

بِسْمِ اللَّهِ الرَّحْمَنِ الرَّحِيمِ



دانشگاه پیام نور

دانشکده علوم پایه

پایان نامه

برای دریافت مدرک کارشناسی ارشد

رشته شیمی (گرایش تجزیه)

گروه شیمی

**عنوان پایان نامه:**

**مطالعه روابط کمی ساختار- فعالیت برای پیش بینی**

**فعالیت دارویی ضد ایدز مشتقات ایمیدازول**

**دانشجو:**

**مریم احمدی**

**استاد راهنما:**

**دکتر اسماعیل تماری**

**استاد راهنمای همکار:**

**دکتر زهرا رستمی**

**استاد مشاور:**

**دکتر محمد حق گو**

**مهر ۱۳۹۲**

تاریخ:

شماره:



دانشگاه پیام نور

استان همدان

بسمه تعالی

### صور تجلسه دفاع از پایان نامه دوره کارشناسی ارشد

جلسه دفاع از پایان نامه دوره کارشناسی ارشد خانم مریم احمدی دانشجوی رشته شیمی (گرایش تجزیه) به شماره دانشجویی ۸۹۰۰۶۸۲۹۵ تحت عنوان ((مطالعه روابط کمی ساختار- فعالیت برای پیش بینی فعالیت دارویی ضد ایدز مشتقات ایمیدازول)) با حضور هیات داوران در روز یکشنبه مورخ ۹۲/۷/۲۸ ساعت ۱۳ بعدظهر در محل ساختمان جدید سالن دفاع دانشگاه پیام نور همدان برگزار شد و پس از بررسی، پایان نامه مذکور با نمره به عدد ۱۹/۷۵ به حروف نوزده و هفتاد و پنج صدم با درجه عالی مورد تایید واقع شد.

ردیف	نام و نام خانوادگی	هیات داوران	مرتبہ دانشگاهی	دانشگاه/مؤسسه	امضا
۱	دکتر اسماعیل تماری	استاد راهنما	استادیار	پیام نور همدان	
۲	دکتر زهرا رستمی	استاد راهنمای همکار	استادیار	پیام نور همدان	
۳	دکتر محمد حق گو	استاد مشاور	استادیار	پیام نور رزن	
۴	دکتر رضا آزادبخت	استاد داور داخلی	استادیار	پیام نور همدان	
۵	دکتر عباس افخمی	استاد داور خارجی	استادیار	بوعلی سینا	
۶	دکتر عباس امینی منش	نماینده گروه و مدیر گروه علوم پایه استان	استادیار	پیام نور همدان	

## تقدیم به:

پیشگاه مقدس و آسمانی امام عشق و نور، حضرت علی بن موسی الرضا (ع)

این پایان نامه را تمام خضوع تقدیم به خانواده عزیزم:

الهه ی مهرم: مادر، که خانه ی عروج با دعای او بنا شد.

اسطوره ی جاودانه ایثار و فداکاری: پدر، که از او آموختم چگونه سبکبال زندگی کنم تا هجرتم نیز سبکبال باشد.

امید بودن و پناه خستگی ام: همسرم، که مرا باور کرد و تمام سختی ها و مشکلات را با سعه صدر تحمل کرد و با وجود پر مهرش در لحظات پر تلاطم زندگی همواره مونس و همراهم بوده است.

سپاس بی کران پروردگار یکتا را که آثار قدرت او بر چهره روز روشن، تابان است و انوار حکمت او در دل شب تار، درخشان. آفریدگاری که خویشتن را به ما شناساند و درهای علم را بر ما گشود و عمری و فرصتی عطا فرمود تا بنده ضعیف خویش را در طریق علم و معرفت بیازماید.

صمیمانه ترین سپاس ها نثار:

استاد با کمالات و شایسته، سرکار خانم زهرا رستمی که با حسن خلق و فروتنی از هیچ کمکی در این عرصه بر من دریغ ننمودند و زحمت راهنمایی این پایان نامه را بر عهده گرفتند. و همچنین اساتید فرهیخته و دلسوز، جناب آقای دکتر اسماعیل تماری و جناب

آقای دکتر محمد حق گو که سخاوتمندانه سرمایه علم و معرفتشان را در اختیارم قرار داده اند. و جناب آقای دکتر امینی منش مدیر تحصیلات تکمیلی و جناب آقای دکتر آزادبخت و آقای دکتر افخمی که زحمت قرائت و داوری این پایان نامه را بر عهده داشتند.

با تشکر فراوان از دوستان خوبم، جناب آقای دکتر علی عادلخانی و خانم بیات ورکشی و همه کسانی که در همه حال بر شادمانی و دانشم افزودند آرزومند آرزوهایشانم،  
فردایتان روشن.

## چکیده:

**عنوان پایان نامه:** مطالعه روابط کمی ساختار- فعالیت برای پیش بینی فعالیت دارویی ضد ایدز مشتقات ایمیدازول.

**هدف:** در این مطالعه مدل یابی فعالیت ضد ایدز و ضد آروماتاز (بیماری سرطان سینه) مبتنی بر ساختار مولکولی برخی از ترکیبات مربوط به مشتقات ایمیدازول، پرمیدازول، فتالازینون و بنزآلفاتالید همچنین مشتقات تری آزول و ایمیدازول با استفاده از روابط کمی ساختار- فعالیت مورد بررسی قرار گرفت.

**روش تحقیق:** مطالعه بر روی دو مجموعه اصلی شامل ۴۲ مولکول با فعالیت ضد ایدز و ۵۷ مولکول با فعالیت ضد آروماتاز (بیماری سرطان سینه) با ۵۰٪ حداقل غلظت بازدارندگی معین انجام شده است. این مطالعه بر اساس مقایسه سه روش رگرسیون خطی چند گانه (MLR)، شبکه عصبی مصنوعی (ANN) و سیستم استنتاج تطبیقی عصبی - فازی (ANFIS) برای پیش بینی فعالیت ترکیبات دارویی، قبل از سنتز آنها می باشد.

**یافته ها:** در روش اول ضریب همبستگی برای هردو دسته ترکیبات، به ترتیب ۰/۳۷۱ و ۰/۵۷۲، روش دوم ضریب همبستگی ۰/۹۱۶ و ۰/۹۱۲ و در روش سوم ضریب همبستگی ۰/۹۱۸ و ۰/۹۱۱ بدست آمده است.

**نتیجه گیری:** با توجه به فعالیت های پیش بینی شده و ضریب همبستگی هردو دسته ترکیبات، میتوان نتیجه گرفت که روشهای غیر خطی شامل (ANN) و (ANFIS) موفق تر از روش خطی (MLR) بوده است.

**واژگان کلیدی:** روابط کمی ساختار - فعالیت، ایمیدازول، پرمیدازول، فتالازینون، بنزآلفاتالید، ایمیدازول، ضریب همبستگی، MLR، ANN، ANFIS.

## فهرست مطالب

عنوان	صفحه
فصل اول: مقدمه	
۱-۱- مقدمه.....	۱
۲-۱- تاریخچه QSAR.....	۱
۱-۲-۱- اصول اساسی.....	۲
۲-۲-۱- روابط انرژی آزاد خطی.....	۴
۳-۲-۱- تجزیه و تحلیل‌های هانش.....	۷
۴-۲-۱- موضوع مهم در QSAR.....	۸
۵-۲-۱- شاخص های مولکولی.....	۹
۶-۲-۱- بیان و شرح متغیرها.....	۹
۷-۲-۱- انتخاب متغیرها.....	۱۰
۸-۲-۱- آنالیز جزء اصلی (PCA).....	۱۱
۹-۲-۱- طبقه بندی از مدل های QSAR.....	۱۲
۱۰-۲-۱- روش های پیش بینی و ارزیابی و تفسیر مدل.....	۱۲
۱۱-۲-۱- هدف QSAR.....	۱۳
۳-۱- شبکه عصبی مصنوعی.....	۱۴
۱-۳-۱- شباهت های شبکه عصبی مصنوعی و بیولوژیکی.....	۱۴
۲-۳-۱- ویژگی شبکه عصبی مصنوعی.....	۱۵
۴-۳-۱- آموزش شبکه عصبی.....	۱۶
۵-۳-۱- روش های مختلف آموزش شبکه عصبی مصنوعی.....	۱۶
۶-۳-۱- الگوریتم های آموزش شبکه عصبی مصنوعی.....	۱۷
۷-۳-۱- توابع فعالیت.....	۱۸
۸-۳-۱- ساختارهای مختلف شبکه عصبی مصنوعی.....	۱۸
۹-۳-۱- شبکه پرسپترون چند لایه.....	۱۹
۴-۱- منطق فازی.....	۲۰
۱-۴-۱- تاریخچه منطق فازی.....	۲۰
۲-۴-۱- تعریف سیستم فازی.....	۲۰
۳-۴-۱- انواع سیستم های فازی.....	۲۱

## ادامه فهرست مطالب

- ۲۲-۴-۱-۴- انواع نافازی ساز..... ۲۲
- ۲۲-۴-۱-۵- قواعد فازی و سیستم های استنتاج فازی..... ۲۲
- ۲۳-۴-۱-۶- مراحل انجام استدلال فازی به وسیله سیستم استنتاج فازی..... ۲۳
- ۲۴-۵-۱- سیستم استنتاج تطبیقی عصبی-فازی (ANFIS)..... ۲۴
- ۲۴-۱-۵-۱- به روزرسانی پارامترها..... ۲۴
- ۲۵-۵-۱-۲- ساختار سیستم استنتاج تطبیقی عصبی-فازی..... ۲۵
- ۲۶-۶-۱- معرفی ترکیبات..... ۲۶
- ۲۷-۱-۶-۱- معرفی ترکیبات دسته اول..... ۲۷
- ۲۷-۱-۶-۱-۱- ایدز چیست؟..... ۲۷
- ۲۷-۱-۶-۱-۲- تاریخچه ایدز..... ۲۷
- ۲۸-۱-۶-۱-۳- تسخیر سیستم دفاعی..... ۲۸
- ۲۸-۱-۶-۱-۴- دارو در مقابل ایدز..... ۲۸
- ۲۹-۱-۶-۱-۵- امیدازول..... ۲۹
- ۳۰-۱-۶-۱-۶- فتالازین..... ۳۰
- ۳۱-۲-۶-۱- معرفی ترکیبات دسته دوم..... ۳۱
- ۳۱-۱-۶-۱-۲- مهار کننده آروماتاز، افزایش دهنده طول عمر مبتلایان به سرطان پستان..... ۳۱
- ۳۱-۲-۶-۱-۲- سرطان پستان..... ۳۱
- ۳۲-۳-۲-۶-۱- درمان سرطان پستان..... ۳۲
- ۳۳-۴-۲-۶-۱- داروهای مورد استفاده در درمان هورمونی..... ۳۳
- ۳۴-۵-۲-۶-۱- سنتز داروهای ضد سرطان لتروزول..... ۳۴
- ۳۴-۶-۲-۶-۱- تری آزول..... ۳۴
- ۳۵-۷-۱- تحقیقات پیشین در زمینه QSAR..... ۳۵
- ۳۵-۱-۷-۱- ترکیبات ضد ایدز و QSAR..... ۳۵
- ۴۰-۲-۷-۱- ترکیبات ضد آروماتاز و QSAR..... ۴۰

## فصل دوم: روشهای محاسباتی

- ۴۴-۱-۲- استخراج داده‌های فعالیت..... ۴۴
- ۴۴-۲-۲- مراحل استفاده از نرم افزارهای بکار برده شده در این مطالعه..... ۴۴
- ۴۴-۱-۲-۲- نرم افزارهای بکار برده شده در روش رگرسیون خطی چند گانه..... ۴۴



## ادامه فهرست مطالب

۴۵	۲-۲-۲- نرم افزارهای به کار برده شده در روش شبکه عصبی مصنوعی.....
۴۵	۳-۲-۲- نرم افزار بکار برده شده در ANFIS.....
۴۶	۳-۲- معرفی برخی از توصیف گرهای مورد استفاده.....
۴۶	۱-۳-۲- توصیف گرهای ۳D-MORSE.....
۴۷	۲-۳-۲- توصیف گرهای Topological.....
۴۷	۳-۳-۲- توصیف گرهای WHIM.....
۴۷	۴-۲- پارامترهای آماری مورد استفاده در رگرسیون خطی چند گانه.....
۴۸	۱-۴-۲- رگرسیون خطی چند گانه.....
۴۸	۲-۴-۲- ضریب همبستگی.....
۵۰	۳-۴-۲- انحراف استاندارد.....
۵۱	۴-۴-۲- واریانس.....
۵۱	۵-۴-۲- کوواریانس.....
۵۱	۶-۴-۲- ضریب آماری فیشر.....
۵۲	۷-۴-۲- روش های پیش بینی و ارزیابی و تفسیر مدل.....
۵۲	۱-۷-۴-۲- روش اعتبار سنجی تقاطعی گذاشتن یکی در بیرون(LOO-CV).....
۵۳	۲-۷-۴-۲- روش اعتبار سنجی تقاطعی گذاشتن تعدادی در بیرون(LMO-CV).....
۵۴	۳-۷-۴-۲- تصادفی کردن(Y).....
۵۵	۵-۲- پارامترهای آماری مورد استفاده در شبکه عصبی مصنوعی و ANFIS.....
۵۵	۱-۵-۲- مجذور میانگین مربعات خطا.....
۵۵	۲-۵-۲- خطای میانگین مطلق.....
۵۵	۳-۵-۲- ضریب تعیین(همبستگی).....
<b>فصل سوم: بحث و نتیجه گیری</b>	
۵۸	۱-۳- داده های تجربی.....
۵۸	۱-۱-۳- داده های ترکیبات دارای فعالیت ضد ایدز.....
۶۲	۲-۱-۳- داده های ترکیبات دارای فعالیت ضد آروماتاز(بیماری سرطان سینه).....
۶۶	۲-۳- ساخت مدل با استفاده از تکنیک رگرسیون خطی چندگانه(MLR).....
۶۶	۱-۲-۳- مدل پیشنهاد شده برای ترکیبات دسته اول.....

## ادامه فهرست مطالب

- ۶۷.....۳-۲-۲- مدل پیشنهاد شده برای ترکیبات دسته دوم.....
- ۶۸.....۳-۳- نمودارهای مدل های پیشنهاد شده رگرسیون خطی چندگانه.....
- ۶۹.....۳-۴- ساخت مدل با استفاده از شبکه عصبی مصنوعی (ANN).....
- ۶۹.....۳-۴-۱- نمودار مدل پیشنهادی برای ترکیبات دسته اول به همراه پارامترها.....
- ۷۰.....۳-۴-۲- نمودار مدل پیشنهادی برای ترکیبات دسته دوم به همراه پارامترها.....
- ۷۲.....۳-۵- ساخت مدل با استفاده از سیستم استنتاج تطبیقی عصبی - فازی.....
- ۷۲.....۳-۵-۱- نمودار مدل پیشنهادی برای ترکیبات دسته اول.....
- ۷۵.....۳-۵-۲- نمودار مدل پیشنهادی برای ترکیبات دسته دوم.....
- ۷۵.....۳-۶- مقایسه نتایج پیش بینی شده مدل های MLR، (ANN) و (ANFIS) برای دسته اول ترکیبات.....
- ۷۸.....۳-۷- معرفی توصیف گر مولکولی مورد استفاده در دسته اول ترکیبات.....
- ۸۰.....۳-۸- مقایسه نتایج پیش بینی شده مدل های MLR، (ANN) و (ANFIS) برای دسته دوم ترکیبات.....
- ۸۱.....۳-۹- معرفی توصیف گر مولکولی مورد استفاده در دسته دوم ترکیبات.....
- ۸۴.....۳-۸- نتیجه گیری.....
- ۸۶.....فهرست منابع.....

## فهرست شکل ها

صفحه	عنوان
۲۳	۱-۱- سیستم استنتاج فازی.....
۲۵	۲-۱- ساختار ANFIS, با دو متغیر ورودی و دو قانون.....
۳۰	۳-۱- فرمول ساختاری ایمیدازول.....
۳۰	۴-۱- فرمول ساختاری فتالازین.....
۳۵	۵-۱- فرمول ساختاری دو ایزومر تری آزول.....
۵۸	۱-۳- ساختار مولکولی مشتقات ترکیبات Imidazo(۲,۱-a)isoindoles.....
۵۹	۲-۳- ساختار مولکولی مشتقات ترکیبات Pyrimido(۲,۱-a)isoindoles.....
۵۹	۳-۳- ساختار مولکولی مشتقات ترکیبات Phthalazinones.....
۶۱	۴-۳- ساختار مولکولی مشتقات ترکیبات Benzalpthalides.....
	۵-۳- سنتز مشتقات ترکیبات Phthalazinones و Pyrimido(۲,۱-a)isoindoles و
۶۲	.....Imidazo(۲,۱-a)isoindoles
۶۲	۶-۳- سنتز مشتقات ترکیبات (۱ و ۲ و ۳) تری آزول.....
۶۴	۷-۳- سنتز ترکیبات مشتقات (۱، ۲، ۴ و ۱، ۲، ۳ و ۱، ۲، ۵) تری آزول و ۱، ۴- ایمیدازول.....
۶۵	۸-۳- سنتز مشتقات ترکیبات ایمیدازول و تری آزول.....
۶۵	۹-۳- سنتز مشتقات ترکیبات (۱ و ۲ و ۳) تری آزول.....
۶۹	۱۰-۳- نمودار تغییرات حداقل غلظت بازدارندگی پیش بینی شده برحسب ( $\mu\text{m}$ ) از طریق آنالیز رگرسیون خطی چندگانه در مقابل حداقل غلظت بازدارندگی تجربی برای ۴۲ مشتق مولکولی دارای فعالیت ضد ایدز.....
۶۹	۱۱-۳- نمودار تغییرات حداقل غلظت بازدارندگی پیش بینی شده برحسب ( $\mu\text{m}$ ) از طریق رگرسیون خطی چندگانه در مقابل حداقل غلظت بازدارندگی تجربی برای ۵۷ مشتق مولکولی دارای فعالیت ضد آروماتاز.....
۶۹	۱۲-۳- نمودار حداقل غلظت بازدارندگی $IC_{50}$ پیش بینی شده برحسب ( $\mu\text{m}$ ) از طریق ANN در مقابل حداقل غلظت بازدارندگی تجربی برای ۴۲ مشتق مولکولی دارای فعالیت ضد ایدز الف) مجموعه آزمایش، ب) مجموعه آموزشی و ج) مجموعه (آموزشی و آزمایشی).....

## ادامه فهرست شکل ها

- ۳-۱۳- نمودار حداقل غلظت بازدارندگی  $IC_{50}$  پیش بینی شده برحسب ( $\mu m$ ) از طریق ANN در مقابل حداقل غلظت بازدارندگی تجربی برای ۵۷ مشتق مولکولی دارای فعالیت ضد آروماتاز (بیماری سرطان سینه) الف) مجموعه آزمایش، ب) مجموعه آموزشی و ج) مجموعه (آموزشی و آزمایشی)..... ۷۱
- ۳-۱۴- ANFIS ساخته شده برای تعداد ۴۲ مولکول و تعداد ۵۷ مولکول..... ۷۲
- ۳-۱۵- نمودار حداقل غلظت بازدارندگی  $IC_{50}$  پیش بینی شده برحسب ( $\mu m$ ) از طریق ANFIS در مقابل حداقل غلظت بازدارندگی تجربی برای ۴۲ مشتق مولکولی دارای فعالیت ضد ایدز الف) مجموعه آزمایش، ب) مجموعه آموزشی و ج) مجموعه (آموزشی و آزمایشی)..... ۷۲
- ۳-۱۶- ساختار مدل فازی ساخته شده برای ۴۲ مشتق مولکولی..... ۷۳
- ۳-۱۷- توابع عضویت ورودی مدل فازی ساخته شده برای ۴۲ مشتق مولکولی..... ۷۴
- ۳-۱۸- قوانین حاکم بر مدل فازی ساخته شده برای ۴۲ مشتق مولکولی..... ۷۴
- ۳-۱۹- رویه ANFIS ساخته شده برای دو ورودی برای ۴۲ مشتق مولکولی..... ۷۵
- ۳-۲۰- نمودار حداقل غلظت بازدارندگی  $IC_{50}$  پیش بینی شده برحسب ( $\mu m$ ) از طریق ANFIS در مقابل حداقل غلظت بازدارندگی تجربی برای ۵۷ مشتق مولکولی دارای فعالیت ضد آروماتاز (بیماری سرطان سینه) الف) مجموعه آزمایش، ب) مجموعه آموزشی و ج) مجموعه (آموزشی و آزمایشی)..... ۷۶
- ۳-۲۱- ساختار مدل فازی ساخته شده برای ۵۷ مشتق مولکولی..... ۷۷
- ۳-۲۲- توابع عضویت ورودی مدل فازی ساخته شده برای ۵۷ مشتق مولکولی..... ۷۷
- ۳-۲۳- قوانین حاکم بر مدل فازی ساخته شده برای ۵۷ مشتق مولکولی..... ۷۸
- ۳-۲۴- رویه ANFIS ساخته شده برای دو ورودی در ۵۷ مشتق مولکولی..... ۷۸

## فهرست جدول ها

عنوان	صفحه
۱-۱- خصوصیات فیزیکی ایمیدازول.....	۳۰
۲-۱- خصوصیات فیزیکی فتالازین.....	۳۱
۳-۱- خصوصیات فیزیکی تری آزول.....	۳۵
۴-۱- تعدادی از مقالات تحقیقاتی در زمینه QSAR در بررسی فعالیت ضد ایدز.....	۳۶
۵-۱- تعدادی از مقالات تحقیقاتی در زمینه QSAR در بررسی ضد آروماتاز(سرطان سینه).....	۴۰
۱-۳- مجموعه ۶ عضوی مشتقات ترکیبات Imidazo(۲,۱-a)isoindoles به همراه حداقل غلظت بازدارندگی تجربی.....	۵۸
۲-۳- مجموعه ۵ عضوی مشتقات ترکیبات Pyrimido(۲,۱-a)isoindoles به همراه حداقل غلظت بازدارندگی تجربی.....	۵۹
۳-۳- مجموعه ۲۰ عضوی مشتقات ترکیبات Phthalazinones به همراه حداقل غلظت بازدارندگی تجربی.....	۶۰
۴-۳- مجموعه ۱۱ عضوی مشتقات ترکیبات Benzalpthalides به همراه حداقل غلظت بازدارندگی تجربی.....	۶۱
۵-۳- ساختار شیمیایی به همراه حداقل غلظت بازدارندگی تجربی مشتقات ترکیبات (۱ و ۲ و ۳) تری آزول.....	۶۳
۶-۳- ساختار شیمیایی به همراه حداقل غلظت بازدارندگی تجربی مشتقات ترکیبات (۱، ۲، ۴، ۱ و ۲، ۳ و ۴) تری آزول و ایمیدازول.....	۶۴
۷-۳- ساختار شیمیایی به همراه حداقل غلظت بازدارندگی تجربی مشتقات ترکیبات ایمیدازول و تری آزول.....	۶۵
۸-۳- ساختار شیمیایی و حداقل غلظت بازدارنده تجربی مشتقات (۱ و ۲ و ۳) تری آزول.....	۶۶
۹-۳- ماتریس ضرایب همبستگی متغیر های انتخاب شده ترکیبات دارای فعالیت ضد ایدز.....	۶۷
۱۰-۳- ماتریس ضرایب همبستگی متغیر های انتخاب شده ترکیبات دارای فعالیت ضد آروماتاز.....	۶۷

ادامه فهرست جدول ها

- ۳-۱۱- مقایسه نتایج پیش بینی شده  $IC_{50}$  برحسب ( $\mu\text{m}$ ) با سه روش MLR ، ANN و ANFIS برای مشتق مولکولی دارای فعالیت ضد ایدز..... ۷۹
- ۳-۱۲- شاخص های مورد استفاده برای مدل سازی روابط ساختار- فعالیت ۴۲ مشتق مولکولی دارای فعالیت ضد ایدز..... ۸۱
- ۳-۱۳- مقایسه نتایج پیش بینی شده  $IC_{50}$  برحسب ( $\mu\text{m}$ ) با سه روش MLR ، ANN و ANFIS برای مشتق مولکولی دارای فعالیت ضد آروماتاز(بیماری سرطان سینه)..... ۸۲
- ۳-۱۴- شاخص های مورد استفاده برای مدل سازی روابط ساختار- فعالیت ۵۷ مشتق مولکولی دارای فعالیت ضد آروماتاز(بیماری سرطان سینه)..... ۸۴
- ۳-۱۵- نتایج مقادیر آماری برای ۴۲ مشتق مولکولی مربوط به فعالیت ضد ایدز..... ۸۴
- ۳-۱۶- نتایج مقادیر آماری برای ۵۷ مشتق مولکولی مربوط به مهارکننده آروماتاز (سرطان سینه)..... ۸۴

# فصل اول

## مقدمه

## ۱-۱ مقدمه

علم کمومتریکس<sup>۱</sup> با کمک دو روش مدل یابی مولکولی و آماری جهت بهینه سازی فرآیندها و آنالیز داده ها در یک شاخه از علم شیمی می پردازد، بطوریکه در روش مدل یابی داده های زیادی حاصل می گردد ولی در روش آماری تعدادی از فاکتورها که فضای زیادی از پارامترها را پوشش می دهد، انتخاب شده. این علم با سرعت فوق العاده ای در حال توسعه می باشد که دلیل آن می تواند دستیابی سریع به داده های مورد نیاز در علوم دیگر باشد، چرا که امروزه برای مقابله با بیماری های نوظهوری که موجب آشفته گی جوامع بشری شده اند، دستیابی و طراحی داروهای نسل جدید با عوارض جانبی کمتر و خواص بهبود یافته در زمان محدود بسیار لازم و ضروری به نظر می رسد. برای دستیابی به انجام یک فرآیند بهینه در محصولات، اطلاعات مربوط به ارزیابی ساختار مواد شیمیایی مورد نیاز بوده که یکی از بهترین راه حل ها جهت کاهش حجم اطلاعات یا کاهش نتایج تجزیه ای که ضریب اطمینان بالا داشته، استفاده از روشهای کمومتریکس می باشد. همزمان با توسعه علم شیمی محاسباتی، استفاده از روش های کمومتریکس آغاز شد و با اصول ریاضی و آمار نه تنها در شیمی تجزیه، بلکه در گستره وسیعی از شاخه های علوم گام موثری برداشته است [۱-۲].

۱-۲- تاریخچه QSAR<sup>۲</sup>

اولین تلاشها برای ایجاد ارتباط بین رفتار فیزیکی- شیمیایی یک ترکیب با ساختار آن حدوداً از ۲۵۰۰ سال پیش با افلاطون<sup>۳</sup> آغاز شد که یک دانش تبدیل به گاز در یک محلول آبی را توصیف کرد. مندلیف<sup>۴</sup>، که جدول تناوبی را در سال ۱۸۶۹ ساخت، را می توان یکی از قدیمی ترین افراد در زمینه

---

<sup>۱</sup> Chemometrics

<sup>۲</sup> Quantitative Structure Activity Relationship

<sup>۳</sup>Plato

<sup>۴</sup>Mendeleev



QSAR دانست. کراس<sup>۱</sup> در سال ۱۸۶۳ مشاهده کرد که افزایش سمیت الکل در پستانداران با کاهش انحلال پذیری آن در آب، مرتبط است و پس از او در سال ۱۹۶۸ به رابطه بین فعالیت های فیزیولوژیکی و ساختار شیمیایی ترکیبات پی بردند [۶].

ریچاردسون<sup>۲</sup> نیز در طول همان سالها (۱۸۶۸) نشان داد که میزان سمیت اترها و الکل ها نسبت معکوس با حلالیت آنها در آب دارد.

اورتون و مایر<sup>۳</sup> به ترتیب در سالهای ۱۸۹۷ و ۱۸۹۹ به طور مستقل نشان دادند که سمیت ترکیبات آلی به مقدار ضریب تقسیم<sup>۴</sup> آنها بین فازهای بیولوژیکی چربی دوستی<sup>۵</sup> و آب دوستی<sup>۶</sup> وابسته است. پس از این دو، دانشمند دیگری برای تخمین ضریب تقسیم از ساختار شیمیایی استفاده کرد، ولی فرگوسن<sup>۷</sup> ضمن اعلام یک اصل علمی برای سمیت ترکیبات با توجه به گروههای عاملی آنها، اولین فرمول بندی ریاضی را بیان کرد. فرگوسن توانست با استفاده از خواص فیزیکی مانند انحلال پذیری در آب، توزیع بین فازها، موینگی و فشار بخار دلیل آن را پیگیری نماید [۷-۸].

### ۱-۲-۱- اصول اساسی

روش QSAR بر نمایش عددی ساختار شیمیایی بنا شده که هدف از آن، فهم چگونگی تاثیر اختلاف ساختاری مربوط به یک گروه از ترکیبات، بر روی فعالیت بیولوژیکی می باشد. فرض اصلی و عمده این است که یک گروه از فاکتورها بر اتفاقات موجود در سیستم بیولوژیکی حاکم هستند که بوسیله شاخصهایی نمایش داده شده و فعالیتهای بیولوژیکی از طریق همان مکانیسم توضیح داده می شوند.

<sup>۱</sup> Cross

<sup>۲</sup> Richardson

<sup>۳</sup> Meyer and Overton

<sup>۴</sup> Partition Coefficient

<sup>۵</sup> Lipophilicity

<sup>۶</sup> Hydrophobicity

<sup>۷</sup> Ferguson

بنابراین QSAR تلاش می‌کند که چگونگی تاثیر خواص یک مولکول بر فعالیت آن را یافته و همچنین تشخیص دهد که چگونه می‌توان یک ترکیب را اصلاح نمود تا خواص آن افزایش یابد. از این رو انتظار می‌رود که برای یک سری از مولکولهای فعال بیولوژیکی، هر اختلاف سیستماتیک در ساختار شیمیایی (از یک تا چندین عدد) بتواند در یک اختلاف مشابه که متناسب با یک پاسخ بیولوژیکی است، منعکس شود.

یک نمونه از QSAR بر این فرض قرار گرفته که یک ارتباط اساسی بین ساختار مولکولی و فعالیت بیولوژیکی وجود دارد که از اختلاف سیستماتیک ناشی شده و همچنین فرض می‌شود که توصیفات فیزیکی - شیمیایی چند متغیری در یک مجموعه از ترکیبات می‌تواند شباهت آنها را آشکار کند. از ساختار مولکولی ماده می‌توان همه خواص بیولوژیکی و فیزیکی - شیمیایی مربوط به یک ماده شیمیایی را محاسبه کرد، که به کمک شاخصهای گوناگون در یک فرم عددی به شکل رمزی در آمده است.

البته اهمیت دارد که تفاوت بین همبستگی و نسبت علت - معلول بیان شود. یک ارتباط رضایت بخش از QSAR بدین معنی نیست که شاخصهای مخصوص سبب یک عمل موثر برای یک ترکیب شوند. نبود یک دلیل واضح برای یک علت می‌تواند با اطلاعات اضافی درباره هدایت مکانیسم‌های گوناگون به سمت فعالیتهای بیولوژیکی تکمیل گردد. در حقیقت QSAR بر اصلی مشهور به نام تشابه<sup>۱</sup> بنا شده است که بیان می‌کند ترکیبات مشابه فعالیتهای مشابهی دارند. به علاوه در کنار اصل فوق الذکر (اصل تشابه) اصلی بنام اصل همسایگی<sup>۲</sup> نیز وجود دارد. این اصل در یک قاعده کلی بیان می‌کند که محل مولکولها در همان ناحیه از فضا تعیین می‌شود که شاخصهای موجود با فعالیتهای مشابه مشخص می‌شوند.

<sup>۱</sup> Similarity principle

<sup>۲</sup> Neighbourhood principle

۱-۲-۲- روابط انرژی آزاد خطی<sup>۱</sup>

در اواسط دهه ۱۹۳۰، هامت<sup>۲</sup> مشاهده نمود که افزایش در استخلافها در یک حلقه بنزوئیک اسید، یک اثر کمی و منظم بر روی ثابت تفکیک دارد. او همچنین اثر مشابهی را روی تفکیک اسیدهای آلی دیگر و بازها مشاهده کرد. از مشاهدات تجربی، روابط خطی را بدست آورد که به معادلات هامت<sup>۳</sup> مشهور هستند:

$$\text{Log} \frac{K}{K_0} = \rho\sigma \quad (1-1)$$

$\rho$  شیب است که متناسب با ثابت واکنش بوده و اثر استخلاف را روی معادله و روی تعادل بنزوئیک اسید بیان می‌کند و  $\sigma$  پارامتری است که خواص الکترونی را در استخلافهای آروماتیک توضیح می‌دهد. این روابط به عنوان روابط انرژی آزاد خطی نام دارند زیرا معادلات مربوط به انرژی آزاد ( $G$ ) و ثابت تعادل ( $K$ ) و ثابت سرعت ( $k$ )، را یادآوری می‌کند چرا که همانطور که می‌دانیم روابط مربوط به انرژی آزاد به اندازه گیریهای غلظت با روابط لگاریتمی، مربوط می‌شوند:

$$\Delta G^{\circ} = -RT \ln K \quad (2-1)$$

این تحقیق اولین تحقیقی بود که در آن مولکولها دسته بندی شده و فعالیت آنها از روی اجزاء آنها به جای استناد به کلیاتشان، توضیح داده می‌شد.

<sup>۱</sup> Linear Free-Energy Relationship(LFER)

<sup>۲</sup> Hammett

<sup>۳</sup>Hammett equation

تافت<sup>۱</sup> بر روی اثرات فضایی<sup>۲</sup> (Es) مطالعه کرد که دقیق به معادله هامت وابسته بود و معادله‌ای از آن مشتق شد که به معادله تافت<sup>۳</sup> مشهور است:

$$\text{Log} \frac{k}{k_0} = \rho^* \sigma^* + \delta E_s \quad (۳-۱)$$

در همین مسیر سواين<sup>۴</sup> نیز فعالیتهایی انجام داد. او اثرات میدان<sup>۵</sup> و رزونانس<sup>۶</sup> را بررسی نموده و تغییرات واکنش پذیری یک ماده الکتروفیل معین را نسبت به یک سری از واکنشگرهای هسته دوست که از رابطه انرژی آزاد خطی مشتق شده‌اند بررسی و به معادله‌ای که به معادله سواين-اسکات<sup>۷</sup> معروف است دست یافت:

$$\text{Log} \frac{k}{k_0} = s n \quad (۴-۱)$$

که  $n$  معیاری از خاصیت هسته دوستی واکنشگر و  $s$  معیاری از حساسیت به هسته دوستی مربوط به مشخصات یک واکنشگر در ماده مورد عمل می‌باشد. او همچنین معادله سواين-لوپتون<sup>۸</sup> را نتیجه گرفت و به پارامترهای دوتایی در آنالیز روابط مربوط به اثر استخلافات دست یافت که شامل ثابت میدان  $F$  و ثابت رزونانس  $R$  بود.

فری<sup>۹</sup> و ویلسون<sup>۱۰</sup> یک مولکول را به روش متفاوتی نسبت به هامت، طبقه بندی کردند. آنها ادعا کردند که فعالیتهای بیولوژیکی یک مجموعه مولکولی، می‌تواند به ترکیبی از تعداد، انواع و موقعیت گروههای استخلافی وابسته باشد.

<sup>۱</sup> Taft

<sup>۲</sup> Steric effect (Es)

<sup>۳</sup> Taft equation

<sup>۴</sup> Swain

<sup>۵</sup> field

<sup>۶</sup> Resonance

<sup>۷</sup> Swain-Scott equation

<sup>۸</sup> Swain-Lupton equation

<sup>۹</sup> Free

<sup>۱۰</sup> Wilson