



***VIRTUAL SCREENING CALON INHIBITOR  
TOPOISOMERASE-1 MENGGUNAKAN QSAR***

**SKRIPSI**

Oleh  
**Rozin Rozaina**  
**171810301015**

**JURUSAN KIMIA  
FAKULTAS MATEMATIKA DAN ILMU PENGE TAHUAN ALAM  
UNIVERSITAS JEMBER**

**2024**



***VIRTUAL SCREENING CALON INHIBITOR  
TOPOISOMERASE-1 MENGGUNAKAN QSAR***

**SKRIPSI**

Diajukan untuk melengkapi tugas akhir dan memenuhi salah satu syarat untuk menyelesaikan Program Studi Kimia (S1) dan mencapai gelar Sarjana Sains

Oleh

**Rozin Rozaina**

**171810301015**

**JURUSAN KIMIA**

**FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM**

**UNIVERSITAS JEMBER**

**2024**

## PERSEMBAHAN

Puji syukur atas berkat dan rahmat Allah S.W.T, serta doa dari orang-orang yang tercinta karya tulis ini dapat terselesaikan. Dengan mengucapkan syukur Alhamdulillah saya persembahkan karya tulis ini kepada:

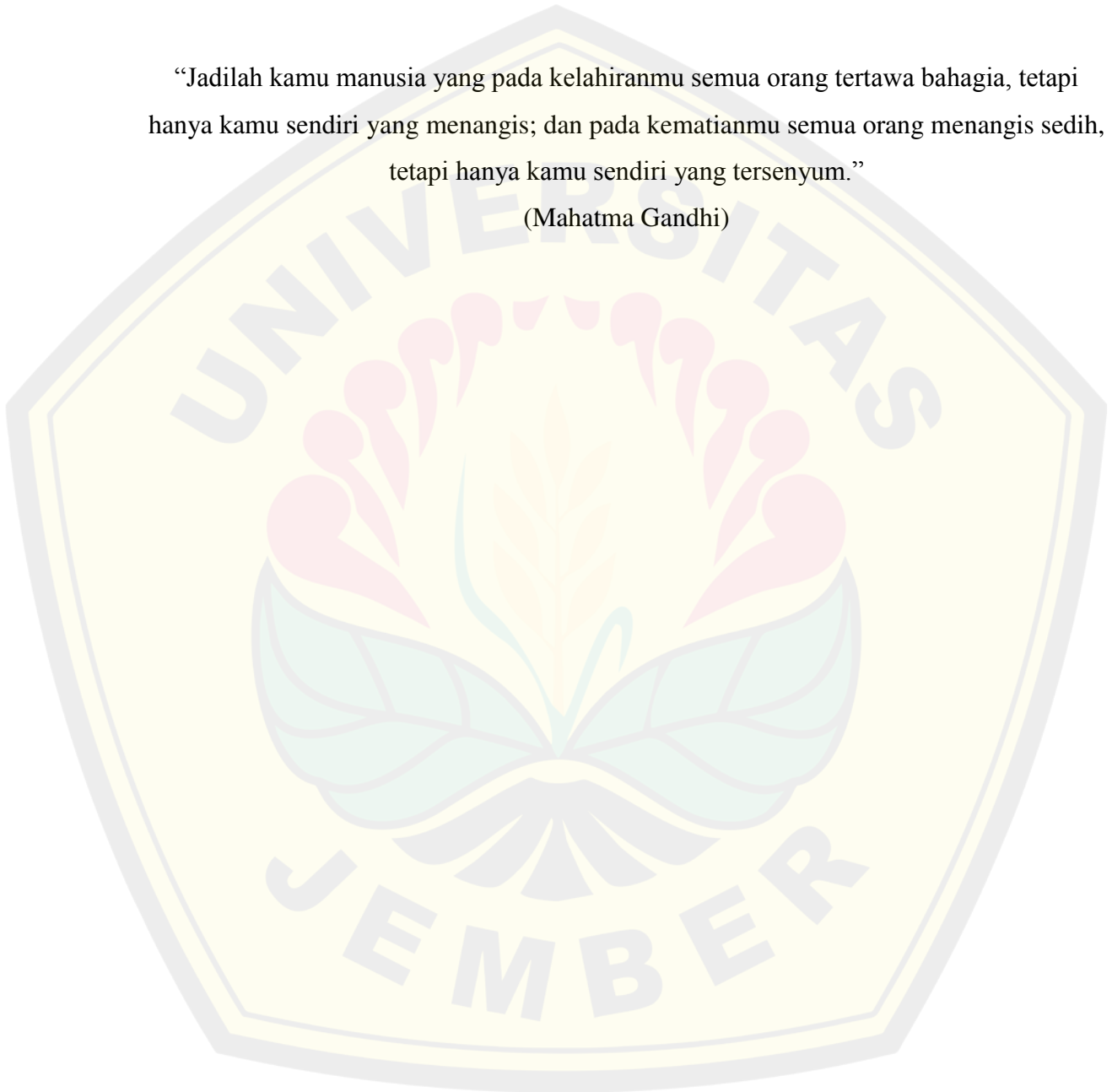
1. Ibunda Kartini dan Ayahanda Salaman atas segala ketulusan cinta, kasih sayang, dan perhatian yang diberikan kepada saya, kerja keras, pengorbanan yang begitu besar, dukungan, dan arahan, serta doa yang tiada terputus demi keberhasilan dan kesuksesan saya;
2. Keempat kakak kandung saya Sufyan Riadi, Ali Imron, Ibtatul awalia dan Shuhufam Muthahharah beserta keluarga besar yang selalu memberikan semangat saat saya mulai lelah dan malas belajar;
3. Bapak/Ibu guru sejak SD hingga Perguruan Tinggi yang saya hormati serta semua orang yang telah memberikan pengetahuan serta pengalaman dengan penuh keikhlasan;
4. Almamater yang saya cintai Jurusan Kimia, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Jember.

**MOTTO**

“Sebaik-baiknya manusia ialah yang bermanfaat bagi manusia lainnya.”

“Jadilah kamu manusia yang pada kelahiranmu semua orang tertawa bahagia, tetapi hanya kamu sendiri yang menangis; dan pada kematianmu semua orang menangis sedih, tetapi hanya kamu sendiri yang tersenyum.”

(Mahatma Gandhi)



**PERNYATAAN**

Saya yang bertanda tangan dibawah ini :

Nama : Rozin Rozaina

NIM : 171810301015

Menyatakan dengan sesungguhnya bahwa karya tulis ilmiah yang berjudul “*Virtual Screening* Calon Inhibitor Topoisomerase-1 Menggunakan QSAR” merupakan benar – benar hasil karya saya sendiri, kecuali kutipan yang telah saya sebutkan sumbernya dan belum pernah diajukan pada institusi manapun, serta bukan hasil jiplakan. Saya bertanggung jawab penuh atas keabsahan dan kebenaran isinya sesuai dengan sikap ilmu yang harus dijunjung tinggi.

Demikian pernyataan ini saya buat dengan sebenarnya, tanpa adanya tekanan dan paksaan dari pihak manapun serta bersedia mendapatkan sanksi akademik jika ternyata dari pernyataan ini tidak benar.

Jember, 29 Juli 2024

Yang menyatakan,

Rozin Rozaina

171810301015

**SKRIPSI**

***VIRTUAL SCREENING CALON INHIBITOR TOPOISOMERASE-1  
MENGUNAKAN QSAR***

Oleh

**Rozin Rozaina**

**NIM 171810301015**

Pembimbing

Dosen Pembimbing Utama : Drs. Sudarko, Ph.D.

Dosen Pembimbing Anggota : Drs. Zulfikar, Ph.D.

**PENGESAHAN**

Skripsi berjudul “*Virtual Screening* Calon Inhibitor Topoisomerase-1 Menggunakan QSAR” karya Rozin Rozaina telah diuji dan disahkan pada :

Hari, Tanggal : Senin, 29 Juli 2024

Tempat : Jurusan Kimia Fakultas Matematika dan Ilmu Pengetahuan Alam  
Universitas Jember

Tim Penguji,

Ketua

Anggota I

Drs. Sudarko, Ph. D.

NIP. 196903121992031002

Drs. Zulfikar, Ph. D.

NIP. 196310121987021001

Anggota II

Anggota III

Dr. Muhammad Reza, S.Si., M.Si

NIP. 199204132022031007

Tanti Haryati S.Si, M.Si.

NIP. 198010292005012002

## RINGKASAN

***Virtual Screening Calon Inhibitor Topoisomerase-1 Menggunakan QSAR***; Rozin Rozaina, 171810301015; 2024; 33 halaman; Jurusan Kimia Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Topoisomerase-1 merupakan protein target yang tergolong dalam kelompok regulator apoptosis sel. Protein Topoisomerase-1 dapat menghambat proses apoptosis pada sel yang rusak dan membuat sel yang telah rusak tetap mengalami regulasi sel dan mengakibatkan terbentuknya sel kanker. Inhibitor sebagai penghambat dibutuhkan untuk menghambat protein Topoisomerase-1 agar proses apoptosis tetap berlangsung. Prediksi inhibitor untuk menghambat Topoisomerase-1 dapat dilakukan dengan menggunakan pemodelan QSAR (*Quantitative Structure-Activity Relationship*) berbasis *virtual screening*.

*Virtual screening* diawali dengan pemodelan QSAR, dimana komputer dapat mempelajari pola data yang ada untuk menghasilkan model terbaik. Model dibuat dengan menggunakan dataset dari *dataset* yang diperlukan sebanyak 429 senyawa yang diambil dari *database* ChEMBL dengan kode ChEMBL1781 dan UniPort ID11387. Kemudian didapatkan data *training* sebanyak 300 senyawa dan data *testing* sebanyak 129 senyawa. Model terbaik yang didapatkan yaitu model yang memiliki *Hidden Layer* (L1, L2, dan L3), dan *batch size* secara berurutan yaitu 8; 16; 32; dan 8; 1024; 2048 dengan *nb\_epoch* 50 sehingga diperoleh model terbaik dengan nilai  $R_2$  sebesar 0,66. Model terbaik tersebut kemudian digunakan untuk proses *virtual screening* pada *database* ZINC20 (<https://zinc20.docking.org/>) dengan jumlah senyawa sebanyak 481 juta senyawa. Hasil *virtual screening* dipilih senyawa dengan nilai  $pIC_{50}$ -nya diatas 5.8 berdasarkan rata-rata obat yang beredar yaitu Topotecan diperoleh sebanyak 335 senyawa. Kemudian dikelompokkan (*clustering*) menggunakan metode *clustering-Butina* yaitu



mengelompokkan senyawa berdasarkan kemiripannya yaitu berdasarkan kemiripan struktur. Jumlah *cluster* yang dihasilkan dari proses *clustering* yaitu 50 *cluster*. *Cluster* tersebut diwakili oleh senyawa dengan nilai  $pIC_{50}$  terbesar dari masing– masing *cluster*. Hasil dari proses *clustering* kemudian dilakukan proses perhitungan nilai *similarity*. Perhitungan nilai *similarity* yang dilakukan pada antar *cluster* dan senyawa yang berada dalam satu *cluster*. Senyawa tersebut kemudian dilakukan uji in-vitro sebagai calonobat baru penyakit kanker lambung menggunakan uji ADMET pKCSM (Absorption, Distribution, Metabolism, Excretion - Toxicity Predictive Model) bertujuan untuk mengevaluasi bagaimana suatu senyawa atau obat diabsorpsi, didistribusikan, dimetabolisme, diekskresikan, serta untuk memprediksi potensi toksisitasnya. Ini penting dalam pengembangan obat untuk memahami bagaimana obat akan berperilaku di dalam tubuh manusia, sehingga dapat menentukan dosis yang efektif dan aman, serta memperkirakan risiko efek samping yang mungkin terjadi. Hasil uji ADMET menggunakan ADMET-pKCSM menghasilkan 10 senyawa yang lolos uji. Berdasarkan keseluruhan proses screening didapatkan 10 calon inhibitor protein Topoisomerase-1 yang dapat direkomendasikan dan dilakukan penelitian lebih lanjut.

## PRAKATA

Puji syukur kehadirat Allah *Subhanahu wa Ta'ala* atas segala rahmat dan karunia-Nya sehingga penulis dapat menyelesaikan karya tulis ilmiah, berupa skripsi berjudul “*Virtual Screening* Calon Inhibitor Topoisomerase-1 Menggunakan QSAR”. Skripsi ini disusun untuk memenuhi salah satu syarat dalam menyelesaikan pendidikan Strata Satu (S1) pada Jurusan Kimia Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Jember.

Penyusunan skripsi ini tidak lepas dari bantuan berbagai pihak, oleh karena itu penulis ingin menyampaikan ucapan terima kasih kepada :

1. Dekan FMIPA Universitas Jember, Drs. Achmad Sjaifullah, M.Sc., Ph.D.;
2. Ketua Jurusan Kimia FMIPA Universitas Jember, Prof. Dr. Anak Agung Istri Ratnadewi, S.Si., M.Si.;
3. Dosen Pembimbing Utama Drs. Sudarko, Ph.D. dan Dosen Pembimbing Anggota, Drs. Zulfikar, Ph.D. yang telah memberikan bimbingan, meluangkan waktu dan memberikan arahan dengan sabar dalam proses penyusunan skripsi ini;
4. Dosen Penguji Utama, Dr. Muhammad Reza, S.Si., M.Si. dan Dosen Penguji Anggota, Tanti Haryati, S.Si., M.Si. yang telah mengevaluasi isi karya tulis dan memberikan saran atau masukan selama proses penyelesaian tugas akhir;
5. Dosen Pembimbing Akademik, Suwardiyanto., M.Si., Ph.D, yang telah membimbing dan memberikan arahan selama menjalani masa perkuliahan
6. Seluruh Bapak Ibu dosen Jurusan Kimia FMIPA Universitas Jember yang telah memberikan banyak ilmu yang bermanfaat selama masa perkuliahan;
7. Seluruh teknisi laboratorium dan admin Jurusan Kimia FMIPA Universitas Jember yang telah banyak membantu
8. Orang tua dan saudara yang selalu mendukung dan mendoakan yang terbaik;
9. Tim Riset Kimia Komputasi ( Kirana, Zihana, Deby, Chelsy, Puput dan Bayu ) yang telah berjuang Bersama dalam penelitian;

10. Semua pihak yang tidak dapat disebutkan satu persatu yang turut membantu dalam penyelesaian skripsi ini.

Penulis meminta maaf atas kekeliruan yang mungkin terdapat dalam penulisan skripsi ini. Penulis menerima segala kritik dan saran untuk menyempurnakan penulisan skripsi ini. Semoga hasil penelitian ini dapat dijadikan referensi bagi penelitian selanjutnya dan dapat bermanfaat bagi orang banyak.

Jember, 29 Juli 2024

Penulis



**DAFTAR ISI**

HALAMAN JUDUL .....	i
HALAMAN SAMPUL .....	ii
HALAMAN PERSEMBAHAN .....	iii
HALAMAN MOTTO .....	iv
HALAMAN PERNYATAAN .....	v
HALAMAN PEMBIMBING.....	vi
HALAMAN PENGESAHAN .....	vii
RINGKASAN.....	viii
PRAKATA.....	x
DAFTAR ISI.....	xi
DAFTAR TABEL.....	xiv
DAFTAR GAMBAR.....	xv
DAFTAR LAMPIRAN.....	xvi
DAFTAR ISTILAH.....	xvii
BAB 1. PENDAHULUAN .....	1
1.1 Latar Belakang .....	1
1.2 Rumusan Masalah .....	3
1.3 Tujuan Penelitian .....	4
1.4 Batasan Masalah.....	4
1.5 Manfaat Penelitian .....	5
BAB 2. TINJAUAN PUSTAKA .....	6
2.1 Topoisomerase-1 .....	6

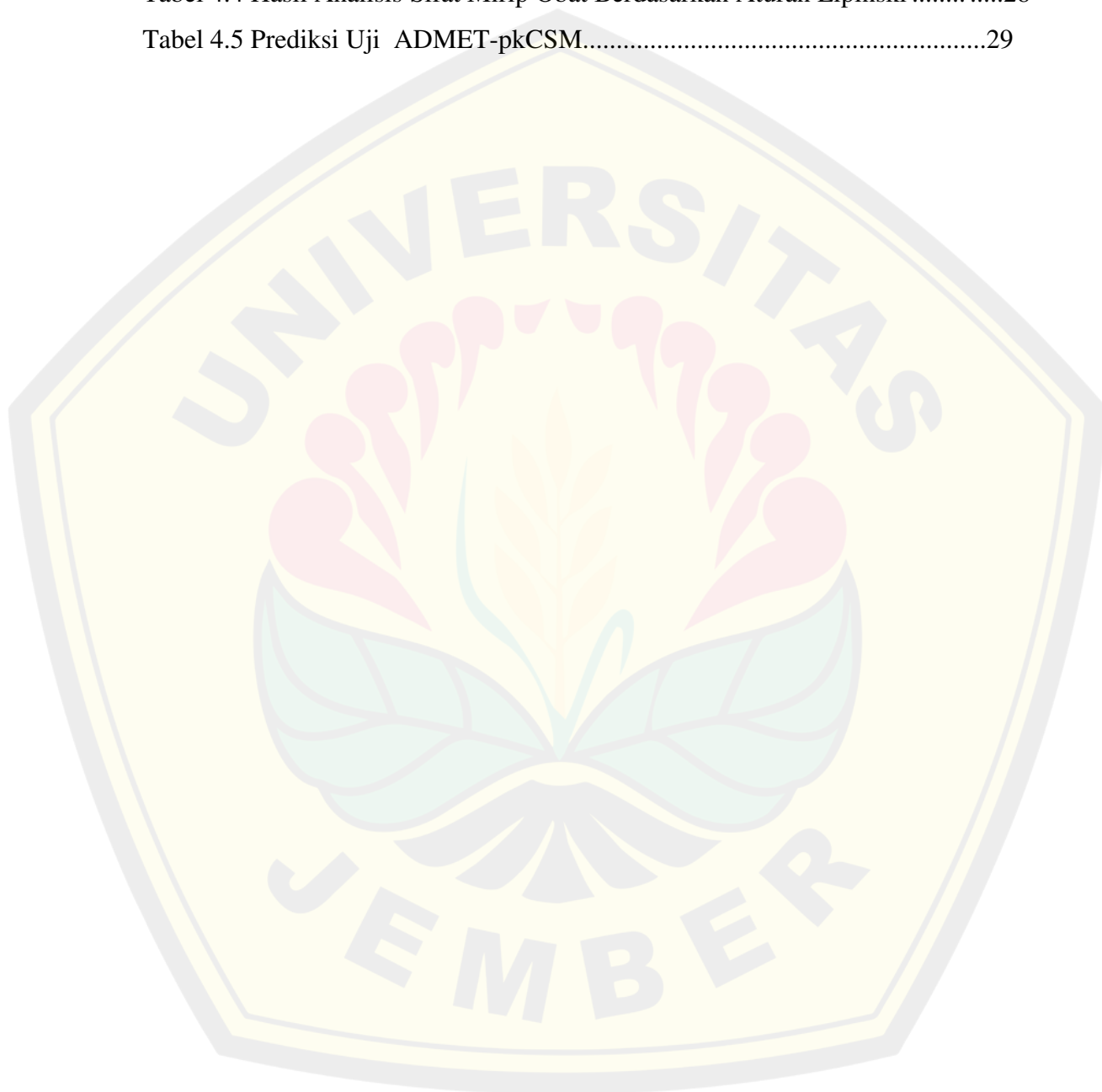
2.2	Kanker Leher Rahim (Kanker Serviks).....	8
2.3	Sel Hela .....	8
2.4	<i>Desain Obat In Sillico</i> .....	9
2.5	<i>Quantitative Structure Activity Relationship (QSAR)</i> .....	10
2.6	Pengobatan Kanker Serviks .....	10
2.7	Inhibitor.....	11
2.8	ChEMBL .....	11
2.9	ZINC .....	12
2.10	<i>Moleculer Fingerprints</i> .....	12
2.11	<i>SMILES (Simplified Molecular-Input Line-Entry)</i> .....	12
2.12	<i>Absorption, Distribution, Metabolism, Excretion and Toxicity (ADMET)</i> .....	13
<b>BAB 3. METODELOGI PENELITIAN</b> .....		15
3.1	Tempat dan Waktu Penelitian .....	15
3.2	Alat Penelitian.....	15
3.3	Diagram Alir Penelitian .....	16
3.4	Prosedure Penelitian.....	17
3.4.1	Preparasi data .....	17
3.4.2	Pemodelan menggunakan QSAR.....	18
3.4.3	Skrining virtual.....	18
<b>BAB 4. HASIL DAN PEMBAHASAN</b> .....		19
4.1	Pemodelan QSAR Menggunakan <i>Machine Learning</i> .....	19
4.2	<i>Virtual Screening</i> Calon Inhibitor Topoisomerase-1 .....	21
<b>BAB 5. PENUTUP</b> .....		31
5.1	Kesimpulan .....	31

5.2 Saran .....	31
DAFTAR PUSTAKA.....	33
LAMPIRAN .....	39



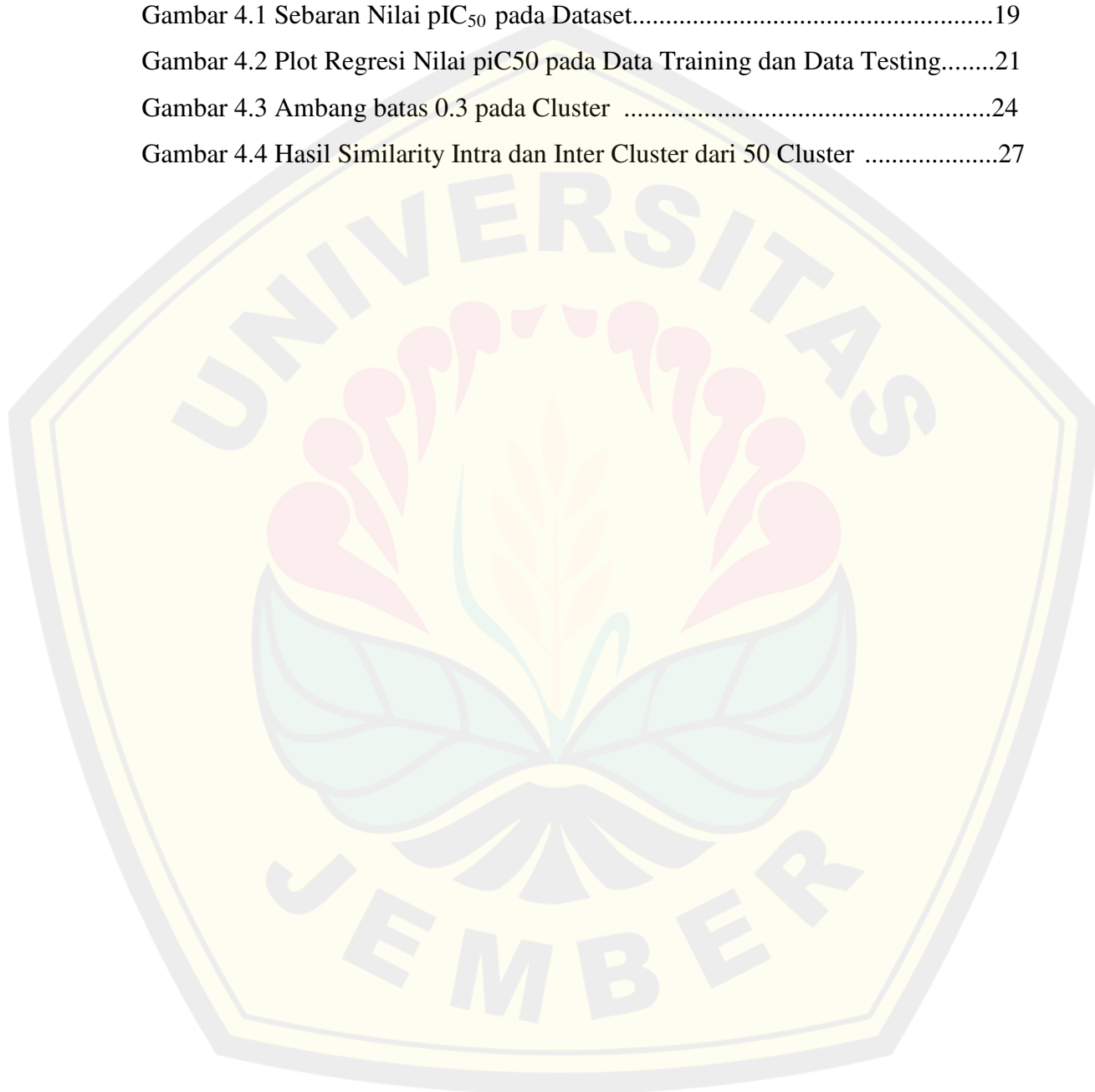
**DAFTAR TABEL**

Tabel 4.1 Nilai $pIC_{50}$ Hasil Prediksi Beberapa Senyawa.....	22
Tabel 4.2 Jumlah Senyawa pada Masing-Masing <i>Cluster</i> .....	24
Tabel 4.3 Senyawa yang Mewakili Masing-Masing <i>Cluster</i> .....	25
Tabel 4.4 Hasil Analisis Sifat Mirip Obat Berdasarkan Aturan Lipinski .....	28
Tabel 4.5 Prediksi Uji ADMET-pkCSM.....	29



**DAFTAR GAMBAR**

Gambar 2.1 Struktur Kristal Topoisomerase-1.....	5
Gambar 2.2 Proses Virtual Screening Berbasis QSAR.....	9
Gambar 4.1 Sebaran Nilai $pIC_{50}$ pada Dataset.....	19
Gambar 4.2 Plot Regresi Nilai $pIC_{50}$ pada Data Training dan Data Testing.....	21
Gambar 4.3 Ambang batas 0.3 pada Cluster .....	24
Gambar 4.4 Hasil Similarity Intra dan Inter Cluster dari 50 Cluster .....	27





## DAFTAR ISTILAH

**A**

*Absorption* Proses penyerapan senyawa obat ke dalam aliran darah

**B**

*Batch Size* Jumlah data yang diinput pada proses *training*

**C**

ChEMBL Salah satu database yang berisi senyawa kimia dan aktivitasnya

**D**

Database Kumpulan data yang digunakan untuk membuat model

*Distribution* Penyebaran senyawa obat di seluruh tubuh

*Drop out rate* Jumlah yang digunakan untuk pencegahan *overfitting* pada proses *training*

**E**

*Excretion* Proses pengeluaran senyawa obat dari dalam tubuh melalui urine atau empedu

**F**

*Fingerprints* Representasi sifat molekul dalam bentuk urutan bit

**I**

Inhibitor protein Senyawa yang berfungsi sebagai penghambat kinerja protein

## L

*Layer* Neuron yang saling terhubung pada proses *training* untuk mempelajari data

*Learning Rate* Kecepatan pada proses *training*

*Lipinski's* Aturan yang menjadi acuan pada kemampuan tubuh untuk menyerap suatu senyawa

## M

*Machine Learning* Metode pada *computer* yang digunakan untuk mempelajari data melalui pola untuk menghasilkan suatu model

*Metabolism* Proses pada senyawa obat dirubah oleh enzim di dalam tubuh menjadi metabolit lain

Model Hasil dari pembelajaran pola data pada *machine learning*

## P

$pIC_{50}$  Nilai  $-\log_{10}$  dari konsentrasi senyawa yang dibutuhkan untuk menghambat aktivitas protein sebanyak 50%

## Q

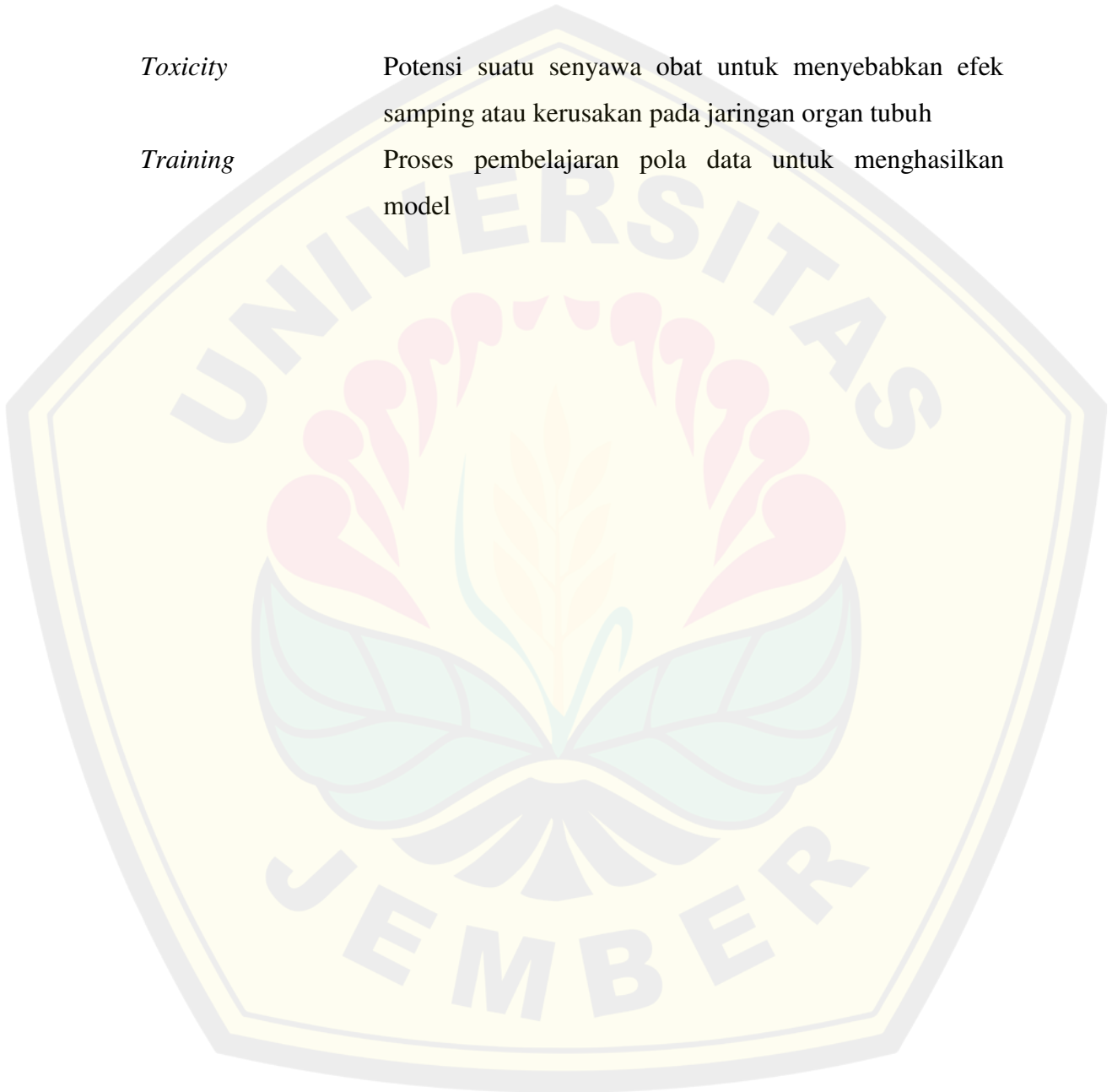
QSAR Metode pada pengembangan obat melalui hubungan kuantitatif struktur dan aktivitas

## S

SMILES Representasi struktur molekul dalam bentuk baris yang berisi angka, huruf dan tanda baca

**T**

- Testing* Proses pengujian model hasil *training* yang sudah melalui *validations*
- Topoisomerase-1 Salah satu jenis protein yang terlibat dalam regulasi pertumbuhan dan proliferasi sel
- Toxicity* Potensi suatu senyawa obat untuk menyebabkan efek samping atau kerusakan pada jaringan organ tubuh
- Training* Proses pembelajaran pola data untuk menghasilkan model



## BAB 1. PENDAHULUAN

### 1.1 Latar Belakang

Kanker merupakan penyebab kematian utama terbesar baik di negara maju maupun negara berkembang. Salah satu jenis keganasan yang di temukan di Indonesia adalah kanker serviks. Penyakit kanker serviks (sel HeLa) merupakan pertumbuhan dari suatu kelompok sel yang mengalami mutasi genetik sehingga terjadi pembelahan sel yang tidak terkendali. Menurut data *Global Burden Of Cancer Study* (Globocan) 2020, terdapat sebanyak 36.633 (17.2%) tercatat kasus baru yang terdiagnosa di Indonesia dan dari data tersebut kanker serviks menempati posisi kedua setelah kanker payudara (Dekpes RI, 2020).

Menurut Hutchinon dan Klein (2008) infeksi *Human Papilloma Virus* (HPV) menjadi penyebab utama terjadinya penyakit kanker serviks pada sel Hela. Sel HeLa diketahui mengekspresikan dua onkogen yaitu E6 dan E7 yang mana E6 berikatan dengan p53 yang terfosforilasi sedangkan E7 berikatan dengan protein pRb. Sel kanker serviks sebagian besar memiliki gen p53 dan pRb dalam bentuk *wild type* yang merupakan gen pengatur pertumbuhan yang aktif dalam sel kanker serviks, namun aktifitasnya dihambat oleh ekspresi protein E6 dan E7 dari *Human Papilloma Virus* (Goodwin dan DiMaio, 2000). Beberapa bukti yang menunjukkan HPV sebagai penyebab kanker serviks adalah karsinogen pada kanker serviks dimulai sejak seseorang terinfeksi HPV yang menyebabkan terjadinya gangguan pada sel serviks (Rasjidi, 2009).

Pengobatan kanker secara umum dapat dilakukan dengan kemoterapi. Efek samping dari kemoterapi timbul karena obat-obat kemoterapi tidak hanya menghancurkan sel-sel kanker tetapi juga akan menyerang sel-sel yang sehat. Efek samping yang terjadi dari kemoterapi adalah mual-mual, diare, stomatitis, alopesia, rentan terinfeksi, trombositopenia, neuropoti (Cordell G, 2003). Obat yang digunakan pada kemoterapi adalah senyawa inhibitor yang berfungsi untuk menghambat kinerja dari protein yang berhubungan dengan kanker serviks. Contoh obat yang dapat digunakan adalah obat topotecan yang merupakan obat

inhibitor dari topoisomerase-1 (Pommier Y, 2006). Topoisomerase-1 merupakan salah satu enzim yang mengkatalisis perubahan dalam keadaan topologi DNA. Topotecan bereaksi dengan cara mengikatkan diri pada kompleks pembelahan topoisomerase-1 untuk mencegah topoisomerase-1 dari pelepasan kembali untai DNA yang tergores setelah menghilangkan ketegangan. Hal ini mencegah terjadinya replikasi DNA yang pada akhirnya menyebabkan kematian sel. Proses ini menyebabkan putusannya untaian DNA yang mengakibatkan apoptosis (Staker et al., 2002).

Perancangan inhibitor baru dilakukan untuk memprediksi calon inhibitor baru topoisomerase-1 yang dapat digunakan sebagai calon obat baru untuk penyakit kanker serviks (sel HeLa). Prediksi inhibitor baru topoisomerase-1 dapat dilakukan dengan proses *virtual screening* berbasis QSAR (*Quantitative Structure-Activity Relationship*) model. Metode QSAR merupakan pemodelan dengan pendekatan *in silico* mengenai hubungan secara kuantitatif antara struktur dan aktifitas. Penggunaan metode QSAR memungkinkan untuk memprediksi dengan valid hubungan secara kuantitatif antara aktivitas biologis terhadap struktur molekul (Pranowo, 2011). Model prediksi QSAR yang telah diperoleh dari *virtual screening* dapat digunakan untuk memprediksi hubungan antara protein target dengan molekul lain (Carpenter Huang, 2018).

Penelitian sebelumnya tentang penggunaan *camptothecin* dan turunannya sebagai kandidat terapi untuk kanker telah dilakukan oleh Sourav Pal et al., (2019). Penelitian ini menggunakan pengembangan model farmakofor berbasis ligan dengan Farmakofor 3D-QSAR pada Discovery Studio 4.1. Untuk set pelatihan, 29 turunan CPT (*Camptothecin*) yang beragam dipertimbangkan untuk pengembangan model farmakofor baru. Farmakofor kuantitatif terbaik (Hypo1) dipilih dari 10 farmakofor lainnya berdasarkan selisih tertinggi (112,58) dan koefisien korelasi terbaik (0,917678), nilai terendah (127,392). Modul set tes yang berisi 33 berbagai turunan CPT telah digunakan untuk mengevaluasi kemampuan prediktif model Hypo1. Koefisien korelasi yang dihasilkan ( $R^2$ ) antara aktivitas estimasi dan aktivitas eksperimen untuk 33 senyawa set pengujian diamati sebesar 0,874718 dan untuk 29 senyawa set pelatihan, nilai  $R^2$  adalah 0,917678.

Berdasarkan penelitiannya terpilih tiga molekul yaitu ZINC68997780, ZINC15018994 dan ZINC38550809. Fakta hasil eksperimen ini menunjukkan bahwa model klasifikasi *mechin learning* di kombinasikan dengan *skrining virtual* dan *docking molekuler* dapat sangat meningkatkan efisiensi dalam aspek menemukan inhibitor potensial dengan aktifitas menarik untuk topoisomerase-1 (Sourav Pal et al., 2019).

Penelitian yang akan dilakukan kali ini berbeda dengan penelitian sebelumnya. Penelitian kali ini dilakukan dengan *dataset* yang digunakan dari *database* ChEMBL1781 yang dibagi menjadi data *training* dan *testing*. *Dataset* berisi senyawa aktif dan tidak aktif terhadap protein target. Hasil model terbaik dilakukan *screening* pada molekul dari *database* ZINC sebanyak 481 juta senyawa. Nilai aktivitas yang didapatkan kemudian dilakukan pemeringkatan berdasarkan nilai pIC<sub>50</sub> tertinggi. Hasil yang diperoleh akan dilakukan proses *clustering* untuk mengelompokkan berdasarkan kemiripan senyawa yang kemudian dilakukan uji ADMET (*Absorbtion* , *Distribution*, *Metabolism*, *Excersion* and *Toxicity*) dengan menggunakan aturan *Lipinski's Rule of Five* untuk menentukan senyawa yang bisa diserap oleh tubuh. Uji ADMET digunakan untuk menentukan senyawa yang cocok sebagai calon inhibitor baru untuk protein Topoisomerase-1.

## 1.2 Rumusan Masalah

Rumusan masalah pada penelitian ini adalah sebagai berikut:

1. Bagaimana model terbaik yang diperoleh untuk proses *virtual screening* pada senyawa target CHEMBL1781?
2. Bagaimana hasil *virtual screening* dari model terbaik dengan pemeringkatan nilai aktivitas (pIC<sub>50</sub>) pada calon inhibitor topoisomerase-1?
3. Bagaimana hasil *clustering* dari calon inhibitor protein topoisomerase-1?
4. Bagaimana hasil uji ADMET pada hasil *clustering* dari calon inhibitor protein topoisomerase-1?



### 1.3 Tujuan Penelitian

Tujuan penelitian ini adalah sebagai berikut:

1. Mengetahui protein target topoisomerase-1 pada *database* ChEMBL dengan kode ChEMBL1781
2. Mengetahui hasil *screening* untuk protein topoisomerase-1 pada *database* ChEMBL1781 yang digunakan sebagai dataset untuk pemodelan QSAR
3. Mengetahui proses *virtual skrining* dengan pendekatan *Machine Learning* menggunakan Python 3.7.1.
4. Mengetahui hasil uji ADMET calon inhibitor protein topoisomerase-1

### 1.4 Batasan Masalah

Batasan masalah pada penelitian ini adalah sebagai berikut:

1. Mengetahui model terbaik yang diperoleh dengan menggunakan metode Pemodelan QSAR untuk senyawa target CHEMBL1781
2. Mengetahui proses hasil *virtual skrining* dari model terbaik dengan pemeringkatan nilai aktivitas (pIC50) calon inhibitor topoisomerase-1
3. Mengetahui proses hasil *clustering* dari calon inhibitor protein topoisomerase-1

### 1.5 Manfaat Penelitian

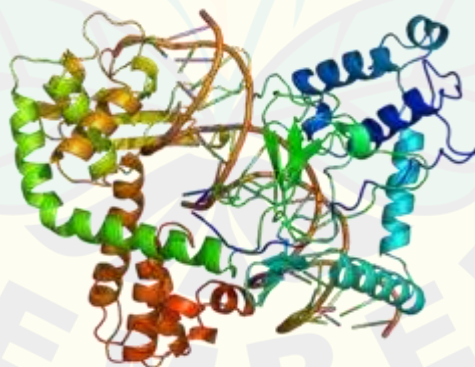
Manfaat yang diperoleh dari penelitian Pemodelan QSAR untuk Screening virtual calon inhibitor topoisomerase-1 menggunakan QSAR pada penyakit kanker serviks adalah sebagai berikut :

1. Memberikan informasi mengenai penggunaan QSAR untuk mencari calon inhibitor suatu senyawa
2. Memberikan informasi mengenai senyawa aktif yang dapat menjadi inhibitor dari protein topoisomerase-1 sebagai calon obat kanker serviks.
3. Memberikan informasi nilai aktivitas senyawa terbaik sebagai inhibitor topoisomerase-1 untuk calon obat kanker serviks.
4. Memberikan informasi senyawa hasil uji ADMET untuk calon inhibitor topoisomerase-1

## BAB 2. TINJAUAN PUSTAKA

### 2.1 Topoisomerase-1

*Topoisomerase-1* atau DNA topoisomerase-1 merupakan salah satu enzim yang mengkatalisis perubahan dalam keadaan topologi DNA. Topoisomerase-1 adalah enzim yang memotong salah satu dari dua untai DNA beruntai ganda untuk menghilangkan batasan heliks yang dapat menghambat replikasi dan transkripsi DNA sehingga menghambat pertumbuhan sel (Wang, 2002). Topoisomerase-1 terdapat tiga kelas yaitu topoisomerase-1A, topoisomerase-1B dan topoisomerase-1C. Topoisomerase-1A adalah enzim yang membuat pemutusan tunggal pada DNA dan melewati untaian atau dupleks kedua melalui pemutusan tersebut (Changela et al., 2001). Topoisomerase-1B berbeda dengan topoisomerase-1A, topoisomerase-1B disebut sebagai superkoil positif atau negatif yang memecahkan masalah *overwound* dan *underwound*. Topoisomerase-1C diidentifikasi sebagai topoisomerase-V di archaeon *Methanopyrus kandleri* yang mana topoisomerase-1C ini bekerja melalui mekanisme putar yang terkontrol seperti topoisomerase-1B (Taneja et al., 2006).



Gambar 2.1 struktur topoisomerase-1  
(Redinbo et al., 1998)

Topoisomerase-1 membentuk perantara fosfotirosin antara fosforil scissile dan tirosin katalitik dari tulang punggung DNA. Topoisomerase-1A membentuk



hubungan kovalen antara katalitik tirosin dan 5'-fosforil. Topoisomerase-1B dan topoisomerase-1C membentuk zat antara 3'-fosfotirosin kovalen (Champoux, 2003). Topoisomerase mempunyai fungsi yang penting dalam proses intraseluler, yaitu berperan dalam proses replikasi, transkripsi, rekombinasi DNA dan proses proliferasi dari sel kanker. Dengan dihambatnya aktivitas enzim DNA Topoisomerase oleh senyawa inhibitor, maka proses terjadinya ikatan antara enzim dengan DNA sel kanker semakin lama. Sehingga akan terbentuk *Protein Linked DNA Breaks* (PLDB), akibatnya terjadi fragmentasi atau kerusakan DNA sel kanker dan selanjutnya berpengaruh terhadap proses di dalam sel khususnya proses replikasi sel yang diakhiri dengan kematian sel kanker. Kerusakan DNA ini bila tidak dikendalikan akan berlangsung terus melalui metastase sel DNA atau mengganas. Keganasan ini bisa dikendalikan melalui proses apoptosis yaitu proses kematian sel yang terprogram (Elmore, 2007).

## 2.2. Kanker Leher Rahim (Kanker Serviks)

Kanker merupakan penyakit yang disebabkan oleh sel-sel abnormal yang tumbuh secara terus menerus sehingga terjadinya pertumbuhan sel yang tidak terkendali dan sel kanker ini tidak mengalami regulasi (Ray, 2014). Penyakit kanker leher rahim atau biasa disebut kanker serviks merupakan kanker leher rahim atau kanker serviks termasuk dalam kategori tumor ganas yang timbul di leher rahim wanita. Kanker serviks adalah kanker yang terjadi karena sel epitel serviks mengalami mutasi genetik sehingga terjadi pembelahan sel yang tidak terkendali. Kanker ini dapat meluas ke vagina, rahim hingga indung telur (Haryani et al., 2016). Menurut data *Global Burden Of Cancer Study* (Globocan) 2018, terdapat sebanyak 32.469 jiwa di Indonesia yang menderita kanker serviks dan dari data tersebut kanker serviks menempati posisi kedua setelah kanker payudara dengan kematian 18.279 jiwa pertahunnya (Benita et al., 2020).

Penyebab utama kanker serviks adalah infeksi virus *Human Papilloma Virus* (HPV). Infeksi HPV mengakibatkan terjadinya integrasi genom DNA HPV dengan host sehingga terjadi gangguan atau hilangnya gen E2 virus yang menyebabkan terekspresinya onkogen virus E6 dan E7. Produk E6 dan E7

menghambat aktivitas tumor supresor p53 dan protein Rb (Chakrabarti & Krishna, 2003). Menurut Cancernet (2001) dalam Price dan Wilson (2005) penelitian epidemiologi di seluruh dunia menegaskan bahwa infeksi HPV adalah faktor penting dalam perkembangan kanker servikal. Lebih dari 20 tipe HPV yang berbeda mempunyai hubungan dengan kanker serviks. Penelitian memperlihatkan bahwa perempuan dengan HPV 16 dan 18 mempunyai angka *neoplasia intraepithelial servikal* (CIN) yang lebih tinggi. Penelitian terbaru memperlihatkan bahwa perempuan dengan HPV strain 18 memiliki angka mortalitas yang lebih tinggi dan prognosis yang lebih buruk (Prabawa et al., 2012).

### 2.3 Sel HeLa

*HeLa cell line* merupakan sel turunan yang tumbuh sebagai sel yang semi melekat. Sel HeLa diturunkan dari sel epitel kanker leher rahim (serviks) manusia. Sel ini diisolasi tahun 1951 dari rahim wanita penderita kanker leher rahim bernama Henrietta Lacks yang berusia 31 tahun. Sel ini secara morfologi merupakan sel epitel yang sudah dimasuki oleh *Human Papilloma Virus* (HPV) tipe 18. Sel hela bersifat imortal dan sangat agresif sehingga mudah untuk dikultivasi tetapi sel ini mudah menginvasi kultur sel lain (DeFilippis et al., 2003).

Sel HeLa adalah sel kanker leher rahim akibat infeksi Human Papillomavirus (HPV 18). Sel kanker leher rahim yang diinfeksi HPV diketahui mengekspresikan 2 onkogen, yaitu E6 dan E7. Protein E6 dan E7 terbukti dapat menyebabkan sifat imortal pada kultur primer keratinosit manusia, namun sel yang imortal ini tidak bersifat tumorigenik hingga suatu proses genetik terjadi. Jadi, onkogen tersebut tidak secara langsung menginduksi pembentukan tumor, tetapi menginduksi serangkaian proses yang pada akhirnya dapat menyebabkan sifat kanker (Horner et al., 2004).

### 2.4 Desain Obat *In Silico*

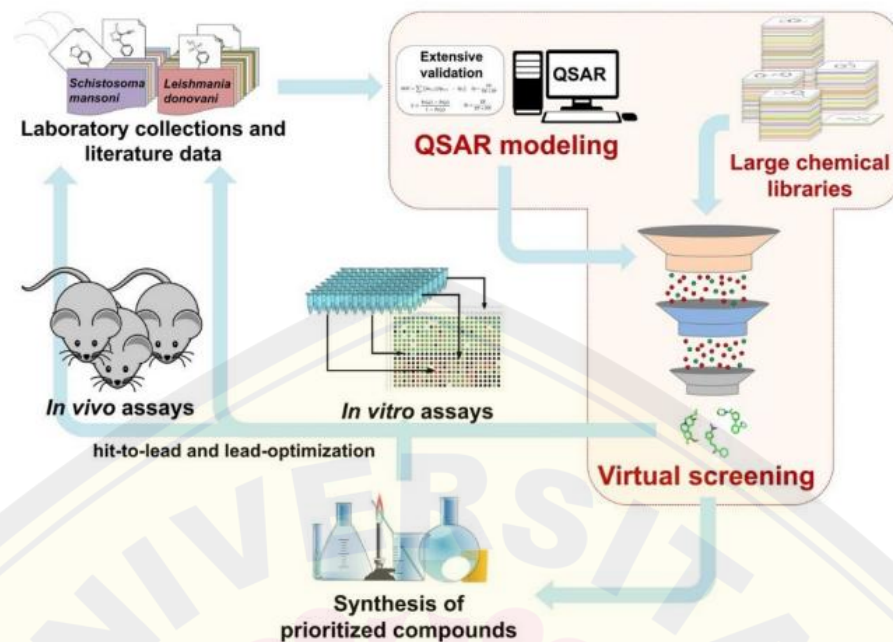
Desain obat (*drug design*) merupakan perancangan obat dalam penemuan obat baru terhadap suatu penyakit dengan aktivitas yang tinggi. Penemuan obat

baru konvensional dilakukan secara eksperimen dengan mengambil senyawa dari alam untuk diuji di laboratorium. Senyawa hasil eksperimen belum tentu memiliki interaksi dan aktivitas hambat terhadap suatu penyakit, sehingga eksperimen yang dilakukan bersifat *trial and error* hingga skrining diperoleh senyawa yang tepat untuk kemudian dapat diuji kinerjanya dengan uji *in vitro*, *in vivo* dan uji klinis. Strategi drug design untuk pengujian aktivitas dapat menjadi sangat rumit dan diperlukan waktu cukup lama untuk sampai pada sintesis obat dan pemanfaatan obat. Pengembangan percobaan untuk efisiensi waktu dan optimasi aktivitas senyawa yaitu dengan melakukan percobaan secara komputerisasi (*in silico*). Uji *in silico* menjadi metode awal dalam penemuan obat baru sebelum pada uji klinis (Hardjono, 2013).

Penerapan kimia komputasi dalam bidang farmasi yaitu melalui penentuan obat berdasarkan sifat biologis penting dan optimasi profil aktivitas ataupun sintesis molekul obat. Desain obat umumnya didasarkan pada uji kemiripan struktural dan klasifikasi molekul *active* dan *inactive*. Pendekatan teoritis telah dikembangkan untuk dapat menghitung secara kuantitatif mengenai hubungan antara aktivitas biologis terhadap perubahan struktur molekul yaitu dikenal sebagai QSAR (Pranowo, 2011). Interaksi molekul obat terhadap reseptor dapat disimulasikan dan dilihat energi interaksinya melalui uji *in silico* menggunakan docking dengan mempertimbangkan sifat-sifat molekulnya (Jensen, 2007).

### **2.5 *Quantitative Structure Activity Relationship (QSAR)***

QSAR (*Quantitative Structure and Activity Relationship*) merupakan pemodelan dengan pendekatan *in silico* mengenai hubungan secara kuantitatif antara struktur dan aktivitas. QSAR penting dalam memahami sifat biologis dan fisikokimia molekul, sehingga QSAR digunakan untuk menyaring (*skrining*) banyak molekul kimia, memahami obat, dan untuk merancang atau desain obat baru (Satbhaiya and Chourasia, 2015). Pengembangan model QSAR dengan skrining virtual yaitu untuk model regresi dan untuk model klasifikasi (Roy, 2015).



Gambar 2.2 Proses Virtual Screening berbasis QSAR

Model yang dikembangkan memungkinkan identifikasi dengan cepat dari sejumlah besar senyawa melalui aktivitas biologisnya (Ai et al., 2019). Pendekatan QSAR menjadi pilihan dalam desain obat berbasis ligan, sebagai alat prediksi dasar (Verma, Khedkar and Coutinho, 2010). Prinsip metode QSAR yaitu pada struktur yang serupa akan memiliki aktivitas yang sama (Martin, Kofron and Traphagen, 2002). Kemampuan prediktif QSAR akan terbatas dan bias untuk skrining virtual pada database yang besar, jika menggunakan dataset yang kecil dan hanya terbatas pada molekul turunannya. Analisis QSAR menggunakan dataset yang menampilkan fitur struktural utama yang bertanggung jawab atas aktivitas molekul (Roy, 2015). Dataset digunakan untuk membangun model pembelajaran mesin (*machine learning*) (Shi, Zhao and Wei, 2018).

Model QSAR regresi bisa dinilai performanya dengan melihat beberapa ukuran statistik seperti koefisien regresi ( $R$ ), koefisien determinasi ( $R^2$ ), *Mean Squared Error* (MSE) untuk setiap algoritma machine learning. Model *hyperparameter* dipilih berdasarkan *Root Mean Square Error* (RMSE) terendah. RMSE mengukur rata-rata perbedaan kuadrat antara nilai yang sebenarnya dan nilai yang diprediksi oleh model. RMSE memperhitungkan perbedaan positif atau



negatif, dan semakin kecil nilai RMSE, semakin baik kualitas modelnya. Angka RMSE selalu positif. Nilai minimum yang direkomendasikan untuk model QSAR regresi agar menghasilkan prediksi yang akurat adalah nilai prediksi  $R \geq 0,8$  dan  $R^2 \geq 0,6$ . Semakin tinggi nilai R dan  $R^2$ , semakin baik performa modelnya dalam memprediksi data (Hermansyah et al., 2020).

## 2.6 Pengobatan Kanker Serviks

Pengobatan kanker serviks dapat dilakukan melalui berbagai metode yaitu kemoterapi, radioterapi dan anti VEGF. Pengobatan kanker serviks yang dinilai lebih efektif adalah dengan cara kemoterapi. Kemoterapi merupakan pengobatan kanker menggunakan obat dengan membunuh sel-sel kanker yang disuntikkan kedalam vena atau diberikan secara oral. Pengobatan secara kemoterapi dapat diklasifikasikan berdasarkan mekanisme kerjanya yaitu antimetabolit, agen alkilasi, inhibitor topoisomerase dan lainnya (Bukowski et al., 2020).

Antimetabolit pada kemoterapi akan mengganggu pembentukan DNA atau RNA sel kanker dengan cara memutus pembentukan untaian DNA melalui penghambatan enzim tertentu. Obat golongan agen alkilasi (cisplatin, carboplatin, dan oxaliplatin) bekerja dengan cara mentransfer gugus alkil pada basa nitrogen DNA yang menyebabkan kesalahan pembentukan pasangan DNA. Inhibitor topoisomerase I (irinotecan dan topotecan) dan inhibitor topoisomerase II (idarubicin, daunorubicin, dan doxorubicin) akan menghambat aktivitas topoisomerase yang terlibat dalam replikasi DNA yang menyebabkan putusnya untaian DNA (Bukowski et al., 2020).

## 2.7 Inhibitor

Inhibitor merupakan senyawa yang dapat menghambat kerja enzim dengan cara menurunkan laju reaksi kimianya. Inhibitor hanya terbatas pada protein yang berfungsi sebagai katalisator yang berfungsi mempercepat reaksi. Fungsi pada mekanisme penghambatan enzim ini dapat digunakan untuk penemuan obat suatu penyakit (Goodman, 2008).

Inhibitor dari topoisomerase-1 yang sudah lolos uji dan sudah di pasaran

contohnya yaitu Topotecan. Mekanisme inhibisi topotecan yaitu menghambat terjadinya ikatan antara topoisomerase-1 dan DNA, dengan kata lain obat ini bekerja untuk mematikan sel kanker atau menghambat sel kanker berkembang. Inhibitor dari topoisomerase-1 ada yang sudah diedarkan dan disetujui oleh *Food and Drug Administration* (FDA) untuk obat kanker serviks dan ada yang dalam tahap uji klinis. Inhibitor yang sudah disetujui oleh FDA yaitu topotecan. Inhibitor topoisomerase-1 yakni topotecan lulus uji FDA pada tahun 2022 berdasarkan database ChEMBL ([www.ebi.ac.uk/chembl](http://www.ebi.ac.uk/chembl)).

## 2.8 ChEMBL

ChEMBL adalah database dengan akses terbuka yang memuat sejumlah senyawa bioaktif dengan informasi pengikatan, fungsional, dan ADMET (*Absorption Distribution Metabolism Excretion dan Toxicology*). ChEMBL merupakan database penemuan obat yang bertujuan untuk menjaring data dan pengetahuan kimia obat pada seluruh penelitian farmasi dan pengembangannya (Mendez et al., 2019). Sumber data pada ChEMBL terkait data hubungan aktivitas yaitu melalui hubungan antara ligan dan molekul target dalam bentuk aksi yang diukur secara eksperimental hingga diperoleh aktivitas konsentrasi penghambatan (IC<sub>50</sub>). Data nama dan sinonim senyawa ditambahkan dengan informasi sifat-sifat, representasi struktur dan aksi mekanisme obat. Situs web ChEMBL dapat diakses melalui laman <https://www.ebi.ac.uk/chembl> secara gratis (Davies et al., 2015).

## 2.9 ZINC

ZINC adalah basis data molekuler yang tersedia secara komersial. Pada dasarnya, data molekuler yang terkandung dalam database ini ada di bentuk sistem entri jalur masukan molekul yang disederhanakan (SMILES), namun dalam perkembangannya, dua dimensi dan tiga dimensi representasi molekul juga tersedia. Zinc database sering digunakan dalam penelitian untuk menemukan ligan, yang merupakan ion atau molekul yang dapat melekat pada atom logam melalui ikatan kovalen. Ligan ini sering digunakan untuk menemukan obat

penyakit atau virus dengan mengganggu siklus hidup virus atau langsung menghancurkannya secara struktural (Irwin et al., 2012).

### 2.10 *Molecular Fingerprint*

*Molecular Fingerprint* (MF) merupakan bentuk representasi digital yang mengkode ciri struktural kimia dari senyawa kedalam bentuk vektor biner yang hanya terdiri dari 0 dan 1. Vektor dalam molekuler fingerprint digunakan dalam berbagai hal, seperti pada *virtual screening*, pencarian kesamaan dan pengelempokan senyawa kimia. Vektor biner 0 pada *molecular fingerprint* menunjukkan bahwa tidak ada struktur kimia tertentu (seperti -OH) pada senyawa kimia, sedangkan untuk nilai 1 menunjukkan bahwa ada struktur senyawa tersebut. Senyawa kimia yang memiliki ciri struktural yang berbeda, seperti adanya gugus -OH dan -Br, akan mendapatkan representasi digital yang berbeda pula pada *molecular fingerprint*. *Molecular fingerprint* memiliki keunggulan dimana ia lebih mudah dipahami jika dibandingkan dengan *molecular docking*. Dimana ia mampu menghasilkan *molecular fingerprint* dari ribuan senyawa dalam waktu kurang dari satu detik dan hal tersebut tidak mungkin dapat dilakukan dengan *molecular docking*. *Molecular fingerprint* dapat digabungkan dengan *machine learning* untuk mengembangkan model QSAR yang dapat digunakan dalam prediksi aktivitas biologis ligan, toksisitas dan konstanta laju radikal OH terhadap kontaminan organik (Ding et al., 2021).

### 2.11 SMILES (*Simplified Molecular-Input Line-Entry System*)

SMILES yaitu notasi kimia yang digunakan untuk pemrosesan informasi dari kimia modern agar dapat menggambarkan struktur kimia dengan karakter-karakter ASCII (Weininger, 1988). Notasi SMILES adalah tatanama senyawa kimia yang sederhana sehingga dapat dibaca dengan baik oleh manusia maupun komputer yang dapat dikonversi menjadi struktur molekul dalam 2D. Notasi SMILES dapat dibangun dari prinsip molekul, yaitu grafik molekul dan notasi dapat diturunkan untuk proses virtual. Notasi SMILES dapat digunakan sebagai pertukaran data dalam bidang informatika. Atom yang digunakan dalam notasi

SMILES adalah B, C, N, O, P, S, F, Cl, Br, I, OH, @, =, #, dan () (Bajusz et al., 2017).

### 2.12 ADMET (*Absorpsi, Distribusi, Metabolisme, Ekskresi, dan Toksisitas*)

ADMET ialah singkatan dari *Absorpsi, Distribusi, Metabolisme, Ekskresi, dan Toksisitas*. Istilah ini mengacu pada sifat-sifat suatu obat yang menentukan seberapa efektif dan aman obat tersebut digunakan, termasuk bagaimana obat tersebut diserap oleh tubuh, bagaimana didistribusikan di seluruh tubuh, bagaimana diolah dan dikeluarkan oleh tubuh, serta potensi efek beracun yang mungkin dimiliki oleh obat tersebut. Sifat-sifat ADMET penting untuk dipertimbangkan dalam proses penemuan dan pengembangan obat, karena dapat mempengaruhi kesuksesan uji klinis serta keselamatan dan efektivitas keseluruhan obat. Alat komputasi seperti AdmetSAR dapat digunakan untuk memprediksi sifat-sifat ADMET dari obat-obatan dengan menggunakan metode pembelajaran mesin dan model QSAR (Sucharitha et al., 2022).



### BAB 3. METODOLOGI PENELITIAN

#### 3.1 Tempat dan Waktu Penelitian

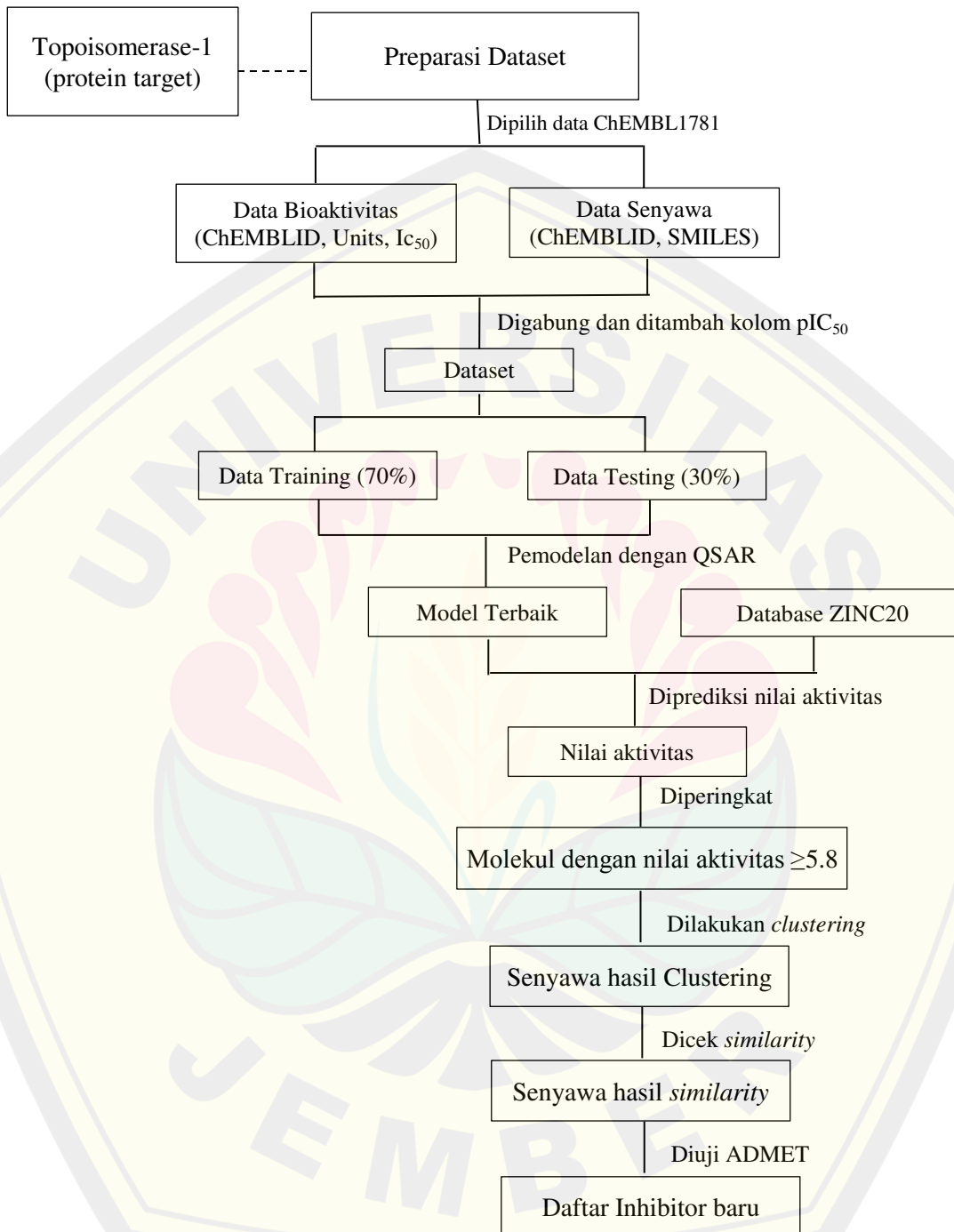
Penelitian ini dilakukan di Laboratorium Komputer Jurusan Kimia Fakultas Matematika dan Ilmu Pengetahuan Alam (FMIPA) Universitas Jember pada bulan Februari 2024 hingga selesai.

#### 3.2 Alat

Alat yang digunakan pada penelitian ini yaitu computer Lenovo IdeaCenter AIO 520-5MID sebanyak 14 buah dengan sistem operasi *Linux Ubuntu 20.04 LTS processor* Intel Core i5-7400T, dan RAM 4GB yang dilengkapi dengan tambahan program aplikasi sebagai berikut :

- Python 3.7.1

### 3.3 Diagram Alir Penelitian



### 3.4 Prosedur Penelitian

#### 3.4.1 Preprasi Data

Dataset yang digunakan pada penelitian ini disiapkan dengan mengambil data dari ChEMBL. Preparasi data dilakukan dengan mencari target protein dengan UniProt ID P11387. Target data dari ChEMBL kemudian dipilih target dengan kode ChEMBL1781. Target dengan ID ChEMBL1781 kemudian didapatkan data bioaktivitasnya. Data bioaktivitas kemudian difilter sehingga data yang tidak bernilai dan data molekul duplikat dihilangkan. Untuk data standar unit akan diubah dalam bentuk data bilangan desimal. Nama kolom untuk standar unit dan standar value diubah menjadi unit dan IC50. Data yang telah diproses kemudian diurutkan kembali. Data senyawa untuk target ChEMBL1781 juga dicari kemudian difilter untuk mendapatkan SMILES dari masing-masing molecule chembl id, menghapus data yang duplikat, menghapus data yang tidak bernilai. Data bioaktivitas dan data senyawa kemudian digabung dengan hanya ada kolom ChEMBL ID, SMILES, units dan IC50. Kemudian pada dataset tersebut ditambahkan kolom PIC50. Dataset yang telah diperoleh kemudian disimpan.

#### 3.4.2 Pemodelan menggunakan QSAR

Dataset dari protein target Topoisomerase-1 dengan kode ChEMBL1781 yang diperoleh pada prosedur sebelumnya digunakan untuk pemodelan QSAR, dimana data akan dibagi menjadi data training 70%, data dan data testing 30%. Data SMILES pada dataset diubah menjadi fingerprint untuk digunakan sebagai fitur dalam model jaringan saraf tiruan. Model jaringan saraf tiruan dibuat dengan menentukan jumlah neuron di setiap lapisan tersembunyi, fungsi aktivasi, dan kompilasi model. Model diuji dengan beberapa ukuran batch yang berbeda dan hasil pelatihannya direkam untuk evaluasi. Evaluasi model akan dihasilkan *Mean Square Error* (MSE), *Mean Absolute Error* (MAE), dan *R-squared* (R<sup>2</sup>). Model dengan nilai terbaik yang didapatkan akan digunakan untuk skrining.

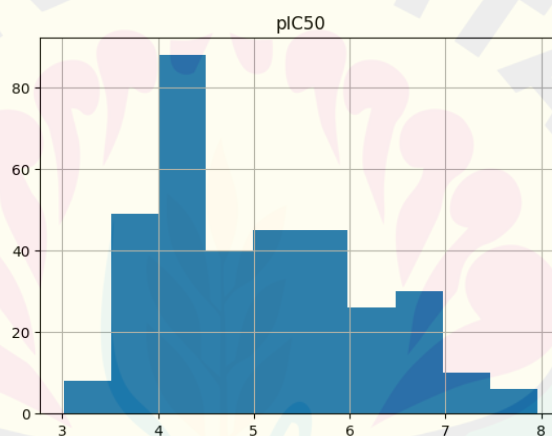
### 3.4.3 Skrining Virtual

Model terbaik yang telah didapatkan dari prosedur sebelumnya kemudian digunakan untuk skrining dengan menggunakan data eksternal. Data eksternal yang digunakan berasal dari database ZINC dengan data sejumlah 431 juta molekul. Data eksternal berisi SMILES dari masing-masing molekul. Model kemudian digunakan untuk memprediksi nilai pIC<sub>50</sub> pada data eksternal. Nilai aktivitas yang diperoleh kemudian dilakukan clustering untuk mengelompokkan berdasarkan kemiripan senyawanya. Hal itu nantinya akan menunjukkan nilai aktivitas terbaik dari kelompok yang sama yang dapat dijadikan sebagai prediksi calon obat baru untuk kanker servik. Hasil clustering dilakukan uji ADMET (*Absorbtion, Distribution, Metabolism, Excretion, and Toxicity*). Hal itu nantinya akan menunjukkan apakah obat tersebut dapat dikonsumsi.

## BAB 4. HASIL DAN PEMBAHASAN

### 4.1 Pemodelan QSAR Menggunakan *Machine Learning*

Pemodelan dilakukan dengan mempersiapkan dataset dengan kode ChEMBL1781 yang diperoleh dari database ChEMBL (<https://ebi.ac.uk/chembl>) dengan protein target UniProt IDP11387. *Dataset* berisi 429 data struktur molekul yang mempresentasikan struktur molekul dalam bentuk smiles dan nilai aktivitas molekul yang berupa nilai  $pIC_{50}$  (Lampiran 4.1). Dataset disimpan dalam format \*.csv yang berisi tiga kolom bagian mengenai informasi chembl\_id, smiles, dan nilai  $pIC_{50}$ . Sebaran nilai  $pIC_{50}$  yang terdapat pada dataset ditunjukkan pada grafik 4.1 berikut:

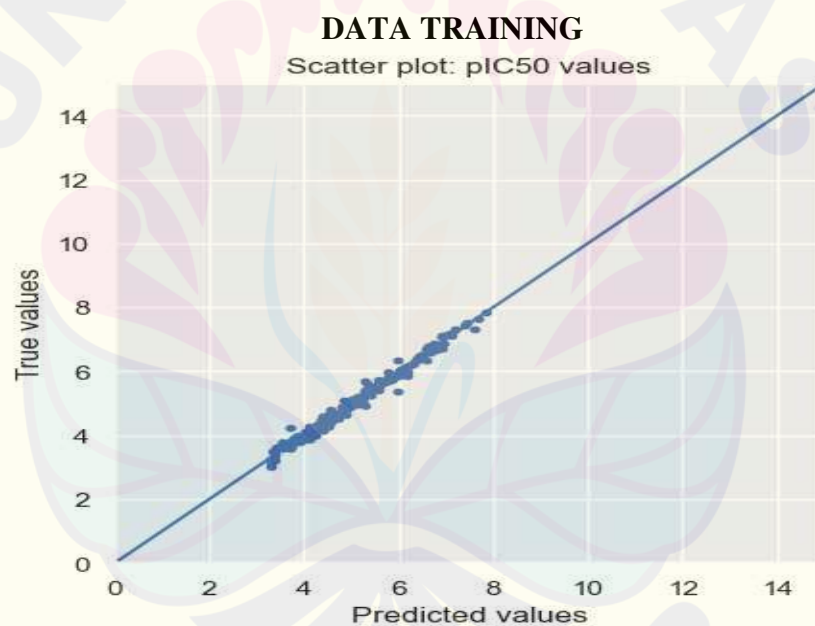


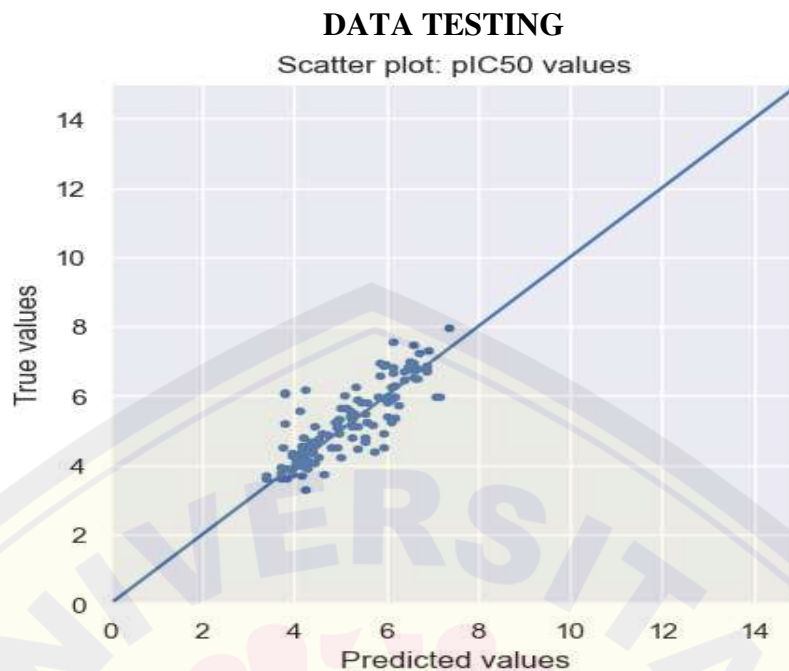
Gambar 4.1 Sebaran Nilai  $pIC_{50}$  pada Dataset

Dataset kemudian dibagi menjadi dua bagian yang berbeda, yaitu data *training* dan *testing* dengan nilai masing-masing sebesar 70% dan 30% dari keseluruhan dataset. Data *training* berfungsi untuk membuat model yang sesuai dengan pola data. Data *testing* berfungsi untuk menguji keakuratan model yang dibuat dengan menggunakan nilai R square dan MSE sebagai parameternya. Data *training* pada penelitian ini diperoleh 300 senyawa untuk melatih model dan data *testing* yang diperoleh 129 senyawa untuk menguji model yang dihasilkan.

Pemodelan QSAR dilakukan dengan cara mempelajari pola data yang ada dengan kombinasi *hyperparameter neural network* menggunakan *machine learning*. Kombinasi dilakukan dengan memvariasikan nilai L1, L2, dan L3 yaitu 8, 16, 32, hingga 2048 dan nilai batch size yang divariasikan adalah 8, 16, 32,

hingga 2048. Kombinasi *hyperparameter* ini bertujuan untuk menemukan model yang terbaik berdasarkan nilai  $R^2$  dan MSE. Syarat nilai  $R^2$  pada QSAR memiliki nilai minimal yaitu  $R^2 > 0.6$ . Nilai  $R^2$  merupakan nilai prediksi yang mana apabila nilai  $R^2$  mendekati 1, maka data prediksi yang diperoleh dikatakan mendekati sempurna (Hermansyah et al., 2021). Model yang terbaik didapatkan pada *Hidden Layer* (L1,L2, dan L3) dan *batch size* secara berurutan yaitu 8; 16; 32 dan 8; 1024; 2048 (Lampiran 4.2). Nilai parameter  $R^2$  untuk data *training* dan data *testing* yang diperoleh yaitu 0.98 dan 0.66. Nilai-nilai yang dihasilkan telah memenuhi syarat minimal untuk nilai  $R^2$ , dimana nilai  $R^2 > 0.6$  (Hermansyah et al., 2021). Nilai data *training* dan data *testing* untuk nilai  $pIC_{50}$  eksperimen dan  $pIC_{50}$  prediksi dibuat plot regresi yang dihasilkan seperti gambar berikut:





Gambar 4.2 Plot Regresi Nilai pIC<sub>50</sub> pada Data Training dan Data Testing

Nilai MSE yang dihasilkan dari pemodelan untuk masing-masing data *training* dan data *testing* adalah 0.01 dan 0.38 yang menunjukkan nilai dari kinerja model. Nilai parameter MSE yang rendah menunjukkan nilai kinerja model yang baik dikarenakan semakin kecil kesalahan atau *error* yang terjadi. Model terbaik yang didapatkan akan digunakan untuk proses *virtual screening* dengan database ZINC ([zinc20.docking.org](http://zinc20.docking.org)).

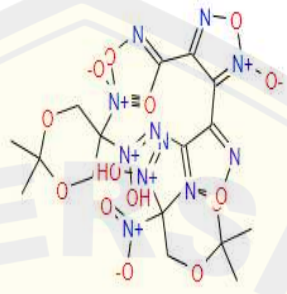
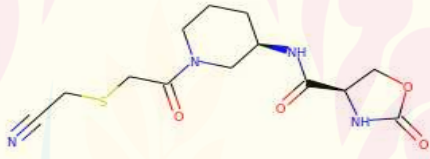
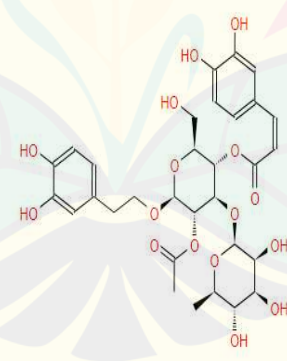
#### 4.2 Virtual Screening Calon Inhibitor Topoisomerase-1

Model terbaik dengan nilai  $R^2$  sebesar 0.66 digunakan untuk proses *virtual screening* yang dilakukan dengan menggunakan database ZINC ([zinc20.docking.org](http://zinc20.docking.org)). Database ZINC digunakan karena memiliki jumlah senyawa yang terdaftar sejumlah 481 juta senyawa. Senyawa yang sudah diprediksi nilai pIC<sub>50</sub>-nya lalu diperingkat dari yang tertinggi dan diambil senyawa dengan nilai pIC<sub>50</sub> diatas rata-rata obat. Nilai IC<sub>50</sub> merupakan konsentrasi obat yang dibutuhkan untuk menghambat 50% dari aktivitas protein target, dimana semakin rendah nilai konsentrasinya akan semakin baik obatnya dan menunjukkan efek samping yang lebih rendah (Gabber et al., 2022). Nilai rata-rata obat yang diambil pIC<sub>50</sub>-nya sebesar 5.8 dikarenakan obat yang sudah



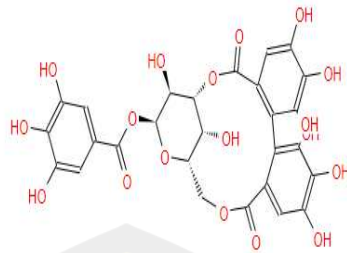
beredar yaitu topotecan memiliki nilai  $pIC_{50}$  rata-rata sebesar 5.8. Jumlah senyawa dengan nilai prediksi diatas 5.8 sebanyak 335 senyawa (Lampiran 4.3). Berikut beberapa contoh senyawa dengan nilai  $pIC_{50}$  diatas 5.8

Tabel 4.1 Nilai  $pIC_{50}$  Hasil Prediksi Beberapa Senyawa

Kode ZINC	Struktur	$pIC_{50}$
ZINC225434563		7.703523
1070452203		5.819089
72148711		5.818800

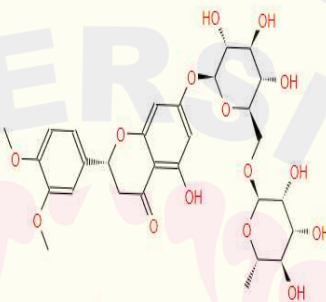


1834878099



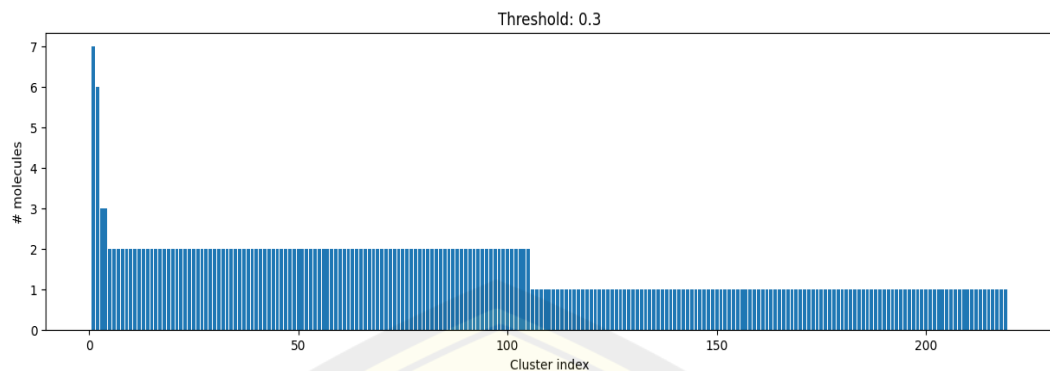
5.817501

1002478097



5.817241

Senyawa hasil screening tersebut kemudian dilakukan pengelompokan kembali (*clustering*) menggunakan metode *cluster-Butina* dan dilakukan menggunakan bahasa pemrograman *python 3.7.1*. Metode *cluster-Butina* merupakan suatu metode yang dikembangkan untuk mengidentifikasi *cluster* yang lebih kecil namun serupa yang dapat digunakan untuk beberapa ukuran *similarity*, dimana *centroid cluster* akan lebih mirip daripada ambang batas yang diberikan pada setiap molekul lain dalam *cluster* (J. Chem. Inf. Model, 1999). Ukuran ambang batas diukur pada skala antara 0 hingga 1, diman 0 diartikan tidak mirip dan 1 adalah paling mirip. Zahoranszky-Kohalmi et al. (2016) menjelaskan bahwa penentuan nilai ambang batas (*threshold*) pada perhitungan nilai *similarity* cukup penting dalam pengelompokan senyawa untuk melihat tingkat kemiripannya. Pada penelitian ini menggunakan ambang batas 0.3 yang berarti kemiripannya 0.7 pada setiap molekul dalam *cluster*. Pengelompokan pada 335 senyawa dengan ambang batas 0.3 dapat ditunjukkan pada gambar 4.3 berikut.



Gambar 4.3 Ambang batas 0.3 pada *Cluster*

Jumlah *cluster* yang dihasilkan dari ambang batas 0.3 sebanyak 219 dari 335 senyawa dengan jumlah molekul terbesar dalam satu *cluster* adalah 7 molekul. Kemiripan antara dua titik acak dalam *cluster* yang sama yaitu 0.72 dan kemiripan antara dua titik acak di *cluster* yang berbeda ialah 0.12. Pengelompokan menghasilkan 219 *cluster* yang selanjutnya hanya diambil 50 *cluster*. Pengelompokan dilakukan berdasarkan kemiripan dari setiap senyawa berdasarkan stukturanya. Hasil pembagian kelompok atau *clustering* menghasilkan jumlah senyawa yang berbeda pada setiap *cluster*-nya. Jumlah senyawa pada masing-masing *cluster* seperti yang ditunjukkan pada tabel 4.2 sebagai berikut.

Tabel 4.2 Jumlah Senyawa pada Masing-Masing *Cluster*

<i>Cluster</i>	Jumlah	<i>Cluster</i>	Jumlah
1	7	26	2
2	6	27	2
3	3	28	2
4	3	29	2
5	2	30	2
6	2	31	2
7	2	32	2
8	2	33	2
9	2	34	2
10	2	35	2
11	2	36	2
12	2	37	2
13	2	38	2

14	2	39	2
15	2	40	2
16	2	41	2
17	2	42	2
18	2	43	2
19	2	44	2
20	2	45	2
21	2	46	2
22	2	47	2
23	2	48	2
24	2	49	2
25	2	50	2

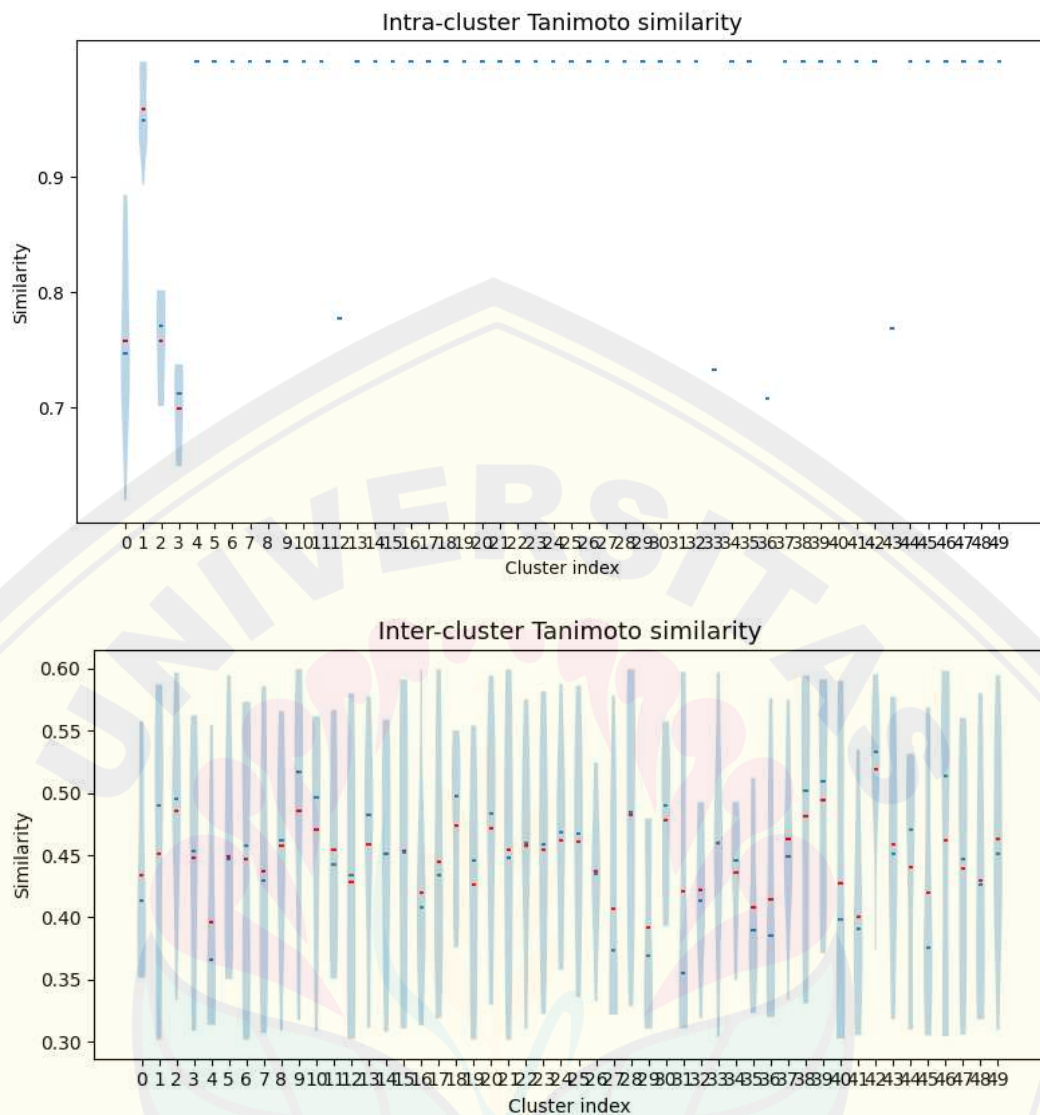
*Cluster* yang dihasilkan berisi daftar senyawa yang setiap *cluster*-nya diwakili oleh senyawa yang memiliki nilai pIC<sub>50</sub> tertinggi, sehingga terdapat 50 senyawa yang mewakili masing-masing pada setiap *cluster* (Lampiran 4.4). Hasil *clustering* juga dapat digunakan sebagai saran untuk uji senyawa calon inhibitor Topoisomerase-1 pada penyakit kanker serviks. Senyawa yang mewakili pada setiap masing-masing *cluster* ditunjukkan pada tabel 4.3 sebagai berikut.

Tabel 4.3 Senyawa yang Mewakili Masing-Masing *Cluster*

<i>Cluster</i>	Kode ZINC	Nilai pIC <sub>50</sub>	<i>Cluster</i>	Kode ZINC	Nilai pIC <sub>50</sub>
1	978812160	5.966101	26	490397741	5.854763
2	2061000274	6.489870	27	491596631	5.854763
3	1218180269	5.823065	28	485446348	5.854763
4	1075058875	6.030778	29	490371532	5.854763
5	796707002	5.8218	30	438599338	5.8561316
6	726030648	5.822926	31	16399818	5.8567896
7	791778342	5.82339	32	443988865	5.8570614
8	609564700	5.824991	33	443732548	5.8570614
9	748379943	5.824992	34	1066891191	5.8616095
10	25501234	5.825045	35	531491814	5.862089
11	1847170840	5.826375	36	433375597	5.863386
12	147056304	5.826794	37	1066509554	5.8638225
13	1837673642	5.833285	38	530660566	5.865591
14	441808795	5.834733	39	796505256	5.867245

15	527265642	5.835548	40	769840707	5.880191
16	796072908	5.842773	41	543954687	5.8809576
17	522921657	5.843547	42	543938258	5.8809576
18	146474435	5.8444204	43	790866447	5.8921604
19	303273693	5.8444204	44	950496003	5.8945603
20	795781765	5.8460016	45	794629907	5.895825
21	533937767	5.848956	46	500935411	5.897694
22	7253324	5.8500595	47	18175031	5.9036655
23	788277226	5.85292	48	752351814	5.9051723
24	792156789	5.8531313	49	751660328	5.9051723
25	795611139	5.853404	50	610424627	5.906119

Hasil yang didapatkan dari proses *clustering* dilakukan pengecekan nilai *similarity*. Proses perhitungan nilai *similarity* dilakukan dengan bertujuan untuk melihat tingkat kemiripan yang dimiliki oleh masing-masing senyawa dengan harapan senyawa yang ada antar cluster berbeda, namun senyawa yang berada dalam satu cluster mirip atau sama. Kemiripan yang dilihat pada suatu senyawa yaitu kemiripan dari strukturnya. Perhitungan nilai *similarity* dilakukan dengan menggunakan pendekatan Tanimoto *similarity*. Tanimoto *similarity* yaitu merupakan pendekatan yang cukup populer untuk menghitung kemiripan dari berbagai data. Pendekatan Tanimoto *similarity* digunakan karena lebih sering digunakan untuk mencari nilai *similarity* dalam dua senyawa dalam bidang bioinformatika dan mudah di implementasikan (Bajusz et al., 2015). Nilai *similarity* yang diperoleh diharapkan berbeda pada setiap antar *cluster* nilai *similarity*-nya, tetapi nilai *similarity* pada satu *cluster* besar (Haranczyk & Holliday, 2008). Pengecekan nilai *similarity* setiap antar *cluster* dilakukan dengan menggunakan 50 senyawa yang mewakili pada setiap masing-masing *cluster*. Hasil perhitungan *similarity* intra dan *inter cluster* ditunjukkan pada gambar 4.4 sebagai berikut.



Gambar 4.4 Hasil *Similarity Intra* dan *Inter Cluster* dari 50 *Cluster*

Gambar 4.4 menjelaskan hasil perhitungan nilai *similarity intra-cluster* dan *inter-cluster* dari 50 *cluster* yang diwakili oleh masing-masing senyawa pada setiap *cluster*. Hasil dari perhitungan nilai *similarity* pada *intra-cluster* menunjukkan bahwa mayoritas senyawa memiliki kemiripan yang cukup besar antar tiap anggota pada suatu *cluster*. Hasil dari perhitungan *similarity inter-cluster* menjelaskan bahwa perbedaan antara suatu *cluster* dan *cluster* lainnya nilainya semakin kecil yang berarti semakin baik kualitas *clustering*. Pendekatan Tanimoto *similarity* yang sering digunakan untuk hasil kemiripan yang ditentukan dengan menggunakan *threshold* dan *cutoff* minimal sebesar 0.6 (Park et al.,



2017). Hasil pada gambar 4.4 menunjukkan nilai *similarity* antar *cluster* berbeda, dan rentang nilai *similarity*-nya kurang dari 0.6. Hal tersebut mengartikan bahwa jumlah frekuensi perhitungan *similarity* pada setiap *cluster* yang memiliki kemiripan di atas 0.6 ialah sebanyak 50 *cluster*.

Hasil yang diperoleh dari perhitungan nilai *similarity* pada 50 *cluster* dilakukan proses uji ADMET (Lampiran 4.5). ADMET merupakan suatu alat komputasi yang digunakan untuk memprediksi sifat-sifat dari obat-obatan dengan menggunakan metode pembelajaran mesin dan model QSAR. ADMET yang digunakan pada penelitian ini adalah ADMET-pkCSM, dimana pada ADMET-pkCSM kita dapat mengetahui seberapa efektif dan aman obat tersebut digunakan, termasuk bagaimana obat tersebut diserap oleh tubuh, bagaimana di distribusikan di seluruh tubuh, bagaimana diolah dan dikeluarkan oleh tubuh, serta potensi efek beracun yang mungkin dimiliki oleh obat tersebut (Sucharitha et al., 2022). Berikut hasil Uji ADMET-pkCSM dari 50 *cluster*

Tabel 4.4 Hasil Analisis Sifat Mirip Obat Berdasarkan Aturan Lipinski

Aturan Lipinski				
<i>Cluster</i>	Massa Molekul	LogP	HBA	HBD
15	411.374	2.498	9	2
41	417.853	2.030	6	1
8	405.173	2.395	5	1
16	407.404	2.217	7	2
18	422.276	2.259	5	1
12	410.496	2.101	6	2
38	416.272	2.497	6	1
21	420.466	2.035	6	2
34	414.528	2.051	7	2
40	418.494	2.176	5	2

Hasil prediksi Uji ADMET-pkCSM nilai parameter-parameter sifat fisikokimia *cluster* 15,41,8,16,18,12,38,21,34, dan 40 dapat dilihat pada tabel 4.4. Lipinski et al.(1997) telah menganalisis 2.245 obat dari data dasar World Drugs

Index dan menyimpulkan bahwa senyawa akan sulit diabsorpsi dan permeabilitasnya rendah apabila berat molekulnya lebih besar 500, mempunyai nilai log koefisien partisi oktanol/air (log P) lebih besar +5; mempunyai ikatanH donor (HBD), yang dinyatakan dengan jumlah gugus O-H dan N-H, lebih besar 5; dan mempunyai ikatan-H aseptor (HBA), yang dinyatakan dengan jumlah atom O dan N, lebih besar 10. Analisis di atas dikenal sebagai hukum lima Lipinski karena semua nilai merupakan kelipatan dari angka lima (D. E. V Pires et al., 2015). Dari tabel 4.4 dapat dianalisis bahwa *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 memiliki berat molekul kurang dari 500, nilai logP kurang dari 5, nilai aseptor dan donor kurang dari 10, sehingga dapat disimpulkan bahwa *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 mudah diabsorpsi.

Tabel 4.5 Prediksi Uji ADMET-pkCSM

ADMET	Cluster				
	15	41	8	16	18
<i>Intestinal absorption (human)(%)</i>	90.211	81.582	85.398	85.276	83.724
<i>Skin Permeability (log Kp)</i>	-2.736	-2.854	-2.81	-2.749	-2.843
<i>VDss (human) (log L/kg)</i>	-0.273	-0.411	-0.691	-0.572	0.384
<i>BBB permeability (log BB)</i>	-0.883	-0.661	-0.974	-1.02	-1.059
<i>CYP3A4 substrate (Ya/Tidak)</i>	Ya	Ya	Tidak	Ya	Tidak
<i>CYP3A4 Inhibitor (Ya/Tidak)</i>	Ya	Ya	Tidak	Ya	Tidak
<i>Total Clearance (log ml/min/kg)</i>	0.566	0.237	-0.303	0.638	0.309
<i>Renal OCT2 substrate (Ya/Tidak)</i>	Tidak	Tidak	Tidak	Tidak	Tidak
<i>AMES toxicity (Ya/Tidak)</i>	Tidak	Tidak	Tidak	Tidak	Tidak
<i>LD50 (Mol/Kg)</i>	2.365	2.489	2.491	2.881	2.426

ADMET	Cluster				
	12	38	21	34	40
<i>Intestinal absorption (human)(%)</i>	79.858	77.705	76.134	72.883	66.894



<i>Skin Permeability</i> (log Kp)	-2.888	-2.809	-2.861	-2.738	-2.963
VDss ( <i>human</i> ) (log L/kg)	-0.268	-0.234	-0.549	0.305	-0.256
BBB permeability (log BB)	-0,712	-0.664	-0.795	-0.913	-0.76
CYP3A4 substrate (Ya/Tidak)	Ya	Tidak	Ya	Ya	Ya
CYP3A4 Inhibitor (Ya/Tidak)	Tidak	Ya	Tidak	Tidak	Tidak
Total Clearance (log ml/min/kg)	1.219	0.376	0.866	0.61	1.474
Renal OCT2 substrate (Ya/Tidak)	Tidak	Tidak	Tidak	Tidak	Tidak
AMES toxicity (Ya/Tidak)	Tidak	Tidak	Tidak	Tidak	Tidak
LD50 (Mol/Kg)	2.706	2.629	2.474	2.757	2.436

Hasil prediksi uji ADMET-pkCSM pada *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 dapat dilihat pada tabel 4.5 tersebut. Menurut Chander et al. (2017), senyawa dikatakan mempunyai absorpsi yang baik bila nilai absorpsinya > 80%, dan absorpsinya kurang baik bila < 30%. Usus merupakan tempat utama untuk penyerapan obat yang diberikan secara oral (Chander et al., 2017). Dari tabel 4.5 dapat dilihat bahwa nilai *intestinal absorption (human)* dari *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 lebih dari 80% dan tidak kurang dari 30%, menunjukkan bahwa *cluster* tersebut mempunyai absorpsi yang baik. Menurut Pires et al. (2015), senyawa dikatakan mempunyai permeabilitas kulit yang relatif rendah bila mempunyai nilai log Kp > -2,5. Dari tabel 4.5 dapat dilihat bahwa nilai *Skin Permeability* (log Kp) dari *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 mempunyai nilai log Kp > -2,5, sehingga dapat diprediksi bahwa *cluster* tersebut mempunyai permeabilitas kulit yang baik. Bahan obat yang mempunyai permeabilitas kulit baik dapat digunakan sebagai pertimbangan untuk pengembangan obat baru dengan pemberian transdermal (D. E. V. Pires et al., 2015).

Volume distribusi (VDss) adalah volume teoritis bahwa dosis total obat perlu didistribusikan secara merata untuk memberikan konsentrasi yang sama seperti pada plasma darah. Semakin tinggi nilai VD, semakin banyak obat yang

didistribusikan di jaringan daripada plasma. Menurut Pires et al. (2015), senyawa dikatakan mempunyai Volume Distribusi rendah bila nilai Log VD < - 0,15, dan tinggi bila > 0,45. Dari tabel 4.5 dapat dilihat bahwa nilai VDss (*Steady State of Volume Distribution*) dari *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 mempunyai nilai VDss < -0,15, sehingga dapat diprediksi bahwa semua senyawa tersebut dapat didistribusikan secara merata untuk memberikan konsentrasi yang sama seperti pada plasma darah.

Kemampuan obat untuk menembus sawar darah otak (Blood Brain Barrier) merupakan parameter penting yang perlu dipertimbangkan untuk membantu mengurangi efek samping dan toksisitas atau untuk meningkatkan khasiat obat yang aktivitas farmakologisnya ada di dalam otak. Permeabilitas otak-darah diukur secara *in vivo* pada model hewan sebagai logBB, yaitu rasio logaritmik konsentrasi pada otak terhadap plasma. Menurut Pires et al. (2015), senyawa dikatakan mampu menembus sawar darah otak dengan baik bila mempunyai nilai Log BB > 0,3, dan tak dapat terdistribusi dengan baik bila log BB < -1. Dari tabel 4.5 dapat dilihat bahwa nilai log BB dari *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 lebih besar dari -1, sehingga dapat diprediksi bahwa semua turunan senyawa tersebut mampu menembus sawar darah otak secara moderat.

Secara umum diketahui bahwa sebagian besar reaksi metabolik akan melibatkan proses oksidasi. Sitokrom P450 adalah enzim detoksifikasi penting dalam tubuh, dan terutama ditemukan di hati. Bekerja dengan cara mengoksidasi senyawa organik asing, termasuk obat, dan memfasilitasi ekskresi senyawa tersebut. Inhibitor enzim ini dapat mempengaruhi metabolisme obat sehingga dikontraindikasikan terhadap enzim sitokrom P450. Oleh karena itu penting untuk menilai kemampuan senyawa yang dapat menghambat sitokrom P450, yang dalam penelitian ini diwakili oleh isoform sitokrom CYP3A4. Dari tabel 4.5 dapat dilihat bahwa *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 tidak mempengaruhi atau menghambat enzim CYP3A4, sehingga dapat diprediksi senyawa turunan tersebut cenderung dimetabolisme oleh enzim P450.

Memprediksi proses ekskresi senyawa dapat dilakukan dengan mengukur

tetapan Total Clearance (CLTOT) dan Renal Organic Cation Transporter 2(OCT2) substrate. CLTOT merupakan kombinasi dari hepatic clearance (metabolisme di hati dan empedu) dan renal clearance (ekskresi melalui ginjal). Hal ini terkait dengan bioavailabilitas, dan penting untuk menentukan tingkat dosis dalam mencapai konsentrasi *steady-state*. Dari tabel 4.5 dapat dilihat bahwa nilai CLTOT *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 dapat diprediksi kecepatan ekskresi senyawanya.

*Organic Cation Transporter 2* adalah transporter pada ginjal yang memegang peran penting dalam disposisi dan clearance obat-obatan dan senyawa endogen. Substrat OCT2 juga berpotensi menimbulkan interaksi sampingan bila diberikan bersama-sama dengan inhibitor OCT2. Dari tabel 4.5 dapat dilihat bahwa *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 tidak mempengaruhi substrat OCT2, sehingga dapat diprediksi senyawa turunan tersebut bukan merupakan substrat OCT2. Menentukan toksisitas senyawa dapat dilakukan dengan uji Ames *Toxicity*. Uji Ames *Toxicity* adalah metode yang digunakan secara luas untuk menilai potensi mutagenik senyawa dengan menggunakan bakteri. Hasil uji *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40 negatif, menunjukkan bahwa senyawa tersebut bersifat tidak menimbulkan efek mutagenik yaitu sebutan untuk zat yang menyebabkan kanker.

## BAB 5 PENUTUP

### 5.1 Kesimpulan

Kesimpulan yang dapat diperoleh berdasarkan penelitian yang dilakukan adalah sebagai berikut:

1. Pemodelan QSAR dari 429 senyawa dari data ChEMBL1781 kemudian didapatkan data training sebanyak 300 senyawa dan data testing 129 senyawa. Data testing yang diperoleh kemudian dilakukan variasi hyperparameter pada *Hidden Layer* (L1,L2, dan L3) dan *batch size* secara berurutan yaitu 8; 16; 32 dan 8; 1024; 2048 dengan *nb\_epoch* 50 sehingga didapatkan model terbaik dengan R<sup>2</sup> 0,66 yang dapat digunakan untuk proses virtual screening.
2. Proses screening yang diperoleh dari 481 juta senyawa dari data ZINC20 diperoleh 335 senyawa pada prediksi pIC<sub>50</sub> obat Topotecan menjadi 219 cluster yang lolos. Data 219 senyawa cluster selanjutnya dilakukan cluster center sehingga didapatkan 50 cluster.
3. Data 50 cluster digunakan untuk uji sintesis pada ADMET pk-CSM yang lolos uji toksisitas sebanyak 10 senyawa yaitu *cluster* 15, 41, 8, 16, 18, 12, 38, 21, 34, dan 40.

### 5.2 Saran

Molekul calon inhibitor Topoisomerase-1 yang diperoleh dari hasil screening perlu dilakukan pengujian dan penelitian lebih lanjut. Pengujian secara *in vitro* diperlukan untuk mengetahui efektivitas molekul hasil screening dalam menghambat protein Topoisomerase-1.

## DAFTAR PUSTAKA

- Alexander, D. L. J., Tropsha, A., & Winkler, D. A. (2015). Be ware of R2: simple, unambiguous assessment of the prediction accuracy of QSAR and QSPR models. *Journal of Chemical Information and Modeling*, 55(7), 1316. <https://doi.org/10.1021/ACS.JCIM.5B00206>
- Bajusz, D., Rácz, A., & Héberger, K. (2017). Chemical Data Formats, Fingerprints, and Other Molecular Descriptions for Database Analysis and Searching. *Comprehensive Medicinal Chemistry III*, 3–8, 329–378. <https://doi.org/10.1016/B978-0-12-409547-2.12345-5>
- Berrouet, C., Dorilas, N., Rejniak, K. A., & Tuncer, N. (2020). Comparison of Drug Inhibitory Effects (IC 50) in Monolayer and Spheroid Cultures. *Bulletin of Mathematical Biology*, 82(6). <https://doi.org/10.1007/s11538-020-00746-7>
- Belete, T. M. (2021). The Current Status of Gene Therapy for the Treatment of Cancer. *Biologics : Targets & Therapy*, 15, 67. <https://doi.org/10.2147/BTT.S302095>
- Benita, S. Mardiah, N. Nurvita. (2020). Analisis Implementasi Program Deteksi Dini Kanker Serviks Dengan Metode Inspeksi Visual Asam Asetat (Iva). *Asian Research Of Midwifery Basic Science Journal* (Vol. 1, pp. 1-12). <https://doi.org/10.37160/arimbi.v1i1.532>
- Bukowski, K., Kciuk, M., & Kontek, R. (2020). Mechanisms of Multidrug Resistance in Cancer Chemotherapy. *International Journal of Molecular Sciences*, 21(9), 1–24.
- Bozzone, D. M. (2009). *Leukemia*. Infobase Publishing. [https://books.google.co.id/books?id=wxVbEUQAhMC&printsec=frontcover&hl=id&source=gbs\\_ge\\_summary\\_r&cad=0#v=o nepage&q&f=false](https://books.google.co.id/books?id=wxVbEUQAhMC&printsec=frontcover&hl=id&source=gbs_ge_summary_r&cad=0#v=o nepage&q&f=false)
- Carpenter, K. A., & Huang, X. (2018). Machine Learning-based Virtual Screening and Its Applications to Alzheimer's Drug Discovery: A Review. *Current Pharmaceutical Design*, 24(28), 3347–3358.



<https://doi.org/10.2174/1381612824666180607124038>

Chander, S., Tang, C. R., Al-Maqtari, H. M., Jamalis, J., Penta, A., Hadda, T. Ben, ... Sankaranarayanan, M. (2017). Synthesis and study of anti-HIV-1 RT activity of 5-benzoyl-4-methyl-1,3,4,5-tetrahydro-2H-1,5-benzodiazepin-2-one derivatives. *Bioorganic Chemistry*, 72, 74–79.

<https://doi.org/10.1016/j.bioorg.2017.03.013>

Changela, A; Digate, R. J.; Mondragón, As (2001). "Crystal structure of a complex of a type IA DNA topoisomerase with a single-stranded DNA molecule". *Nature*. 411 (6841):107781. [Bibcode:2001Natur.411.1077C](#). [doi:10.1038/35082615](#)

Cereto-Massagué, A., Ojeda, M. J., Valls, C., Mulero, M., Garcia-Vallvé, S., & Pujadas, G. (2015). Molecular fingerprint similarity search in virtual screening. *Methods*, 71(C), 58–63.

<https://doi.org/10.1016/J.YMETH.2014.08.005>

Champoux J.J. (2003). DNA topoisomerases: structure, function, and mechanism. *Annu. Rev. Biochem.* 70: 369–413. [doi:10.1146/annurev.biochem.70.1.369](#).

DeFilippis; Rosa, Anna; Edward, Goodwin; Lingling Wu and Daniel DiMaio. (2003). Endogenous Human Papillomavirus E6 and E7 Proteins Differentially Regulate Proliferation, Senescence, and Apoptosis in HeLa Cervical Carcinoma Cells. *Journal of Virology*. Volume 77. nomor 2. <https://doi.org/10.1128/JVI.77.2.1551-1563.2003>

Delano, W. L., & Bromberg, S. (2004). PyMOL User's Guide. *DeLano Scientific LLC*, 1–66.

Diananda, R. (2008). *Mengenal Seluk Beluk Kanker*. Mengenal Seluk Beluk Kanker. <http://sastramangutama.badungkab.go.id/inlislite3/opac/detail-opac?id=11673>

Elmore, Susan (2007). "Apoptosis: a review of programmed cell death". *Toxicologic Pathology*. 35 (4):495516. [doi:10.1080/01926230701320337](#)

Goodman, S. R. (2008). *Medical cell biology*. 320.

Gramatica, P. (2013). On the Development and Validation of QSAR Models. In

- Methods in Molecular Biology* (Vol. 930, pp. 499–526).  
[https://doi.org/10.1007/978-1-62703-059-5\\_21](https://doi.org/10.1007/978-1-62703-059-5_21)
- Hanahan, D., & Weinberg, R. A. (2011). Hallmarks of cancer: the next generation. *Cell*, 144(5), 646–674. <https://doi.org/10.1016/J.CELL.2011.02.013770-776>. <https://doi.org/10.1038/35037710>
- Haranczyk, M., & Holliday, J. (2008). Comparison of similarity coefficients for clustering and compound selection. *Journal of Chemical Information and Modeling*, 48(3), 498–508. <https://doi.org/10.1021/ci700413a>
- Hermansyah, O., Bustamam, A., & Yanuar, A. (2021). Virtual screening of dipeptidyl peptidase-4 inhibitors using quantitative structure–activity relationship-based artificial intelligence and molecular docking of hit compounds. *Computational Biology and Chemistry*, 95(April), 107597. <https://doi.org/10.1016/j.compbiolchem.2021.107597>
- Höltje, H.-Dieter. (2008). *Molecular modeling: basic principles and applications*.310.
- Horner, Stacy; Ros, Anna; DeFilippis; Laertes, Manuelidis and Daniel DiMaio. (2004). Repression of the Human Papillomavirus E6 Gene Initiates p53-Dependent, Telomerase-Independent Senescence and Apoptosis in HeLa Cervical Carcinoma Cells. *Journal of Virology*. Volume 78. Nomor 8. <https://doi.org/10.1128/JVI.78.8.4063-4073.2004>
- Irwin, J. J., Tang, K. G., Young, J., Dandarchuluun, C., Wong, B. R., Khurelbaatar, M., Moroz, Y. S., Mayfield, J., & Sayle, R. A. (2020). ZINC20- A Free Ultralarge-Scale Chemical Database for Ligand Discovery. *Journal of Chemical Information and Modeling*, 60(12), 6065–6073. <https://doi.org/10.1021/ACS.JCIM.0C00675>
- Kantarjian, H., Kadia, T., DiNardo, C., Daver, N., Borthakur, G., Jabbour, E., Garcia-Manero, G., Konopleva, M., & Ravandi, F. (2021). Acute myeloid leukemia: current progress and future directions. *Blood Cancer Journal*, 11(2). <https://doi.org/10.1038/s41408-021-00425-3>
- Kemmish, H., Fasnacht, M., & Yan, L. (2017). Fully automated antibody structure prediction using BIOVIA tools: Validation study. *PloS One*, 12(5).



<https://doi.org/10.1371/JOURNAL.PONE.0177923>

- Leach, A. (2001). Molecular modelling : principles and applications. In *Molecular Modelling Principles and Applications* (2nd ed.). Prentice Hall.
- Lipinski, C. A., Lombardo, F., Dominy, B. W., & Feeney, P. J. (1997). Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings, 23.
- Mao, J., Akhtar, J., Zhang, X., Sun, L., Guan, S., Li, X., Chen, G., Liu, J., Jeon, H. N., Kim, M. S., No, K. T., & Wang, G. (2021). Comprehensive strategies of machine-learning-based quantitative structure-activity relationship models. *IScience*, 24(9). <https://doi.org/10.1016/J.ISCI.2021.103052>
- Mendez, D., Gaulton, A., Bento, A. P., Chambers, J., de Veij, M., Félix, E., Magariños, M. P., Mosquera, J. F., Mutowo, P., Nowotka, M., Gordillo-Marañón, M., Hunter, F., Junco, L., Mugumbate, G., Rodriguez-Lopez, M., Atkinson, F., Bosc, N., Radoux, C. J., Segura-Cabrera, A., Leach, A. R. (2019). ChEMBL: Towards direct deposition of bioassay data. *Nucleic Acids Research*, 47(D1), D930–D940. <https://doi.org/10.1093/NAR/GKY1075>
- Mohan, V., Gibbs, A., Cummings, M., Jaeger, E., & DesJarlais, R. (2005). Docking: successes and challenges. *Current Pharmaceutical Design*, 11(3), 323–333. <https://doi.org/10.2174/1381612053382106>
- Morris, G. M., Huey, R., & Olson, A. J. (2008). Using AutoDock for ligand-receptor docking. *Current Protocols in Bioinformatics*, Chapter 8(SUPPL. 24). <https://doi.org/10.1002/0471250953.BI0814S24>
- Neves, B. J., Braga, R. C., Melo-Filho, C. C., Moreira-Filho, J. T., Muratov, E. N., & Andrade, C. H. (2018). QSAR-based virtual screening: Advances and applications in drug discovery. *Frontiers in Pharmacology*, 9(NOV), 1275. <https://doi.org/10.3389/FPHAR.2018.01275/BIBTEX>
- O. Chakrabati, S. Krishna. (2003). Molecular interactions of 'high risk' human papillomaviruses E6 and E7 oncoproteins: implications for tumour progression. *Journal Of Biosciences*, (Vol. 28, pp. 337-348). <https://pubmed.ncbi.nlm.nih.gov/12734411/>

- Park, S. H., Haddad, P. R., Amos, R. I. J., Talebi, M., Szucs, R., Pohl, C. A., & Dolan, J. W. (2017). Towards a chromatographic similarity index to establish localised Quantitative Structure-Retention Relationships for retention prediction. III Combination of Tanimoto similarity index, logP, and retention factor ratio to identify optimal analyte training . *Journal of Chromatography A*, 1520, 107–116.  
<https://doi.org/10.1016/j.chroma.2017.09.016>
- Pires, D. E. V., Blundell, T. L., & Ascher, D. B. (2015). pkCSM: Predicting smallmolecule pharmacokinetic and toxicity properties using graph-based signatures. *Journal of Medicinal Chemistry*, 58(9), 4066–4072.  
<https://doi.org/10.1021/acs.jmedchem.5b00104>
- Pires, D. E. V, Blundell, T. L., & Ascher, D. B. (2015). pkCSM : predicting smallmolecule pharmacokinetic properties using graph-based signatures (Theory- How to Enterpret pkCSM Result). *PKCSM*, 5. Retrieved from  
<http://biosig.unimelb.edu.au/pkcsm/theory>
- Pranowo, H. D. (2011). *Pengantar Kimia Komputasi*. 4-5:118.
- Purnomo, H. (2011). *Kimia Komputasi : Molecular Docking Plants-Penambatan Molekul Plants [Protein-Ligand-Ant-System]* . Pustaka Pelajar.
- Ray, S. (2014). The Cell: A Molecular Approach. *The Yale Journal Of Biology and Medicine*, (Vol. 87, pp. 603). /pmc/articles/PMC4257047/.  
<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4257047/>
- Redinbo, M.R., Stewart, L., Kuhn, P., Champoux, J.J., Hol, W.G. (1998). Crystal Structures of Human Topoisomerase I in Covalent and Noncovalent Complexes with DNA. *Science* 279:1504-1513.  
<https://www.science.org/doi/10.1126/science.279.5356.1504>
- Rogers, D., & Hahn, M. (2010). Extended-connectivity fingerprints. *Journal of Chemical Information and Modeling*, 50(5), 742–754.  
[https://doi.org/10.1021/CI100050T/ASSET/IMAGES/MEDIUM/CI-2010-00050T\\_0018.GIF](https://doi.org/10.1021/CI100050T/ASSET/IMAGES/MEDIUM/CI-2010-00050T_0018.GIF)
- Ruddon, R. W. (1998). Cancer biology. *Current Opinion in Oncology*, 10(1), 530.
- Septia Haryani, Defrin Defrin, Yenita Yenita. (2016). Prevalensi Kanker Serviks

- Berdasarkan Paritas di RSUP. Dr. M. Djamil Padang Periode Januari 2011-Desember 2012. *Jurnal Kesehatan Andalas*. doi: <https://doi.org/10.25077/jka.v5i3.592>
- Stewart, M., & Watson, I. (1983). Standard units for expressing drug concentrations in biological fluids. *British Journal of Clinical Pharmacology*, 16(1), 3–7. <https://doi.org/10.1111/J.1365-2125.1983.TB02136.X>
- Sucharitha, P., Ramesh Reddy, K., Satyanarayana, S. V., & Garg, T. (2022). Absorption, distribution, metabolism, excretion, and toxicity assessment of drugs using computational tools. In *Computational Approaches for Novel Therapeutic and Diagnostic Designing to Mitigate SARS-CoV2 Infection: Revolutionary Strategies to Combat Pandemics*. INC. <https://doi.org/10.1016/B978-0-323-91172-6.00012-1>
- Sourav Pal, Vinay Kumar, Biswajit Kundu, Debomita Bhattacharya, Nagothy sPreety, Mamindla Prashanth Reddy, Arindam Talukdar. (2019). "[Ligand-based Pharmacophore Modeling, Virtual Screening and Molecular Docking Studies for Discovery of Potential Topoisomerase I Inhibitors](#)". *Comput Struct Biotechnol J*. 2019; 17: 291–310. <https://doi.org/10.1016%2Fj.csbj.2019.02.006>
- Taneja B, Patel A, Slesarev A, Mondragón A (January 2006). "[Structure of the N-terminal fragment of topoisomerase V reveals a new family of topoisomerases](#)". *EMBO J*. 25 (2): 398–408. doi:10.1038/sj.emboj.7600922
- Verma, R. P., & Hansch, C. (2011). Use of <sup>13</sup>C NMR chemical shift as QSAR/QSPR descriptor. *Chemical Reviews*, 111(4), 2865–2899. <https://doi.org/10.1021/CR100125D>
- Wang JC (June 2002). "Cellular roles of DNA topoisomerases: a molecular perspective". *Nat. Rev. Mol. Cell Biol*. 3 (6): 430–40. doi:10.1038/nrm831
- Zahoránszky-Kohalmi, G., Bologa, C. G., & Oprea, T. I. (2016). Impact of similarity threshold on the topology of molecular similarity networks and clustering outcomes. *Journal of Cheminformatics*, 8(1), 1–17. <https://doi.org/10.1186/s13321-016-0127-5>

LAMPIRAN

Lampiran 4.1 *Datase*



Lampiran 4.2 Hasil Pemodelan QSAR



Lampiran 4.3 Hasil *Virtual Screening* dengan *Database Zinc20*



Lampiran 4.4 Hasil *Clustering*



Lampiran 4.5 Hasil Uji ADMET-pkCSM

