey Dik Uzaklığı (Å)			
557	S1	563	C10
558	C9	564	H7
559	C2	565	N4
560	C7	566	C12
561	N1	567	C8
562	N2	568	C22
C11 C1 C4 Düzlemine Yüzey Dik Uzaklığı (Å) + Van der Waals (Å)			
569	S1(S1)	574	N4(H19)
570	C7(O1)	575	C12(C14)
571	N1(H7)	576	C8(C8)
572	N2(N2)	577	C22(H24)
573	H7(H7)		

	Açı (Radyan)		
578	N2 C10 N1	592	N4 C7 C22
579	C11 C10 N1	593	N4 N1 C22
580	C11 C10 N2	594	N4 N2 C22
581	C10 C11 N3	595	C12 C1 N3
582	C12 C5 C4	596	C12 S1 N3
583	C12 C5 C6	597	C12 O1 N3
584	N4 C5 C4	598	C12 N1 N3
585	N4 C5 C6	599	N4 C22 C1
586	C22 C11 O1	600	N4 C22 S1
587	C22 C11 S1	601	N4 C22 O1
588	N3 C11 O1	602	N4 C22 C7
589	N3 C11 S1	603	N4 C22 N1
590	N4 S1 C22	604	N4 C22 N2
591	N4 O1 C22	605	N4 C22 H7
	*Yüzey Doğru Açısı (Radyan)		
606	C1 C5 C8 (S1 C11)	...	...
...	Çeşitli Yüzey Doğru Açıları	...	...
...	...	...	...
...	...	...	...
...	...	1330	C11 C1 C4 (H15 C21)
	Mulliken Bond Order		
1331	C1 C2	1341	C9 C7
1332	C2 C3	1342	C7 O1
1333	C3 C4	1343	C7 N1
1334	C4 C5	1344	N1 H7
1335	C5 C6	1345	N1 C10
1336	C6 C1	1346	N2 C10
1337	S1 C2	1347	C10 C11
1338	S1 C9	1348	C11 N3
1339	C3 C8	1349	C5 C12
1340	C8 C9	1350	C12 C15
	Lowdin Bond Order		
1351	C1 C2	1361	C9 C7
1352	C2 C3	1362	C7 O1
1353	C3 C4	1363	C7 N1
1354	C4 C5	1364	N1 H7
1355	C5 C6	1365	N1 C10
1356	C6 C1	1366	N2 C10
1357	S1 C2	1367	C10 C11
1358	S1 C9	1368	C11 N3
1359	C3 C8	1369	C5 C12
1360	C8 C9	1370	C12 C15

	Mulliken Yüğü (e-)		
1371	S1	1377	C5
1372	N1	1378	C11
1373	O1	1379	C10
1374	C7	1380	C12
1375	N2	1381	C8
1376	C1	1382	C9
	Natural Yüğü (e-)		
1383	S1	1389	C5
1384	N1	1390	C11
1385	O1	1391	C10
1386	C7	1392	C12
1387	N2	1393	C8
1388	C1	1394	C9
	Elektrostatik Yüğü (e-)		
1395	S1	1401	C5
1396	N1	1402	C11
1397	O1	1403	C10
1398	C7	1404	C12
1399	N2	1405	C8
1400	C1	1406	C9
	Mulliken Valens (e-)		
1407	S1	1413	C5
1408	N1	1414	C11
1409	O1	1415	C10
1410	C7	1416	C12
1411	N2	1417	C8
1412	C1	1418	C9
	Lowdin Valens (e-)		
1419	S1	1425	C5
1420	N1	1426	C11
1421	O1	1427	C10
1422	C7	1428	C12
1423	N2	1429	C8
1424	C1	1430	C9
	Elektrofilik Atom Sınır Elektron Yoğunluğu		
1431	S1	1437	C5
1432	N1	1438	C11
1433	O1	1439	C10
1434	C7	1440	C12
1435	N2	1441	C8
1436	C1	1442	C9
	Elektrofilik Atom Sınır Elektron Yoğunluğu İndeksi		
1443	S1	1449	C5

1444	N1	1450	C11
1445	O1	1451	C10
1446	C7	1452	C12
1447	N2	1453	C8
1448	C1	1454	C9
Nükleofilik Atom Sınır Elektron Yoğunluğu			
1455	S1	1461	C5
1456	N1	1462	C11
1457	O1	1463	C10
1458	C7	1464	C12
1459	N2	1465	C8
1460	C1	1466	C9
Nükleofilik Atom Sınır Elektron Yoğunluğu İndeksi			
1467	S1	1473	C5
1468	N1	1474	C11
1469	O1	1475	C10
1470	C7	1476	C12
1471	N2	1477	C8
1472	C1	1478	C9
Fukui Atomik Nükleofilik Reaktivite İndeksi			
1479	S1	1485	C5
1480	N1	1486	C11
1481	O1	1487	C10
1482	C7	1488	C12
1483	N2	1489	C8
1484	C1	1490	C9
Fukui Atomik Elektrofilik Reaktivite İndeksi			
1491	S1	1497	C5
1492	N1	1498	C11
1493	O1	1499	C10
1494	C7	1500	C12
1495	N2	1501	C8
1496	C1	1502	C9
Elektrofilik Atom Sınır Elektron Yoğunluğu İndeksi (ev)			
1503	S1	1509	C5
1504	N1	1510	C11
1505	O1	1511	C10
1506	C7	1512	C12
1507	N2	1513	C8
1508	C1	1514	C9
Nükleofilik Atom Sınır Elektron Yoğunluğu İndeksi (ev)			
1515	S1	1521	C5
1516	N1	1522	C11
1517	O1	1523	C10

1518	C7	1524	C12
1519	N2	1525	C8
1520	C1	1526	C9
	Atomik Nükleofilik Reaktivite İndeksi (ev)		
1527	S1	1533	C5
1528	N1	1534	C11
1529	O1	1535	C10
1530	C7	1536	C12
1531	N2	1537	C8
1532	C1	1538	C9
	Atomik Elektrofilik Reaktivite İndeksi (ev)		
1539	S1	1545	C5
1540	N1	1546	C11
1541	O1	1547	C10
1542	C7	1548	C12
1543	N2	1549	C8
1544	C1	1550	C9
1551	E_{L-H} = Lumo-Homo Farkı	1562	E_{L-H} = Lumo-Homo Farkı
1552	Kararlılık İndeksi	1563	Kararlılık İndeksi
1553	Aktivasyon Enerji İndeksi	1564	Aktivasyon Enerji İndeksi
1554	Hardness	1565	Hardness
1555	Elektronegatiflik İndeksi	1566	Elektronegatiflik İndeksi
1556	Kimyasal Potansiyel	1567	Kimyasal Potansiyel
1557	Softness	1568	Softness
1558	Electrofilik İndeks	1569	Electrofilik İndeks
1559	Homo	1570	Homo
1560	Lumo	1571	Lumo
1561	Nükleofilik İndeks	1572	Nükleofilik İndeks
1573	relE	1576	HBD
1574	E	1577	HBA
1575	PSA	1578	POS
	Torsiyon Açısı (Radyan)		
1579	C11 O1 S1 C12	1587	C22 N2 S1 C12
1580	C11 O1 S1 N4	1588	C22 N2 S1 N4
1581	C11 N2 S1 C12	1589	C22 N1 S1 C12
1582	C11 N2 S1 N4	1590	C22 N1 S1 N4
1583	C11 C7 S1 C12	1591	C22 N2 O1 C12
1584	C11 C7 S1 N4	1592	C22 N2 O1 N4
1585	C22 O1 S1 C12	1593	C22 C7 C1 C12
1586	C22 O1 S1 N4	1594	C22 C7 C1 N4

Tabloda * ile işaretlenmiş olan yüzey doğru açısı değerleri 606 ile 1330 numaralı parametreleri kapsamaktadır. Çeşitli yüzeyler ile çeşitli iki atom arasında geçen yüzey doğru açısı değerlerini göstermektedir. Çok yer kapladığından bu aralık tabloda gösterilmemiştir.

Benzotiyenopirimidin bileşik serisinde 96 bileşiğin herbir konformeri için belirlenen değişkenler hazırlandıktan sonra Matlab 7.04 paket programında yazılan kod ile eşitlik 3.5'deki formüle göre aktivite hesaplamaları yapıldı. Bileşiklerin literatürde Pim-1 ve Pim-2 için inhibisyon değerleri (K_i) nanomolar cinsinden verilmiştir [20]. Bu değerler molar birimine dönüştürülerek eksi logaritması ($-\log K_i$) alınarak hesaplamalarda kullanılmıştır. Pim-1 ve Pim-2 inhibisyon değerleri biyolojik veya biyokimyasal fonksiyonların engellenmesinde bir bileşiğin etkinliğinin bir ölçümü olarak tarif edilmektedir.

4.3. Benzotiyenopirimidin Bileşik Serisinin Aktivite Hesaplamaları

Benzotiyenopirimidin bileşik serisi için hazırlanan 1594 parametre içerisinde aktivite üzerine en çok etkisi olan parametrelerin seçimi genetik algoritma optimizasyon tekniği kullanılarak yapılmıştır. En iyi parametreler ve bu parametrelere ait k_j değerleri hesaplandıktan sonra aktivite hesaplamaları eşitlik 3.5 kullanılarak yapılmıştır. Kullanılan bu denklem lineer olmadığı için bu denklemin çözümü lsqnonlin fonksiyonu ile gerçekleştirilmiştir. Benzotiyenopirimidin serisi için yapılan hesaplamalar ilerleyen sayfalarda 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 benzotiyenopirimidin serisinde oluşturulan modelin etkinliğini arttırmak ve modelin performansını ölçmek için 4 farklı bileşik setinde hesaplamalar yapılarak modelleme geliştirilmiştir. Bu hesaplamalara ilişkin sonuçlar aşağıda verilmektedir.

4.3.1. Benzotiyenopirimidin Bileşik Serisinin Eğitim ve Test Seti Olarak Hesaplanması

Geliştirilen bir modelde en önemli etken verilerin eğitim ve test seti şeklinde ikiye ayrılmasıdır. Nedeni genetik algortmada en iyi parametrelerin yani kappa değerlerinin ve regresyon katsayılarının belirlenebilmesi, geliştirilen optimizasyon modelinin

biyolojik aktivite hesaplamalarında deneye ihtiyaç duyulmadan aktivitesi bilinmeyen bileşiklerin aktivite hesaplamalarını tahmin edebilmesi ve ayrıca 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. 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 olarak sıfır olacak şekilde işleme başlanarak geniş bir uzay alanında tarama yapılır. Bu taramadaki amaç uygun parametre ve kappa değerleri için doğru çıktıları üretecek değerleri bulmaktır. Bunun için defalarca (ortalama 2000) iterasyon yapılır. Uygun kappa değerlerine, regresyon katsayısının maksimum ve standart hatanın minimum olduğunda ulaşılmış olur. Eğitim setinin uygun bir şekilde hazırlanmış olduğunu gösterir. Bu duruma eğitim setinin öğrenmesi denir. Setin öğrenip öğrenmediğini (performansını ölçme) yani aktivitesi bilinmeyen bileşiklerin (test setinin) aktivitesinin belirlenmesi için yapılan denemelere ise setin test edilmesi denir. Eğitim seti için bulunan moleküler parametreler ve kappa değerleri değiştirilmeden aynı aktivite formülü (eşitlik 3.5) 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. Elde edilen çıktıların doğruluk değerleri setin öğrenmesi hakkında bilgiler verir. Geçerliliği en iyi olan modeli belirleyebilmek, modelin performansını ölçmek ve sonuçların doğruluğunu test etmek için 4 ayrı hesaplama seti ile çalışılmıştır. Model 1’de bileşikler 50 eğitim 17 test olarak, model 2’de ise 45 eğitim 22 test olarak bölünerek hesap yapılmıştır. Model 3’te ise 67 bileşik eğitim olarak alınmış ve bölmeden hesap verilmiştir. Model 4’te ise her bir bileşiğin bir konformerini hesaba katılarak hesap verilmiştir. Pim-1 ve Pim-2 için bu hesaplar ayrı ayrı yapılmıştır.

Aşağıda Pim-1 ve Pim-2 için bu modellere ait bulunan sonuçlar verilmiştir.

Tablo.4.4. Pim-1 için deneysel aktivite değerleri ve hesaplanan aktivite değerleri

Bileşik No	Pim1- A_{deneysel}	Pim1- $A_{\text{hesaplanan}}$			
		Model 1	Model 2	Model 3	Model 4
1 ^{a,b,c}	6,4461	5,4898	6,8400	6,4464	5,4469
2	7,8539	7,5425	7,8464	7,9210	7,5501
3 ^b	7,9208	7,1862	6,5077	7,2842	7,6039
4 ^b	7,9586	7,9405	7,9224	7,8831	7,5947
5	6,4789	7,2636	7,0571	7,4606	7,4078
6 ^d		7,9245	8,0451	7,2147	7,5450
7	8,0458	7,5081	7,8528	7,6997	7,4815
8 ^d		10,5231	10,6505	9,7403	13,2444
9	7,3768	7,5688	7,1972	7,2994	7,4822
10 ^{a,c}	7,1487	7,1129	7,4749	7,2649	7,2819
11 ^b	7,7696	7,6658	6,3822	7,5572	7,6221
12 ^{a,c}	7,9586	8,3501	8,0299	8,1137	7,9917
13	7,3565	7,4386	7,5990	7,2598	7,7840
14	7,0362	6,9128	7,5079	7,3553	6,8558
15 ^{a,c}	9,0458	7,0102	7,9113	8,5088	6,8336
16	6,9706	7,3611	7,2389	6,8654	7,7807
17 ^{a,c}	7,0362	6,7869	7,1213	7,0405	7,3035
18	7,1805	7,1567	7,4172	7,6608	7,5063
19	7,4202	7,3012	7,2373	7,2177	7,6735
20	7,7447	7,5884	6,9489	7,4277	7,4486
21 ^{a,c}	7,3098	6,7430	7,2478	7,5821	7,9149
22 ^b	7,7959	7,6791	7,7017	7,5157	7,7894
23 ^{a,b,c}	7,8539	7,7920	8,5969	7,9270	7,7256
24 ^{a,c}	8,3010	7,0069	7,9012	8,3978	7,1958
25	6,5287	6,5447	6,5337	6,6799	6,5683
26	7,4949	7,1560	7,4714	7,5866	7,1818
27 ^{a,c}	7,9586	6,7094	7,7277	7,3818	7,6408
28 ^b	7,4437	7,3055	7,0650	7,3166	7,5256
29 ^b	7,6990	7,4944	7,1636	7,6348	7,3117
30	8,0458	7,8011	7,9629	8,0057	7,8646
31	9,0000	8,6996	8,2404	9,4978	8,5742
32 ^b	8,3010	8,0293	8,7895	8,1499	8,0624
33 ^b	7,3979	8,2823	8,2660	8,2503	8,5425
34	8,5229	8,6617	8,8291	8,6693	8,9002
35	8,3979	8,0774	8,1988	8,2538	8,3187
36 ^{a,b,c}	7,1871	7,7533	7,5839	7,9501	7,6640

37	8,2218	8,2912	8,1295	8,3895	8,2186
38	8,5229	8,4247	8,2543	8,4671	8,3773
39	8,3010	8,5064	8,3501	8,3855	8,1162
40 ^{a,c}	9,0000	8,3851	8,8001	8,2805	7,9540
41 ^{a,c}	7,8539	7,7616	8,3316	8,1486	7,9203
42 ^b	8,3010	8,3546	8,0915	8,7203	8,5819
43 ^{a,b,c}	8,3010	8,7454	8,7789	8,6645	8,5663
44	8,6990	8,5419	8,7739	8,5775	8,3157
45	8,6990	8,8095	9,2176	8,9914	8,6110
46 ^{a,c}	7,9208	8,2862	8,1257	8,2480	8,3210
47	9,0000	8,8505	8,8875	8,5540	8,6222
48 ^b	8,3979	8,1245	8,2724	8,0892	8,7451
49	8,0969	8,4209	8,1431	8,3304	8,1533
50	8,5229	8,3574	8,1616	8,2558	8,4672
51b	8,3979	8,1036	8,1206	8,1735	8,1202
52	8,3979	8,2277	8,2921	8,4145	7,9836
53	8,0969	8,2800	8,4396	7,7687	7,9145
54	8,0969	8,4199	8,4412	8,4226	8,3265
55	7,9586	7,7760	7,9046	7,4158	7,7385
56 ^b	8,6990	8,2945	8,3548	8,3324	8,2700
57 ^b	7,5376	8,5416	8,5798	8,3528	8,4113
58	8,3979	8,5140	8,2164	8,3958	8,3798
59 ^b	6,5670	7,1985	8,3129	6,5492	7,2288
60 ^d		9,3811	9,5035	8,7001	8,4595
61	9,2218	9,2218	9,2218	9,2218	9,2218
62 ^{a,b,c}	9,0969	7,2150	8,0882	8,5580	6,9546
63 ^b	8,6990	8,7864	9,7322	8,7360	8,2930
64 ^{a,c}	8,6990	8,4073	8,9498	8,4125	8,3310
65 ^b	8,3010	8,2642	8,3819	7,8185	7,8964
66 ^d		8,4546	8,6302	8,0062	8,1124
67 ^d		9,0269	9,6067	8,5882	8,2029
68 ^d		9,3162	9,2116	7,9660	8,3143
69	9,0969	8,7194	8,9319	8,5529	7,7593
70	7,3665	7,3829	7,5058	7,5938	7,2867
71 ^d		7,9026	7,8595	7,0992	7,6484
72 ^d		12,6311	10,5304	11,2665	13,9751
73	8,5229	15,0318	15,0460	14,7967	14,7835
74	9,1549	16,3311	16,2106	15,1447	14,9617
75 ^{a,b,c}	8,3979	16,5593	16,1025	15,1336	15,5379
76 ^d		12,8324	10,4068	10,2870	14,3102
77 ^d		6,3138	6,1468	6,1238	7,4494

78 ^d		6,8840	6,2706	6,5076	6,6128
79 ^d		6,6129	5,5857	6,3580	6,5585
80 ^d		6,6671	5,3208	5,6864	7,6878
81 ^d		5,7753	5,9929	6,0619	6,0879
82 ^d		6,5118	6,1130	6,5114	6,9896
83 ^d		8,9102	13,8576	6,9806	7,8620
84 ^d		7,4195	9,2964	6,5851	8,2905
85 ^d		8,7149	8,6394	6,2840	10,2078
86 ^d		6,2538	5,3808	6,3246	6,7906
87 ^d		6,4007	13,0224	6,6308	7,4882
88 ^d		7,1974	7,9383	7,7584	6,5725
89 ^d		7,3440	7,3161	7,3513	7,7591
90 ^d		8,5522	10,0764	7,9644	8,6138
91 ^d		7,9798	8,1410	7,4803	8,1797
92 ^d		7,8981	8,1440	7,7260	7,7939
93 ^d		9,5689	9,7517	9,9771	7,7784
94 ^d		8,6666	8,8504	8,5342	8,8068
95 ^d		8,4295	8,6856	8,0857	8,1514
96 ^d		8,0554	8,2421	8,1512	8,5356
Tablodaki simgeler a: Model 1 için test bileşikleri b: Model 2 için test bileşikleri c: Model 4 için test bileşikleri d: Tüm modeller için aktivitesi bilinmeyen bileşikler					

Serideki bileşikler rastgele (random) eğitim ve test seti olarak seçildikten sonra bileşik seti sabit tutulup 1 parametreden (1 κ_j) 15 parametreye (15 κ_j) kadar taramalar yapılmıştır. 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 aktivite hesaplamalarında kullanılan parametre sayısının az olması gerektiğinden Pim-1 hesaplamalarında model-1 için 10, model-2 için 11, model-3 için 13; Pim-2 için ise model-1 için 10, model-2 için 9, model-3 için 12 için parametreden sonrası için olan hesaplama sonuçları tabloda verilmemiştir.

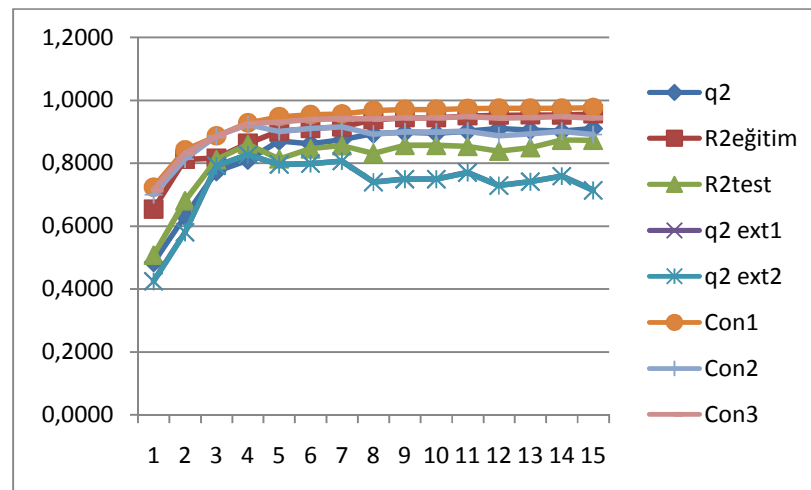
Tablo 4.5.'de Benzotiyenopirimidin serisinde Pim-1 Model 1 (50 bileşik eğitim, 17 bileşik test seti) için aktivite hesaplamalarında aktiviteye en fazla etkisi olan parametre numaraları ile buna karşılık gelen κ_j değerleri görülmektedir.

Tablo 4.5. Benzotiyenopirimidin serisi Pim-1 Model 1 (50 bileşik eğitim, 17 bileşik test) için aktivite hesaplamalarında elde edilen parametre numaraları ve κ_j değerleri

κ_j adedi	Parametre No	κ_j değeri
1	356	-2,4086
2	356, 1604	-2.2733, -6.4867
3	356, 1575, 1604	-1.6614, -0.0037, -6.6872
4	356, 1575, 1604, 1623	-1.5127, -0.0040, -6.6238, -1.0162
5	39, 356, 1575, 1604, 1623	-0.0178, -0.7255, -0.0014, -4.1883, -2.2943
6	39, 356, 1378, 1575, 1604, 1623	-0.0197, -0.5868, -0.5111, -0.0017, -3.0001
7	39, 356, 1378, 1405, 1575, 1604, 1623	-0.0194, -0.8055, -0.6319, 0.6535, -0.0012, -2.6824, -2.311384
8	39, 356, 1333, 1378, 1405, 1575, 1604, 1623	-0.0195, -0.8802, -5.7276, -0.6404, 0.6061, -0.0013, -3.0245, 2.1920
9	39, 356, 1333, 1378, 1381, 1405, 1575, 1604, 1623	-0.01623, -0.6752, -6.8654, -0.6607, 8.8950, 0.8446, -0.0020, -4.3606, -1.9405
10	39, 82, 356, 1333, 1378, 1381, 1405, 1575, 1604, 1623	-0.0174, -0.0195, -0.4089, -7.1321, -0.8370, 13.6307, 0.8891, -0.0018, -3.7387, -2.2817

Şekil 4.5'te bileşik serisinde 50 bileşik eğitim, 17 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği verilmiştir. Bu hesaplamalarda belli bir parametre sayısına kadar regresyon katsayısının ve q^2 'nin yükseldiği görülmektedir. Hem baş parmak kuralına göre hem de şekil 4.5'te yapılan hesaplamalara ilişkin sonuçlarda serisinde eğitim ve test seti için 10 parametre sayısından sonra istatistiksel hesaplamalarda önemli bir değişiklik olmadığı için en uygun parametre sayısı olarak 10 parametre belirlenmiştir.

Şekil 4.5 Benzo-tiyenopirimidin serisi Pim-1 Model 1 (50 bileşik eğitim, 17 bileşik test) için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği

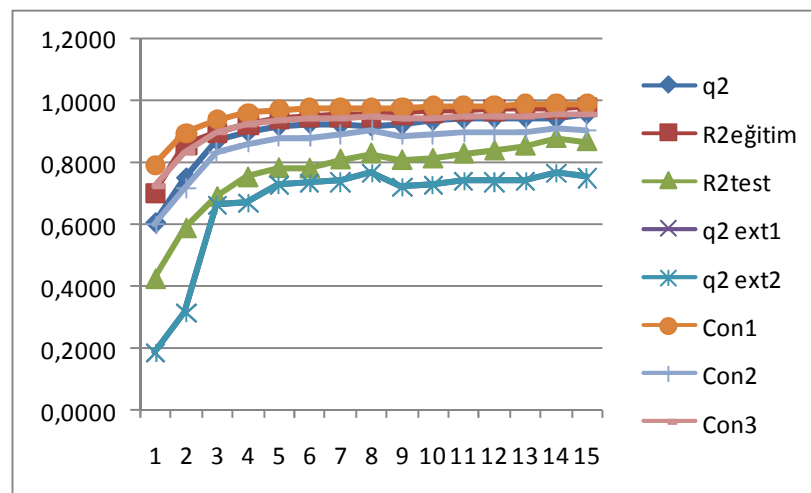


Şekil 4.3'te Benzotiyenopirimidin serisinde 50 bileşik eğitim, 17 bileşik test seti için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği verilmiştir. Bu hesaplamalarda belli bir parametre sayısına kadar regresyon katsayısının ve q^2 nin yükseldiği görülmektedir. Hem baş parmak kuralına göre hem de şekil 4.5'te yapılan hesaplamalara ilişkin sonuçlarda serisinde eğitim ve test seti için 14 parametre sayısından sonra istatistiksel hesaplamalarda önemli bir değişiklik olmadığı için en uygun parametre sayısı olarak 14 parametre belirlenmiştir

Tablo 4.6. Benzotiyenopirimidin serisi Pim-1 Model 2 (45 bileşik eğitim, 22 bileşik test) için aktivite hesaplamalarında elde edilen parametre numaraları ve k_j değerleri.

K_j adedi	Parametre No	k_j değeri
1	356	-2.4583
2	356, 1604,	-2.2733, -6.4867
3	356, 1575, 1604	-1.7678, -0.0033, -7.4987,
4	356, 1391, 1575, 1604	-1.5045, 5.4109, -0.0035, -7.751825
5	51, 356, 1391, 1575, 1604	0.0331, -1.0159, 6.5230, -0.0038, -7.1806
6	51, 356, 1354, 1391, 1575, 1604	0.0324, -0.6419, 1.7066, 8.0188, -0.0038
7	51, 356, 1354, 1391, 1549, 1575, 1604	0.0371, -0.6660, 1.7488, 8.6511, 26.9303, -0.0040, -7.1477

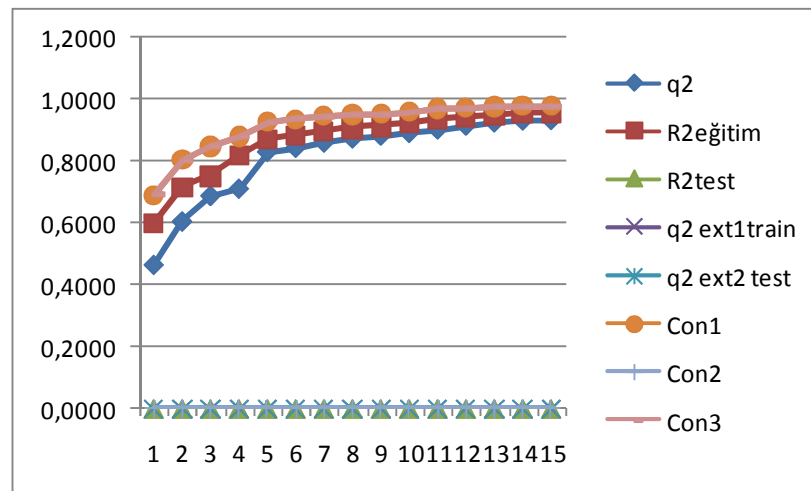
Şekil 4.5. Benzotiyenopirimidin serisi Pim-1 Model 2 (45 bileşik eğitim, 22 bileşik test) için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği



Tablo 4.7. Benzotiyenopirimidin serisi Pim-1 Model 3 (67 bileşik eğitim) için aktivite hesaplamalarında elde edilen parametre numaraları ve k_j değerleri

κ_j adedi	Parametre No	κ_j değeri
1	299	-5.7422
2	299, 356	-3.4295, -1.4422
3	299, 356, 566	-3.5615, -2.4367, 7.823147
4	31, 299, 356, 566	12.5292, -3.5536, -2.4163, 8.1784
5	31, 39, 299, 356, 566	37.4077, -0.0187, -2.3068, -0.9930, 3.8145
6	31, 39, 194, 299, 356, 566	31.4130, -0.0197, -0.0762, -3.756754, -0.838765, 3.752391
7	31, 39, 49, 194, 299, 356, 566	34.7859, -0.0199, 0.0240, -0.100436, -4.227994, -0.555548, 3.462675
8	31, 39, 49, 194, 299, 356, 566, 1388	32.7089, -0.0207, 0.0300, -0.1107, -4.1329, -0.9632, 5.0447, -5.0790
9	31, 39, 49, 194, 299, 356, 566, 1332, 1388	30.9722, -0.0208, 0.0303, -0.1180, -4.1471, -0.9463, 4.5555, 9.3107, -6.6526
10	31, 39, 49, 194, 299, 356, 566, 1332, 1388, 1604	22.1031, -0.0211, 0.0248, -0.1367, -3.8832, -1.1113, 4.556416, 10.3999, -9.0549, -3.2657
11	31, 39, 49, 194, 299, 353, 356, 566, 1332, 1388, 1604	15.8959, -0.0209, 0.0285, -0.1429, -4.0093, -6.2299, -1.3222, 5.6611, 8.7546, -9.6893, -4.0817
12	31, 39, 49, 194, 299, 353, 356, 446, 566, 1332, 1388, 1604	15.5435, -0.0205, 0.0245, -0.1550, -4.2474, -9.4209, -1.5404, 8.9834, 6.3020, 7.8157, -9.3353, -4.4374
13	31, 39, 49, 194, 299, 353, 356, 446, 566, 1332, 1388, 1407, 1604	14.0311, -0.0158, 0.0228, -0.1433, -3.8256, -10.4497, -1.7149, 10.9365, 7.2123, 11.5054, -9.8526, 5.0426, -4.3886

Şekil 4.5. Benzotiyenopirimidin serisi Pim-1 Model 3 (67 bileşik eğitim) için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği

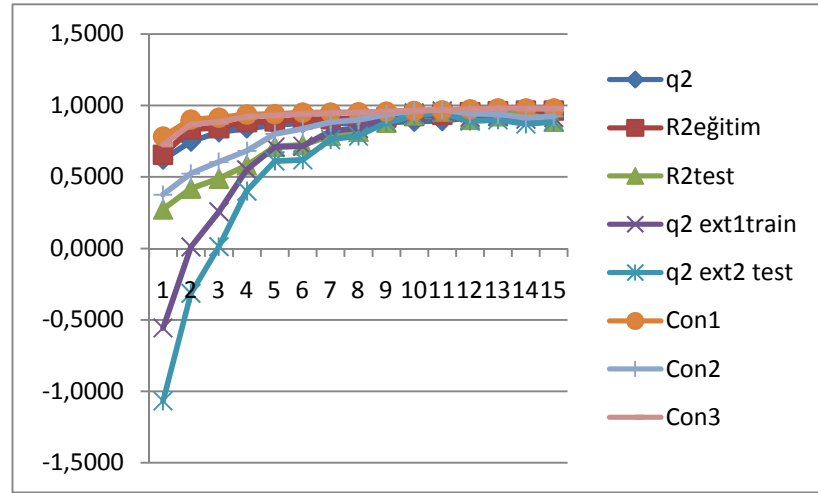


Pim-1 Model 4'te ise her bir bileşiğin tek konformerini dikkate alınarak hesap yapılmıştır. Model 4'te κ_j değerleri Pim-1 için 10 kappa baz alınarak parametre numaralarına göre şöyledir: 39, 282, 356, 1333, 1378, 1381, 1405, 1575, 1604, 1623 için sırasıyla; -0.0180, -0.0284, -0.3857, -4.5276, -0.9272, 0.3195, 0.1844, 0.0006, -0.9021, -3.1339'dur.

Tablo 4.8. Benzotiyenopirimidin serisi Pim-2 Model 1 (50 bileşik eğitim, 17 bileşik test) için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği

κ_j adedi	Parametre No	κ_j değeri
1	356	-2.858135
2	356, 1604	-2.5943, -9.4514
3	356, 1528, 1604	-1.9908, 100.0000, -11.1614
4	356, 1528, 1604, 1646	-2.2221, 100.0000, -8.5717, 0.2679
5	356, 357, 1528, 1604, 1646	-2.6656, 1.8300, 100.0000, -7.5731, 0.3429
6	356, 357, 1368, 1528, 1604, 1646	-2.7127, 2.3549, -4.1738, 100.0000, -7.4656
7	63, 356, 357, 1368, 1528, 1604, 1646	-13.6527, -2.3387, 7.7022, -3.6924, 100.0000, -7.1660, 0.2820
8	63, 356, 357, 446, 1368, 1528, 1604, 1646	-20.3380, -2.3912, 10.4875, 5.1426, -4.2038, 100.0000, -6.8176, 0.2829
9	63, 356, 357, 390, 446, 1368, 1528, 1604, 1646	-36.3445, -2.0363, 16.3815, -33.9620, 15.6580, -3.9596, 100.0000, -7.92151, 0.2464
10	63, 356, 357, 390, 446, 1331, 1368, 1528, 1604, 1646	-43.5783, -1.7383, 19.1075, -35.4683, 17.0989, -10.0295, -3.4817, 100.0000, -7.4599, 0.2476

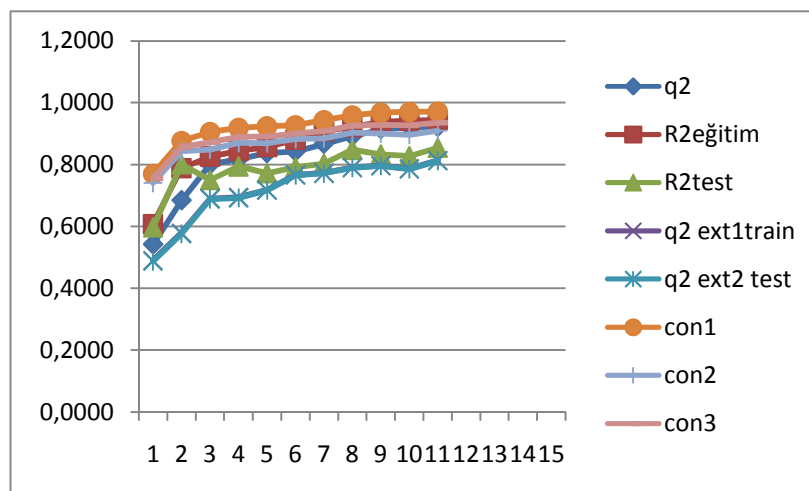
Şekil 4.5. Benzotiyenopirimidin serisi Pim-2 Model 1 (50 bileşik eğitim, 17 bileşik test) için optimum parametre sayısına karşılık regresyon katsayısı ve q^2 grafiği



Tablo 4.9. Benzotiyenopirimidin serisi Pim-2 Model 2 (45 bileşik eğitim, 22 bileşik test) için aktivite hesaplamalarında elde edilen parametre numaraları ve κ_j değerleri

κ_j adedi	Parametre No	κ_j değeri
1	356	-3.2179
2	356, 1600	-3.1100, -5.8857
3	356, 1600, 1601	-2.5769, -5.7312, 0.6968
4	132, 356, 1600, 1601	3.0890, -2.6251, -6.6087, 0.6590
5	132, 356, 1510, 1600, 1601	2.6875, -2.3994, 100.0000, -6.2412, 0.6745
6	132, 356, 1508, 1510, 1600, 1601	2.5347, -2.2366, 15.1802, 100.0000, -6.1173, 0.6270
7	132, 356, 471, 1508, 1510, 1600, 1601	2.6184, -2.1468, 0.0409, 21.8676, 100.0000, -6.6977, 0.6052
8	132, 299, 356, 471, 1508, 1510, 1600, 1601	2.5952, -1.4700, -1.8827, 0.0470, 21.9544, 100.0000, -6.8021, 0.6202
9	132, 299, 356, 471, 1469, 1508, 1510, 1600, 1601	1.8212, -1.7641, -1.1640, 0.0499, 1.5569, 30.6038, 100.0000, -7.7052, 0.6747

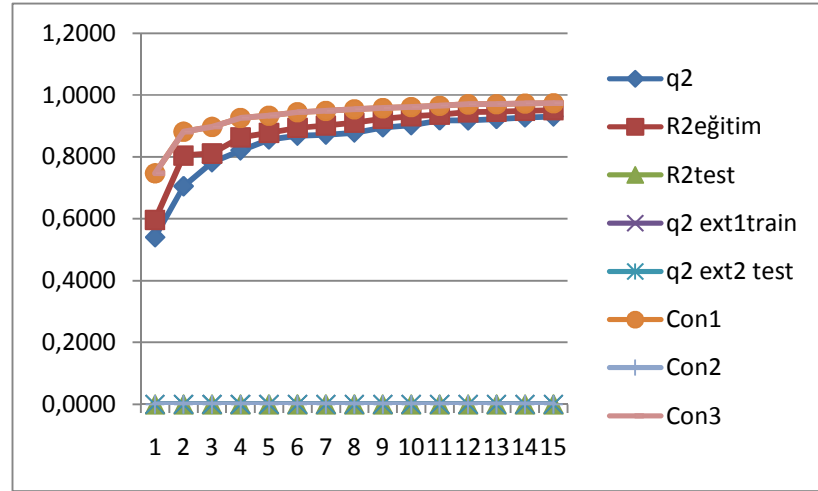
Şekil 4.5. Benzotiyenopirimidin serisi Pim-2 Model 2 (45 bileşik eğitim, 22 bileşik test) için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği



Tablo 4.10 Benzotiyenopirimidin serisi Pim-2 Model 3 (67 bileşik eğitim) için aktivite hesaplamalarında elde edilen parametre numaraları ve κ_j değerleri

κ_j adedi	Parametre No	κ_j değeri
1	356	-2.9475
2	356, 1604	-2.6332, -10.1214
3	356, 1528, 1604	-2.1135, 100.0000, -10.7815
4	356, 1528, 1604, 1646	-2.3291, 100.0000, -8.3094, 0.2854
5	356, 1375, 1528, 1604, 1646,	-2.0714, -12.6373, 100.0000, -7.4397, 0.2809
6	162, 356, 1375, 1528, 1604, 1646,	-1.5107, -2.0216, -14.4238, 100.0000, -7.2174, 0.2425
7	162, 315, 356, 1375, 1528, 1604, 1646	-1.9387, 3.3453, -2.3620, -10.7229, 100.0000, -6.3972, 0.2696
8	162, 315, 356, 1375, 1482, 1528, 1604, 1646	-3.6452, 4.0603, -2.0367, -12.7485, 100.0000, 18.7752, -5.5538, 0.2557
9	162, 315, 356, 1375, 1482, 1528, 1604, 1631, 1646	-4.0991, 2.0266, -1.8028, -15.0906, 100.0000, -0.7734, -6.4095, 0.5826, 0.2748
10	162, 315, 356, 558, 1375, 1482, 1528, 1604, 1631, 1646	-5.2169, 9.6901, -1.3791, -8.7334, -14.9333, 100.0000, -13.0489, -5.9890, 0.6379, 0.2158
11	162, 315, 356, 392, 558, 1375, 1482, 1528, 1604, 1631, 1646	-7.5491, 12.1710, -1.3463, 3.5328, -12.5839, -17.1331, 10.0000, -20.3632, -5.4874, 0.6995, 0.2223
12	134, 162, 315, 356, 392, 558, 1375, 1482, 1528, 1604, 1631, 1646	-0.0704, -7.5616, 11.9395, -1.5670, 3.8659, -11.6301, -16.8794, 99.9998, -22.6559, -5.0466, 0.6934, 0.2366

Şekil 4.5 Benzotiyenopirimidin serisi Pim-2 Model 3 (67 bileşik eğitim) için optimum parametre sayısına karşılık regresyon katsayısı ve q2 grafiği



Model 4'te κ_j değerleri Pim-2 için 10 kappa baz alınarak parametre numaralarına göre şöyledir: 63, 356, 357, 390, 446, 1331, 1368, 1528, 1604, 1646 için sırasıyla; -20.9773, -1.9800, 12.0942, 25.8693, 10.4395, -3.5967, -6.0599, 100.0000, -1.4940, 0.2601.

5. BÖLÜM

TARTIŞMA VE SONUÇ

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. 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 grubumuz tarafımızdan ortaya koyulan yeni yazılımlar geliştirilmiş ve bu çalışma kapsamında benzotiyenopirimidin 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.

Geliştirilen EC-GA Metodu ve EMRE yazılımları modelleme ve simülasyon sonuçları doğrultusunda deneysel çalışmalara temel oluşturması amacıyla zaman kaybetmeden, özgün yeni aktif inhibitörleri önerme potansiyeli bulundurmaktadır.

Çalışmamızda ligand temelli bilgisayar destekli ilaç tasarımı yapılmıştır. EC metot ve genetik algoritma optimizasyon tekniğini birleştiren ve yeni bir 4D-QSAR yaklaşımı olan EC-GA metodu kullanılmıştır. Benzotiyenopirimidin serisi için, geliştirilen EC-GA modelleri bileşiklerin biyolojik aktivitelerinin kontrolünde önemli rol oynayan moleküler özelliklerin anlaşılmasına imkan sağlamıştır.

EC-GA metodu iki kısımdan oluşmaktadır: farmakofor tanımlaması ve aktiviteye etki eden grupların/özelliklerin parametrelendirilerek biyoaktivite tahmininin yapılması. Bu metot ile güvenilir deneysel veriler mevcut olduğu sürece nicel olarak aktivitenin tahmini mümkün olmaktadır.

Bu çalışmadaki diğer önemli bir özellik de hem farmakofor tanımlamasında hem de biyoaktivite tahmininde bütün bileşiklerin konformasyonel esnekliğinin Boltzmann dağılımına bağlı olarak dikkate alınmasıdır. Bu amaç için farmakofor tanımlama, moleküler parametre ve aktivite hesaplamasında kullanılmak üzere kapsamlı bir şekilde geliştirilen EMRE yazılım paketi benzotiyenopirimidin serisi için başarılı bir şekilde uygulanmıştır. Seri için farmakofor tanımlanmış, buna bağlı olarak aktiviteye etki eden önemli parametreler ve parametrelere karşılık gelen κ_j değerleri bulunmuştur. İstenildiğinde her seri için konjenerik herhangi bir bileşiğin ilgili parametre değerleri ve κ_j değerleri aktivite formülünde yerine koyularak çözüldüğünde aktivitesi bilinmeyen ve üzerinde deneysel çalışması yapılmayan bileşiklerin aktivite tahminlerinin başarılı bir şekilde yapılabileceği hesaplanan R^2 (eğitim ve test seti için), q^2 , q^2_{ext1} , q^2_{ext2} , q^2_{ext3} ve ccc ve standart hata değerlerinden anlaşılmaktadır.

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 gerçekleştirildiği için bu çalışma bir 4D-QSAR çalışmasıdır. Geçerliliği en iyi olan modeli belirleyebilmek, modelin performansını ölçmek ve sonuçların doğruluğunu test etmek için iki ayrı model kullanılmıştır.

Modellerin geçerliliğini ve güvenilirliğini doğrulamak için LOO-çapraz doğrulama (Leave-one-out Cross Validation), regresyon, dahili ve harici doğrulama ve uyum korelasyon katsayısı (CCC) analizleri istatistiksel olarak yapılmıştır. Çalışmanın sonucunda sentez edilmemiş aktivitesi bilinmeyen bileşiklerin aktiviteleri tahmin edilmiştir.

EC-GA metodunun sonuçlarına bakıldığında yeni bileşikleri dizayn etmek için bu metodun kullanışlı olduğu görülmektedir. Kullanılan bilgisayarların hızları yeterli ise

geliştirilen bu programlar ile Hartree Fock, Density Functional Theory ve Møller Plesset hesaplamalarının sonuçları da aktivite hesaplamalarında kullanılabilir.

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

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

KAYNAKLAR

1. Itai, A., Tomioka, N., Kato, K., 1995. QSAR and Drug Design: New Developments and Applications, Cilt 23, (Ed.: Toshio Fujita), s. 3-4, Elsevier Science Pub Co., the Netherlands.
2. Thomas, G., 2003. Fundamentals of Medicinal Chemistry, s.56-70, Wiley-Blackwell, Weinheim.
3. Hopfinger, A.J., Wang, S., Tokarski, J.S., Jin, B., Albuquerque, M., Madhav, P.J., Duraiswami, C., 1997. Construction of 3D-QSAR Models Using the 4D-QSAR Analysis Formalism, J. Am. Chem. Soc. 119, 10509-10524.
4. Wermuth, C.G., 2006. Pharmacophores and Pharmacophore Searches, Cilt 32 (Ed.: Langer, T., Hoffmann, R.D.), s. 3-11, Wiley-VCH, Weinheim.
5. Guner, O.F., 2002. History and Evolution of the Pharmacophore Concept in Computer-Aided Drug Design, Curr. Top. Med. Chem. 2, 1321-1332.
6. Dimoglo, A.S., et al., 1995. Electronic-Topological Investigations of the Relationships Between Chemical Structure and Ambergri Odor, New Journal of Chemistry 19, 1217-1226.
7. Saripinar, E., Guzel, Y., Patat, S., Yildirim, I., Akcamur, Y., Dimoglo, A., 1996. Electron-Topological Investigation of the Structure-Antitubercular Activity Relationship of Thiosemicarbazone Derivatives, Arzneim. Forsch./Drug Res., 46, 824-828.
8. Guzel, Y., Saripinar, E., Yildirim, I., 1997. Electron-Topological (ET) Investigation of Structure-Antagonist Activity of a Series of Dibenzo[a,d]Cycloalkenimines, J. Mol. Structure (Theochem), 418, 83-91.
9. Bersuker, I.B., Boggs, J.E., 2001. Pharmacophore Identification and Bioactivity Prediction for Group I Metabotropic Glutamate Receptor Agonists by the Electron-Conformational QSAR Method, Quant. Struct.-Act. Relat., 20, 327-334.
10. Makkouk, Al H., Bersuker, I.B., Boggs, J.E., 2004. Quantitative Drug Activity Prediction for Inhibitors of Human Breast Carcinoma, Int. J. Pharm. Med., 18, 81-89.
11. Marenich, A.V., et al., 2008. 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.

12. Bersuker, I.B., Bahceci, S., Boggs, J.E., Pearlman, R.S., 1999. A Novel Electron-Conformational Approach to Molecular Modeling for QSAR by Identification of Pharmacophore and Anti-Pharmacophore Shielding, *SAR QSAR Environ. Res.*, 10, 157-173.
13. Ehresmann, B., Groot, M.J., Alex, A., Clark, T., 2004. New Molecular Descriptors Based on Local Properties at the Molecular Surface and a Boiling-Point Model Derived from Them, *J. Chem. Inf. Comput. Sci.*, 44 (2), 658–668.
14. Duchowicz, P.R., Bucknum M. J., Castro, E.A., 2007. New Molecular Descriptors Based upon the Euler Equations for Chemical Graphs, *Journal of Mathematical Chemistry*, 41, 193-208.
15. Eriksson, L., Andersson, P.L., Johansson, E., Tysklind, M., 2006. Analysis of Environmental QSAR Data. Part I – A Basic Framework Founded on Principal Component Analysis (PCA), Partial Least Squares (PLS), and Statistical Molecular Design (SMD), *Mol. Divers.* 10, 169–186.
16. Dudek, A.Z., Arodz, T., Galvez, J., 2006. Computational Methods in Developing Quantitative Structure-Activity Relationships (QSAR): A Review, *Comb. Chem. High T. Scr.*, 9, 213-228.
17. Guha, R., Jurs, P.C., 2005. Interpreting Computational Neural Network QSAR Models: A Measure of Descriptor Importance, *J. Chem. Inf. Model.*, 45, 800-806.
18. Holland, J.H., 1975. *Adaptation in Artificial and Natural Systems*, Dissertation, University of Michigan, Ann Arbor, Michigan.
19. Terfloth, L., Gasteiger, J., 2001. Neural Networks and Genetic Algorithms in Drug Design, *DDT.*, 6, 102-108.
20. Tao Z.F., *et al.*, 2009. Discovery of 3H-benzo[4,5]thieno[3,2-d]pyrimidin-4-ones as potent, highly selective and orally bioavailable inhibitors of the human protooncogene proviral insertion site in moloney murine leukemia virus (Pim) kinases, *Journal of Medicinal Chemistry*, 52 (21): 6621–6636.
21. http://en.wikipedia.org/wiki/Serine/threonine-specific_protein_kinase.
22. <http://en.wikipedia.org/wiki/PIM1>.
23. Tosun, M., 2007. *Farmakolojiye Giriş Ders Notları*, Ege Üniversitesi, İzmir.
24. Bersuker, I.B., Dimoglo, A.S., Gorbachov, M.Yu., 1987. The Electron-Topologic Approach to The QSAR Problem Illustrated by Inhibitor Activity for Thymidine Phosphorylase and á-Chymo-Trypsin, In *QSAR in Drug Design and Toxicology*, ed: Hadzi, D., Jerman-Blazic, B., Vol:10, Elsevier, Amsterdam, pp:43-48.

25. Hopfinger, A. J., Tokarski, J. S., 1997. Three-Dimensional Quantitative Structure-Activity Relationships. In *Practical Application of Computer-Aided Drug Design*, Charifson, P. S., Ed., Marcel Dekker, New York, pp 105-164.
26. Oprea, T. I., Waller, C. L., 1997. Theoretical and Practical Aspects of Three-Dimensional Quantitative Structure-Activity Relationships. In *Reviews in Computational Chemistry*, Lipkowitz, K. B., Boyd, D. B., Eds., Wiley-VCH, New York, Vol. 11, pp 127-182.
27. Karacan, N., 2008. 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.
28. Hansch, C., et al., 1962. Correlation of biological activity of phenoxyacetic acids with hammett substituent constants and partition coefficients, *Nature*, 194, 178-180.
29. Debnath, A.K., 2001. Quantitative Structure-Activity Relationship (QSAR) Paradigm-Hansch Era to New Millennium, *Mini-Rev. Med. Chem.*, 1, 187–195.
30. Richet, C., 1993. 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.
31. Meyer, H., 1899. On the theory of alcohol narcosis I. Which property of anesthetics gives them their narcotic activity, *Arch. Exp. Pathol. Pharmacol.*, 42, 109-118.
32. Overton, E., 1991. Studien uber die Narkose, zugleich ein Beitrag zur allgemeinen Pharmakologie, G. Fischer, Jena, 1901; English translation by Lipnick, R.L. 136 (ed.), *Studies on Narcosis*, Charles Ernest Overton, Chapman and Hall, London.
33. Fühner, H., Neubauer, E., 1907. Hämolyse durch Substanzen homologen Reihen, *Arch. Exp. Path. Pharm.*, 56, 333-345.
34. Bruice, T.C., Kharasch, N., Winzler, R.J., 1956. A correlation of thyroxine-like activity and chemical structure, *Arch. Biochem. Biophys.*, 62, 305 -317.
35. Zahradnik, R., Chvapil, M., 1960. Study of the Relationships Between the Magnitude of Biological Activity and the Structure of Aliphatic Compounds, *Experientia*, 16, 511-512.
36. Hammett, L.P., 1970. *Physical Organic Chemistry. Reaction Rates, Equilibria and Mechanism*, 2nd Ed., McGraw-Hill, New York.
- [37]. Hansen, O.R., 1962. Hammett Series With Biological Activity, *Acta Chem. Scand.*, 16, 1593-1600.
38. Hansch, C., Fujita, T., 1964. Rho-sigma-pi analysis. Method for correlation of biological activity and chemical structure, *J. Am. Chem. Soc.*, 86, 1616–1626.

39. Free, Jr., S.M., Wilson, J.W., 1964. A mathematical contribution to structure activity studies, *J. Med. Chem.*, 7, 395-399.
40. Todeschini, R., Consonni, V., 2000. *Handbook of Molecular Descriptors*, v.11, Mannhold, R., Kubinyi, H., Timmerman, H. (eds.), Wiley-VCH, Weinheim.
41. Karelson, M., 2004. Quantum chemical descriptors in QSAR. In *Computational Medicinal Chemistry for Drug Discovery*, pp. 1-32, Bultinck, P., De Winter, H., Langenaeker, W., Tollenaere, J.P. (eds.), Dekker Inc., New York, Basel.
42. Fruhbeis, H., Klein, R., Wallmeier, H., 1987. Computer-Assisted Molecular Design (CAMD), An Overview. *Angew. Chem. Int. Ed. Engl.*, 26, 403-418.
43. Perkins, R., Fang, H., Tong, W., Welsh, W.J., 2003. Quantitative structure-activity relationship methods: Perspectives on drug discovery and toxicology, *Environmental Toxicology and Chemistry*, 22(8), 1666-1679.
44. Lowis, D.R., 1997. HQSAR: A New, Highly Predictive QSAR Technique, *Tripos Technical Notes*, 1(5), 1-17.
45. Kubinyi, H., 2002. From Narcosis to Hyperspace: The History of QSAR, *Quant. Struct.-Act. Relat.*, 21, 348-356.
46. Klebe, G., 1998. 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.
47. Vedani, A., et al., 2000. Multiple-Conformation and Protonation-State Representation in 4D-QSAR: The Neurokinin-1 Receptor System, *J. Med. Chem.*, 43, 4416-4427.
48. Lill, M. A., 2007. Multi-dimensional QSAR in drug discovery, *Drug Discovery Today*, 12, 1013-1017.
49. Chitty, M., 2010. In Silico & molecular drug modeling glossary & taxonomy, Cambridge Healthtech Institute, Needham, USA, http://www.genomicglossaries.com/content/molecular_modeling_gloss.asp, Jan.
50. Smith, C., 1998. Molecular Modeling-Seeing the Whole Picture with Modeling Software Packages, *Scientist*, 12(17), 31.
51. Dimoglo, A. S., 1987. The Electron-Topological Approach in the Search and Design of Biologically Active Compounds. Thesis of Doctor of Sciences, Rostow/Don University. (Russ.)
52. Bersuker I.B., Dimoglo, A.S., Gorbachov, M.Yu. , Vlad, P.F., Pesaro, M. 1991. Origin of Musk Fragrance Activity, The Electron-Topological Approach. *New J. Chem.* 15, 307-320.

53. Bersuker, I.B., Bahceci, S., Boggs, J.E., and Pearlman, R.S., 1999. 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.
54. Bersuker, I.B., Bahceci, S., and Boggs, J.E., 2000. Improved Electron-Conformational Method of Pharmacophore Identification and Bioactivity Prediction. Application to Angiotensin Converting Enzyme Inhibitors, *J. Chem. Inf. Comput. Sci.* 40, 1363–1376.
55. Bersuker, I.B., and Dimoglo, A.S. 1991. The Electron-Topological Approach to the QSAR Problem, in: Lipkowitz, K.B., Boyd, D.B. (Eds.), *Reviews in Computational Chemistry*, Vol. 2, VCH, , New York, , pp. 423–460.
56. Bersuker, I.B., 2003. Pharmacophore Identification and Quantitative Bioactivity Prediction Using the Electron-Conformational Method, *Current Pharmaceutical Design*, 9, 1575-1606.
57. Mason, J. S., Good, A. C., Martin, E. J., 2001. 3-D Pharmacophores in Drug Discovery *Current Pharmaceutical Design*, 7, 567-597.
58. Goldberg D.E., 1989. *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, USA.: 1-7.
59. Yeniay, 2001: 37 Yeniay Ö. 2001. “An Overview of Genetic Algorithms”, *Anadolu Üniversitesi Bilim ve Teknoloji Dergisi*, Cilt: 2, Sayı: 1, s. 37-49.
60. Engin, 2001: 21 Engin O. 2001. Akış Tipi Çizelgeleme Problemlerinin Genetik Algoritma ile Çözüm Performansının Arttırılmasında Parametre Optimizasyonu, İTÜ, Fen Bilimleri Enstitüsü, Yayınlanmamış Doktora Tezi.
61. Jang, 1997: 175-176, Jang J. S. R., 1997. *Neuro-Fuzzy and Soft Computing: A Computational Approach To Learning and Machine Intelligence*, Chapter 7: Derivative-Free Optimization, Prentice-Hall, USA, s. 173-196.
62. Yeo M. F. ve Agyel E. O., 1996. “Optimising Engineering Problems Using Genetic Algorithms”, *Engineering Computations*, Volume: 15, Number: 2, s. 268-280.
63. Braysy, 2001: 54 Braysy O. 2001. *Local Search and Variable Neighborhood Search Algorithms for The Vehicle Routing Problem With Time Windows*, PH D Thesis.
64. Fung R.Y.K. , Tang J. ve Wang D., 2001. “Extension Of A Hybrid Genetic Algorithm For Nonlinear Programming Problems With Equality And Inequality Constraints”, *Computers & Operations Research*, Volume: 29, Issue: 3, s. 261-274.
65. Altıparmak, Dengiz ve Smith, 2000: 61-62 Altıparmak F., Dengiz B. ve Smith A.E. (2000), “An Evolutionary Approach For Reliability Optimization in Fixed Topology

Computer Networks”, *Transactions On Operational Research*, Volume: 12, Number: 1-2, s. 57- 75.

66. Gonzales E. L. ve Fernandez M.A.R. 2000. “Genetic Optimisation of A Fuzzy Distribution Model”, *International Journal of Physical Distribution & Logistics Management*, Volume 30, Number 7/8, s. 681-696.

67. Sarıpınar, E., Geçen, N., Sahin, K., Yanmaz, E., 2010. Pharmacophore identification and bioactivity prediction for triaminotriazine derivatives by electron conformational-genetic algorithm QSAR method, *European Journal of Medicinal Chemistry*, 45 (9): 4157-4168. ____ Akyüz L., Sarıpınar E., Kaya E., Yanmaz E., 2012. 4D –QSAR study of HEPT derivatives by electron conformational-genetic algorithm QSAR method, *SAR and QSAR in environmental research*, 23 (5-6): 409-433.

68. Sarıpınar, E., Geçen, N., Sahin, K., Yanmaz, E., Çopur F., 2012. 4D-QSAR analysis and pharmacophore modeling : electron conformational-genetic algorithm approach for penicillins, *Bioorganic & Medicinal Chemistry*, 19 (7): 2199–2210.

69. Geçen, N., Sarıpınar, E., Yanmaz, E., Sahin, K., 2011. Application of electron conformational–genetic algorithm approach to 1,4-dihydropyridines as calcium channel antagonists: pharmacophore identification and bioactivity prediction. *Journal of Molecular Modeling*, 18 (1):65-82.

70. Akyüz L., Sarıpınar E., Conformation depends on 4D-QSAR analysis using EC–GA method: pharmacophore identification and bioactivity prediction of TIBOs as non-nucleoside reverse transcriptase inhibit, *Journal of Enzyme Inhibition and Medicinal Chemistry*, doi:10.3109/14756366.2012.684051.ors

71. Geçen N., 2010. Artemisinin, Raloksifen, Tetrahidroizokuinolin, 1,4-Dihidropiridin ve Benzoksazinon Türevlerinin Nicel Yapı Aktivite İlişkilerinin İncelenmesi ve Aktif Yapının Belirlenmesi, Erciyes Üniversitesi, Doktora Tezi, Kayseri, 182 s.

72. Şahin K., 2010. İnden Pirozol Piridin, Geldanamisin, Triazin Piroolidin Amin, Primidin, Fosfonat Türevlerinin ve Farnasiltransferaz İnhibitörlerinin Genetik Algoritma Kullanılarak EC-QSAR Yöntemi İle Aktivite Hesabı. Erciyes Üniversitesi, Doktora Tezi, Kayseri, 158 s.

73. Yanmaz E., 2010. Elektron Konformasyonel-Genetik Algoritma Qsar Metodu ile Penisilin Türevlerine Ait Farmakofor Grubunun Belirlenmesi ve Biyoaktivite Hesabı. Erciyes Üniversitesi, Yüksek Lisans Tezi, , Kayseri, 142 s.

74. Aydın, S., 2010. 3-(1,1-Diokso-2H-(1,2,4)-Benzotiyodiazin-3-il)-4-Hidroksi-2(1H)-Kinolin Bileşik Serisinin Hepatit C İnhibitörü olarak Elektron Konformasyonel-Genetik Algoritma (EC-GA) Yöntemi ile QSAR İncelenmesi. Erciyes Üniversitesi, Yüksek Lisans Tezi, Kayseri, 127 s.

75. Akyüz , L., 2011. Timin, Benzodiazepin, Tiyazolidin, Pirimidin ve İmidazolidinon Türevlerinin Kantitatif Aktivitelerinin Belirlenmesi. Erciyes Üniversitesi, Yüksek Lisans Tezi, Kayseri, 187 s.
76. Bersuker I.B., Dimoglo, A.S., Gorbachov, M.Yu. , Vlad, P.F., Pesaro, M. Origin of Musk Fragrance Activity, The Electron-Topological Approach. *New J. Chem.* (1991), 15, 307-320.
77. Aleksandr V. Marenich , Pei-Han Yong , Isaac B. Bersuker ,and James E. Boggs, 2008. *J.Chem. Inf. Model.* ,48,556-568.
78. Bersuker I.B., 2008. *J. Comput Aided Mol Des* 22:423:430.
79. Parthasarathi R., Subramanian, V., Roy, D.R., Chattaraj P:K., 2004. Electrophilicity index as a possible descriptor of biological activity, *Bioorganic & Medicinal Chemical* 12, 5533-5543.
80. F. Fan, J. Cheng, Z.Li, X. Xu and X. Qian, 2009. Novel Dimer Based Descriptors with Solvational Computation for QSAR Study of Oxadiazoylbenzoyl-ureas as Novel Insect-growth Regulators, *Journal of Computational Chemistry*. 31 pp. 586-591.
81. P. Gramatica, P. Pilutti and E. Papa, 2004. Validated QSAR Prediction of OH Tropospheric Degradation of VOCs: Splitting into Training-Test Sets and Consensus Modeling, *J. Chem. Inf. Comput. Sci.* 44 pp. 1794-1802.
82. G. Schuurmann, R. U. Ebert, J. Chen, B. Wang and R. Kuhne, 2008. External Validation and Prediction Employing the Predictive Squared Correlation Coefficients Test Set Activity Mean vs Training Set Activity Mean, *J. Chem. Inf. Model.* 48 pp. 2140–2145.
83. V. Consonni, D. Ballabio and R. Todeschini, 2009). Comments on the Definition of the Q2 Parameter for QSAR Validation, *J. Chem. Inf. Model.* 49 pp. 1669–1678.
84. Chirico N., Gramatica P., 2011. Real external predictivity of QSAR models: how to evaluate it? Comparison of different validation criteria and proposal of using the concordance correlation coefficient, *Journal of Chemical Information and Modeling* Article 51 (9): 2320-2335.

Bu proje kapsamında katıldığımız bilimsel kongreler:

Sevtap ÇAĞLAR, Semiha AYDIN, Harun KESKİN, Sevinç ÇATALKAYA, Emin SARIPINAR, '4D-QSAR EC-GA Metodu İle Benzotiyenopirimidin Bileşik Serisinin Farmakofor Grubunun Belirlenmesi Ve Biyoaktivite Hesabı' III. Fiziksel Kimya Günleri Kongresi, (Balıkesir) Burhaniye, 12-15 Temmuz 2012.

Sevtap Çağlar, Semiha AYDIN, Harun KESKİN and Emin Sarıpınar 'Determination of Pharmacophore Group and Bioactivity Account to Benzothienopyrimidinones, Compounds by Electron Conformational-Genetic Algorithm Method', 10. Kimyasal Fizik Kongresi (CPC-X), Ankara, 10-12 Ekim 2012.