



PENAPISAN SENYAWA HERBAL DENGAN KOMPUTASI GPU PADA BASIS DATA ZINC MELALUI PENDEKATAN PEMBANDINGAN SIMILARITAS

REFIANTO DAMAI DARMAWAN



**DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2021**

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

PERNYATAAN MENGENAI SKRIPSI DAN SUMBER INFORMASI SERTA PELIMPAHAN HAK CIPTA

Dengan ini saya menyatakan bahwa skripsi dengan judul “Penapisan Senyawa Herbal dengan Komputasi GPU pada Basis Data ZINC melalui Pendekatan Perbandingan Similaritas” adalah karya saya dengan arahan dari dosen pembimbing dan belum diajukan dalam bentuk apa pun kepada perguruan tinggi mana pun. Sumber informasi yang berasal atau dikutip dari karya yang diterbitkan maupun tidak diterbitkan dari penulis lain telah disebutkan dalam teks dan dicantumkan dalam Daftar Pustaka di bagian akhir skripsi ini.

Dengan ini saya melimpahkan hak cipta dari karya tulis saya kepada Institut Pertanian Bogor.

Bogor, Juni 2021

Refianto Damai Darmawan
G64170001

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



ABSTRAK

REFIANTO DAMAI DARMAWAN. Penapisan Senyawa Herbal dengan Komputasi GPU pada Basis Data ZINC melalui Pendekatan Perbandingan Similaritas. Dibimbing oleh WISNU ANANTA KUSUMA dan HENDRA RAHMAWAN.

Covid-19 adalah sebuah pandemi yang membuat banyak peneliti berlomba mencari solusi untuk kemaslahatan sesama, terutama dalam segi kesehatan, pengobatan, dan penanggulangan penyakit tersebut. Penapisan awal dengan proses *in-silico* berperan penting dalam mengurangi jumlah bahan kimia yang perlu diuji untuk menemukan suatu obat. Penelitian ini bertujuan untuk mencari kandidat senyawa obat terbaik yang ada dalam basis data ZINC sehingga bisa diteliti lebih lanjut secara *in-vitro*. Sebanyak 997.402.117 senyawa kimia dicari kemiripannya dengan beberapa senyawa obat yang sudah teruji untuk melawan virus corona. Komputasi sekuensial akan memakan banyak waktu beberapa bulan untuk menyelesaikannya. Pendekatan *general programming graphic processing unit* digunakan melalui implementasi algoritme perbandingan similaritas secara paralel sehingga dapat mempercepat proses pencarian. Hasil dari penelitian ini menunjukkan bahwa implementasi paralel algoritme ini dapat mempercepat proses komputasi hingga 55 kali, serta beberapa kandidat senyawa herbal yang dapat ditemukan pada bahan-bahan alam, memiliki nilai similaritas yang tinggi dengan obat coronavirus.

Kata kunci: covid-19, perbandingan similaritas, pemrograman GPU, pemrograman paralel.

ABSTRACT

REFIANTO DAMAI DARMAWAN. *Herbal Compound Screening with GPU Computation on ZINC Database through Sequence Alignment Approach*. Supervised by WISNU ANANTA KUSUMA and HENDRA RAHMAWAN.

Covid-19 is a global pandemic that drives many researcher strive to look for its solution, especially in the field of health, medicine, and total countermeasures. Early screening with in-silico processes is crucial to minimize the search space of the potential drugs to cure a disease. This research aims to find potential drugs of covid-19 disease in ZINC database to be further investigated through in-vitro method. About 997.402.117 chemical compounds are searched about its similarity to some of confirmed drugs to combat coronavirus. Sequential computation would take months to accomplish this task. General programming graphic processing unit approach is used to implement similarity comparison algorithm in parallel, in order to speed up the process. The result of this study shows the parallel algorithm implementation can speed-up the computation process up to 55 times faster, and also that some of the chemical compounds have high similarity score and can be found in nature.

Keywords: covid-19, GPU programming, parallel programming, similarity comparison.

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

© Hak Cipta milik IPB, tahun 2021
Hak Cipta dilindungi Undang-Undang

Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan atau menyebutkan sumbernya. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik, atau tinjauan suatu masalah, dan pengutipan tersebut tidak merugikan kepentingan IPB.

Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apa pun tanpa izin IPB.



PENAPISAN SENYAWA HERBAL DENGAN KOMPUTASI GPU PADA BASIS DATA ZINC MELALUI PENDEKATAN PEMBANDINGAN SIMILARITAS

REFIANTO DAMAI DARMAWAN

Skripsi
sebagai salah satu syarat untuk memperoleh gelar
Sarjana pada
Program Studi Ilmu Komputer

**DEPARTEMEN ILMU KOMPUTER
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
INSTITUT PERTANIAN BOGOR
BOGOR
2021**

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Tim Penguji pada Ujian Skripsi:

1 Auriza Rahmad Akbar, S.Komp., M.Kom. (Ketua)



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Judul Skripsi : Penapisan Senyawa Herbal dengan Komputasi GPU pada Basis Data ZINC melalui Pendekatan Perbandingan Similaritas
Nama : Refianto Damai Darmawan
NIM : G64170001

Disetujui oleh

Pembimbing 1:
Dr.Eng. Wisnu Ananta Kusuma, S.T., M.T.

Pembimbing 2:
Dr. Hendra Rahmawan, S.Kom., M.T.



Diketahui oleh

Ketua Departemen Ilmu Komputer:
Dr. Sony Hartono Wijaya, S.Kom., M.Kom.
19810809 200812 1 002



Tanggal Ujian:
16 Juni 2021

Tanggal Lulus:
(tanggal penandatanganan oleh Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam)



PRAKATA

Puji dan syukur penulis panjatkan kepada Allah subhanaahu wa ta'ala atas segala karunia-Nya sehingga karya ilmiah ini berhasil diselesaikan. Tema yang dipilih dalam penelitian yang dilaksanakan sejak bulan Desember 2020 sampai bulan Juni 2021 ini ialah utilitas GPU untuk optimasi bioinformatika, dengan judul “Penapisan Senyawa Herbal dengan Komputasi GPU pada Basis Data ZINC melalui Pendekatan Perbandingan Similaritas”.

Terima kasih penulis ucapkan kepada para pembimbing, Dr.Eng. Wisnu Ananta Kusuma, S.Si, M.Kom. dan Dr. Hendra Rahmawan, S.Kom., M.T. yang telah membimbing dan banyak memberi saran. Ucapan terima kasih juga disampaikan kepada pembimbing akademik, moderator seminar, dan penguji luar Komisi pembimbing. Di samping itu, penghargaan penulis sampaikan kepada Dr.Eng. Wisnu Ananta Kusuma, S.Si, M.Kom. beserta staf Laboratorium Biofarmaka IPB yang telah membantu selama pengolahan data. Ungkapan terima kasih juga disampaikan kepada ayah, ibu, serta seluruh keluarga yang telah memberikan dukungan, doa, dan kasih sayangnya.

Semoga karya ilmiah ini bermanfaat bagi pihak yang membutuhkan dan bagi kemajuan ilmu pengetahuan..

Bogor, Juni 2021

Refianto Damai Darmawan

DAFTAR ISI

DAFTAR TABEL	x
DAFTAR GAMBAR	x
DAFTAR LAMPIRAN	x
I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	2
1.3 Tujuan	2
1.4 Manfaat	3
1.5 Ruang Lingkup Penelitian	3
II TINJAUAN PUSTAKA	4
2.1 Penapisan Senyawa berbasis Ligan	4
2.2 Senyawa Herbal	4
2.3 <i>Graphic Processing Unit</i>	4
2.4 Basis Data ZINC	6
2.5 Algoritme Perbandingan Similaritas	6
III METODE	8
3.1 Tahapan Penelitian	8
3.2 Data Penelitian	8
3.3 Pengolahan <i>Fingerprint</i>	9
3.4 Penentuan Ukuran Blok	11
3.5 Penentuan Jumlah Aliran (<i>Streamsize</i>)	12
3.6 Perbandingan Similaritas	12
3.7 Evaluasi Algoritme	16
3.8 Aplikasi pada Basis Data ZINC	16
3.9 Lingkungan Pengembangan	16
IV HASIL DAN PEMBAHASAN	18
4.1 Pengolahan <i>Fingerprint</i>	18
4.2 Penentuan Ukuran Blok	20
4.3 Penentuan Jumlah Aliran (<i>Streamsize</i>)	21
4.4 Perbandingan Similaritas	22
4.5 Aplikasi pada Basis Data ZINC	24
V SIMPULAN DAN SARAN	28
5.1 Simpulan	28
5.2 Saran	28
DAFTAR PUSTAKA	29
LAMPIRAN	32
RIWAYAT HIDUP	49



DAFTAR TABEL

1	Daftar senyawa obat yang digunakan sebagai acuan	8
2	Hasil Tiga teratas perbandingan similaritas pada seluruh basis data ZINC	24

DAFTAR GAMBAR

1	Peran Ligan dalam interaksi virus dan sel (Sethi <i>et al.</i> 2020)	2
2	Ilustrasi desain arsitektur CPU dan GPU	5
3	Contoh file unduhan dari basis data ZINC	6
4	Diagram alir tahapan penelitian	8
5	Data yang diunduh dari basis data ZINC	10
6	Diagram alir proses pengolahan <i>fingerprint</i>	10
7	Contoh representasi <i>fingerprint</i> PubChem	11
8	Pseudocode untuk algoritme sekuensial CPU pengolahan <i>fingerprint</i>	11
9	Pseudocode untuk algoritme paralel CPU pengolahan <i>fingerprint</i>	11
10	Pseudocode untuk algoritme penentuan ukuran blok	12
11	Pseudocode untuk algoritme paralel CPU pengolahan <i>fingerprint</i>	12
12	Ilustrasi penghitungan similaritas tanimoto	13
13	Representasi proses perbandingan secara sekuensial	13
14	Pseudocode untuk algoritme sekuensial CPU perbandingan similaritas	14
15	Representasi proses perbandingan secara paralel	15
16	Pseudocode untuk algoritme paralel CPU perbandingan similaritas	16
17	Grafik waktu pengolahan <i>fingerprint</i> MACCS (paralel CPU)	19
18	Grafik nilai <i>speed-up</i> pengolahan <i>fingerprint</i> MACCS (paralel CPU)	19
19	Grafik waktu pengolahan <i>fingerprint</i> PubChem (paralel CPU)	20
20	Grafik nilai <i>speed-up</i> pengolahan <i>fingerprint</i> PubChem (paralel CPU)	20
21	Grafik nilai ukuran blok dengan waktu pengerjaan algoritme perbandingan similaritas	21
22	Grafik nilai jumlah aliran terhadap waktu pengerjaan algoritme perbandingan similaritas	22
23	Grafik waktu proses perbandingan similaritas dengan <i>fingerprint</i> MACCS (paralel GPU)	23
24	Grafik nilai <i>speed-up</i> perbandingan similaritas dengan <i>fingerprint</i> MACCS (paralel GPU)	23
25	Grafik waktu proses perbandingan similaritas dengan <i>fingerprint</i> PubChem (paralel GPU)	24
26	Grafik nilai <i>speed-up</i> perbandingan similaritas dengan <i>fingerprint</i> PubChem (paralel GPU)	24

DAFTAR LAMPIRAN

	Hasil nilai <i>speed-up</i> pengolahan <i>fingerprint</i> MACCS	33
--	---	----

2	Hasil nilai <i>speed-up</i> pengolahan <i>fingerprint</i> PubChem	34
3	Hasil waktu perbandingan similaritas dengan berbagai ukuran blok	35
4	Hasil waktu perbandingan similaritas dengan berbagai jumlah aliran (<i>stream</i>)	36
5	Hasil nilai <i>speed-up</i> algoritme perbandingan similaritas untuk <i>fingerprint</i> MACCS	39
6	Hasil nilai <i>speed-up</i> algoritme perbandingan similaritas untuk <i>fingerprint</i> PubChem	40
7	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap remdesivir	41
8	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap favipiravir	42
9	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap lopinavir	43
10	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap hydroxychloroquine	44
11	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap chloroquine	45
12	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap nitazoxanide	46
13	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap ivermectin	47
14	Daftar senyawa yang memiliki nilai similaritas tinggi terhadap oseltamivir	48

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :

- a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
- b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.

2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

I PENDAHULUAN

1.1 Latar Belakang

Covid-19 merupakan suatu penyakit yang begitu fenomenal pada tahun 2020. Penyakit ini disebabkan oleh infeksi virus SARS COV-2. Karena sifat virus ini yang mudah menyebar dengan cepat serta dibantu dengan mudahnya akses manusia melakukan transportasi lintas negara, permasalahan lokal virus ini di China menjadi pandemi global yang memberi dampak pada semua negara di dunia. Selama tanggal 12 hingga 18 Juli, 32 dari 34 provinsi di Indonesia melaporkan peningkatan kasus di mana 17 provinsi mengalami peningkatan sebesar 50% atau lebih (WHO Indonesia 2021). Para peneliti di seluruh dunia turut berusaha dalam mengembangkan vaksin dan obat dari penyakit COVID-19 (*coronavirus disease*). Cara virus ini masuk ke dalam tubuh manusia adalah melalui reseptor ACE-2, yang dapat dilihat pada Gambar 1. Ligan yang dimaksud dalam gambar tersebut adalah sebuah molekul yang berikatan dengan molekul lain, pada kasus ini adalah protein *spike* dari virus corona dengan reseptor ACE-2.

Kebutuhan manusia terhadap obat dan sarana pencegahan membuat peneliti melakukan riset untuk menemukan alternatif obat dan vaksin virus corona. Salah satu alternatif yang dikembangkan adalah penggunaan tanaman herbal untuk mencegah dan mengobati penyakit ini (Bologa *et al.* 2019). Strategi *drug repurposing* sangat berguna mengingat proses penemuan obat secara konvensional membutuhkan waktu yang lama. *Drug repurposing* dilakukan dengan cara menemukan manfaat atau khasiat baru dari senyawa obat yang sudah terdaftar. *Drug repurposing* biasanya dilakukan dengan cara menganalisis interaksi dari senyawa obat dengan protein-protein yang berkaitan dari suatu penyakit, kemudian memprediksi interaksi obat-target baru yang belum diketahui sebelumnya (Ashburn dan Thor 2004).

Pendekatan obat herbal, atau biasa disebut jamu, dinilai sangat berguna untuk pencegahan dan pengobatan penyakit COVID-19 karena beberapa faktor, yakni tingkat ketersediaan obat, keamanan obat, dan tingkat kepercayaan masyarakat Indonesia. Sebagai obat tradisional yang berbahan dasar tumbuh-tumbuhan, tingkat ketersediaan obat herbal di Indonesia begitu melimpah, apalagi setelah dinyatakan sebagai salah satu negara dengan biodiversitas tanaman yang sangat luas dan besar. (Medail dan Quezel 1999). Keamanan obat dari jamu atau tanaman herbal sudah teruji dari waktu ke waktu karena metode ini telah digunakan secara turun-temurun di masyarakat tradisional. Tingkat kepercayaan masyarakat Indonesia terhadap obat herbal juga cukup tinggi, mengingat semakin banyak masyarakat yang menggunakan herbal sebagai alternatif pengobatan terhadap berbagai macam penyakit.

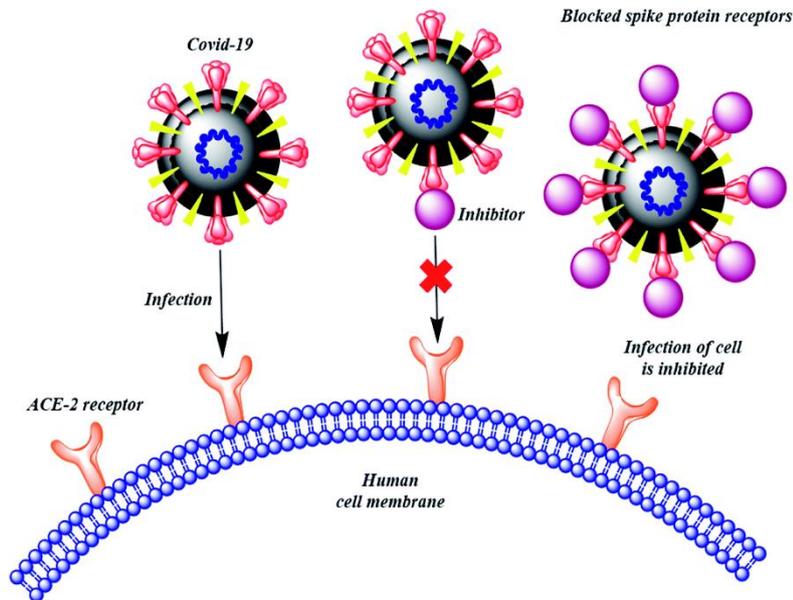
Besarnya data senyawa yang tersedia membuat proses pencarian similaritas secara sekuensial memakan waktu yang sangat lama, sehingga memerlukan pendekatan yang lebih efisien. Selain penggunaan perangkat keras yang memadai, kecepatan komputasi juga dipengaruhi oleh algoritme atau cara kerja komputer. Konsep paralelisasi dan penggunaan *Graphics Processing Unit* (GPU) untuk mempercepat penghitungan membuat waktu komputasi lebih cepat dibandingkan *Central Processing Unit* (CPU) pada jenis-jenis komputasi tertentu



(Athanasopoulos *et al.* 2017). Pencarian similaritas senyawa ini bisa menggunakan konsep paralelisasi yang disediakan GPU untuk mempercepat prosesnya.

Penelitian *drug-target interaction* (DTI) secara paralel telah dilakukan dengan beberapa skema, diantaranya adalah dengan *breadth-first search* (Reinaldo *et al.* 2020), *molecular docking* dengan GPU (Imbernon *et al.* 2021), serta BINDSURF yang merupakan metodologi penapisan virtual untuk mencari *binding site* protein terhadap suatu ligan (Sanchez-Linares *et al.* 2012). Penelitian ini bertujuan untuk mencari tahu potensi GPU dalam pencarian similaritas senyawa dengan memanfaatkan potensi paralelisasinya. Dengan mengetahui performa GPU pada pencarian senyawa obat ini, diharapkan dapat menjadi referensi penelitian *in-silico* berikutnya.

Penelitian ini bertujuan untuk menemukan potensi senyawa herbal yang ada di alam sebagai obat Covid-19, melalui pencarian similaritas dengan beberapa obat yang sudah ada. Senyawa-senyawa herbal yang diidentifikasi mirip dengan obat Covid-19 tersebut dapat dilakukan penelitian lanjutan secara *in-vitro* untuk verifikasi, dan apabila terbukti secara ilmiah akan dapat membantu masyarakat dalam menemukan alternatif obat untuk menghadapi pandemi Covid-19.



Gambar 1 Peran ligan dalam interaksi virus dan sel (Sethi *et al.* 2020)

1.2 Rumusan Masalah

Perumusan masalah pada penelitian ini adalah:

- Bagaimana menerapkan algoritme perbandingan similaritas pada data senyawa molekul dengan komputasi GPU untuk menentukan tingkat similaritasnya?
- Apa saja tanaman herbal yang memiliki potensi tinggi menjadi kandidat obat dari penyakit covid-19?

1.3 Tujuan

Tujuan dari penelitian ini adalah:

- a. Merancang, mengimplementasikan, dan mengevaluasi solusi komputasi paralel perbandingan similaritas menggunakan GPU.
- b. Menemukan kandidat senyawa yang berpotensi untuk mencegah atau mengobati Covid-19.

1.4 Manfaat

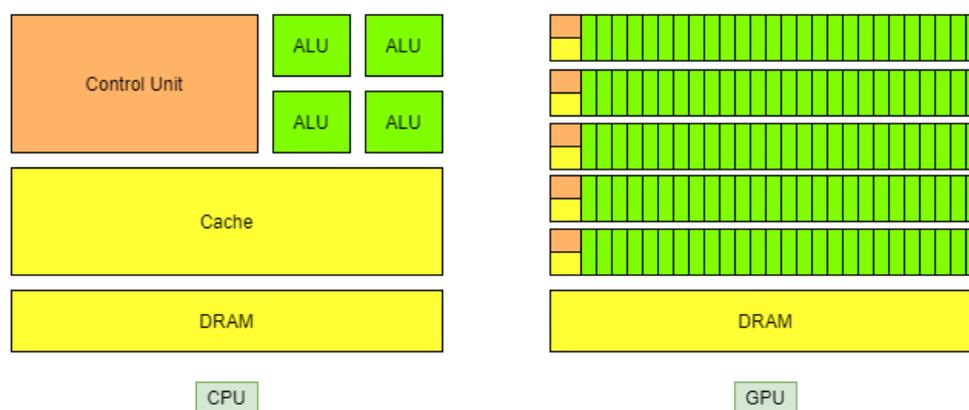
Hasil dari penelitian ini diharapkan dapat memberikan manfaat untuk meningkatkan kualitas dan kecepatan proses *drug discovery* dan *drug repurposing* dalam dunia jamu atau obat-obatan herbal sehingga pada akhirnya dapat membantu masyarakat dalam mengatasi penyakit yang relatif baru dengan harga terjangkau.

1.5 Ruang Lingkup Penelitian

Ruang lingkup pada penelitian ini adalah mencari data senyawa secara *in-silico* tanpa disertai uji *in-vitro* dan *in-vivo*. Data yang dicari terbatas pada informasi string SMILES (*simplified molecular input line entry system*) senyawa



GPU memiliki banyak *computational core* yang dapat digunakan untuk proses komputasi paralel. Hal yang perlu disiapkan pengguna adalah bagaimana memisah data-data tersebut agar GPU dapat mengerjakannya secara paralel, lalu penggabungan dari hasil-hasil penghitungan tersebut agar tidak terjadi galat dan hasilnya valid.



Gambar 2 Ilustrasi desain arsitektur CPU dan GPU

Pada Gambar 2, dapat dilihat perbedaan arsitektur pada *Central Processing Unit* (CPU) dan GPU. Pada gambar arsitektur CPU, terlihat bahwa cukup banyak tempat digunakan untuk *control unit* dan *cache*. Hal ini masuk akal karena CPU akan menerima banyak sekali data dan perintah yang cenderung unik untuk setiap data yang akan diolah, sehingga perlu diakomodasi dengan *cache* dan *control unit* yang memadai. Pada sisi GPU, terlihat bahwa alokasi tempat untuk *cache* dan *control unit* relatif kecil dan minimal, serta sebagian besar terdiri dari *arithmetic and logic unit* (ALU) yang relatif kecil namun banyak. Hal ini terjadi karena tujuan awal dari GPU adalah meningkatkan performa komputer dalam mengolah data grafis agar lebih cepat diproses sehingga dapat lebih cepat pula tampil ke layar. Dan hal tersebut dicapai dengan memperbanyak unit pengolah data ALU karena data yang diolah cukup besar dan instruksi untuk mengolah data tersebut berupa aritmatika sederhana, seperti tambah kurang kali bagi (Li *et al.* 2019).

Berangkat dari analogi pemrosesan grafis, di mana setiap pikselnya akan diproses secara paralel, GPU memiliki banyak *thread* yang digunakan sebagai tempat untuk memproses data. Suatu kumpulan *thread* disebut sebagai *block* yang memiliki *cache* atau memori yang dapat digunakan bersama. Kumpulan *block* di suatu tempat yang sama akan membentuk *grid* (NVIDIA 2021). Pada GPU, *processing element* yang digunakan untuk mengolah data adalah *thread*. Abstraksi-abstraksi ini akan memudahkan pemrograman paralel saat melakukan implementasi pada penelitian.

GPU mengenal istilah hierarki memori, di mana terdapat beberapa memori yang memiliki ukuran dan hak akses berbeda. Program dapat berjalan lebih efisien apabila mampu memanfaatkan memori-memori ini secara optimal. Pada satuan terkecil, dikenal istilah *local memory* yang hanya bisa diakses oleh satu *thread*. Setiap *thread* memiliki *local memory*-nya masing-masing. Beberapa *thread* yang terkumpul dalam suatu *block* dapat mengakses memori yang disebut *shared memory*. Setiap *block* memiliki *shared memory*-nya masing-masing. Dan yang paling luas dan dapat diakses oleh semua *thread* adalah *global memory*. Ukuran

shared memory jauh lebih kecil daripada global memory, dan waktu akses memori dari shared memory ke global memory memakan waktu yang cukup signifikan. Pembagian data yang efisien diperlukan akan shared memory melakukan sesedikit mungkin pengambilan data dari global memory.

2.4 Basis Data ZINC

ZINC adalah sebuah basis data molekul yang tersedia secara komersil. Pada dasarnya, data molekul yang terkandung dalam basis data ini berbentuk *simplified molecular input line entry system* (SMILES), namun dalam perkembangannya, tersedia juga representasi molekul dalam dua dimensi dan tiga dimensi (Irwin *et al.* 2012). Basis data ZINC sering digunakan dalam penelitian untuk menemukan ligan, yakni sebuah ion atau molekul yang dapat menempel pada atom metal dengan ikatan kovalen. Ligan ini sering digunakan untuk menemukan obat dari suatu penyakit atau virus dengan cara mengganggu siklus hidup virus tersebut, atau langsung menghancurkan virusnya secara struktural. Untuk mengetahui bentuk data dari basis data ZINC, silakan melihat Gambar 3.

zinc_id	smiles
ZINC000245189325	O=P(=O)O
ZINC000029747110	COc1cccc2c1[C@@H]1CN(CCCCN3C(O)nc4c(sc5ncc(-c6cccc6)nc54)c3=O)C[C@@H]1C2=O
ZINC000137550338	CN(C)c1cc(CNCC(C)(C)C)O)c2c1C[C@H]1C[C@H]3[C@H](N(C)C)C(=O)[C@H](C(N)=O)C(=O)[C@@]3(O)C(=O)[C@H]1C2=O
ZINC000137550409	CN(C)c1cc(CNCC(C)(C)C)O)c2c1C[C@H]1C[C@H]3[C@H](N(C)C)C(=O)[C@H](C(N)=O)C(=O)[C@@]3(O)C(=O)[C@H]1C2=O
ZINC000137550489	CN(C)c1cc(CNCC(C)(C)C)O)c2c1C[C@H]1C[C@H]3[C@H](N(C)C)C(=O)[C@H](C(N)=O)C(=O)[C@@]3(O)C(=O)[C@H]1C2=O
ZINC000005161047	Cc1ccc2c(n1)Oe1ccc([C@@H](C)C(=O)OCC(=O)N(C)C)cc1C2
ZINC000137550260	CN(C)c1cc(CNCC(C)(C)C)O)c2c1C[C@H]1C[C@H]3[C@H](N(C)C)C(=O)[C@H](C(N)=O)C(=O)[C@@]3(O)C(=O)[C@H]1C2=O

Gambar 3 Contoh file unduhan dari basis data ZINC

2.5 Algoritme Perbandingan Similaritas

Algoritme perbandingan similaritas adalah suatu algoritme untuk menentukan kemiripan antara dua sekuens data. Dalam bioinformatika, algoritme ini sudah ditemukan dan dipakai cukup lama, yakni sejak abad 20 (Corpet 1984). Penggunaan dan implementasi dari algoritme ini juga luas, dan bergantung pada tipe data yang akan dibandingkan, seperti data sekuens protein, data *fingerprind*, data biner, dan lain-lain. Pada penelitian ini, data yang dibandingkan adalah data biner berupa hasil keluaran *fingerprind* dari molekul senyawa.

Secara garis besar, perbandingan similaritas biner yang dilakukan pada penelitian ini bertujuan untuk memberikan nilai similaritas terhadap dua string biner yang berbeda (Bajusz *et al.* 2015). Misalnya string biner yang dibandingkan adalah 0110 dan 1010. Setiap indeks pada string tersebut menunjukkan ada atau tidaknya suatu ciri pada senyawa yang dirujuk. Algoritme Tanimoto akan mencari jumlah digit 1 pada kedua string yang berada di indeks yang sama (dalam contoh ini adalah indeks ketiga), dan membaginya dengan jumlah digit 1 di kedua string pada indeks yang berbeda (dalam contoh ini, ada nilai 1 pada indeks pertama, kedua, dan ketiga). Algoritme Tanimoto akan memberikan nilai 1/3 untuk kedua string biner di atas. Dengan kata lain, algoritme Tanimoto akan membagi jumlah ciri yang sama pada kedua senyawa dengan jumlah ciri keseluruhan yang ada pada kedua senyawa tersebut, baik hanya dimiliki oleh salah satu senyawa atau keduanya.

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.
Perpustakaan IPB University

Selain Tanimoto, ada beberapa algoritme lain untuk menentukan nilai similaritas biner, yakni Dice dan Kosinus. Algoritme Dice mengalikan jumlah ciri yang sama pada kedua senyawa dengan angka 2, lalu membaginya dengan jumlah ciri yang ada pada kedua senyawa. Algoritme kosinus akan membagi jumlah ciri yang sama pada kedua senyawa dengan nilai akar dari perkalian jumlah ciri yang hanya dimiliki oleh salah satu senyawa. Penelitian ini menggunakan algoritme Tanimoto karena cukup sederhana dan terbukti sesuai untuk digunakan sebagai dasar penghitungan similaritas senyawa kimia menurut Bajusz *et al.* (2015).

@Hak cipta milik IPB University

IPB University

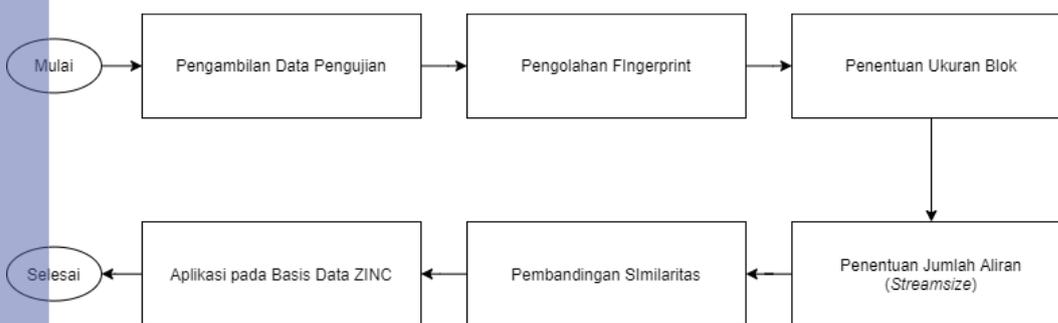


III METODE

3.1 Tahapan Penelitian

Gambar 4 menjelaskan tahapan dan metode penelitian yang digunakan. Sebagai data awal, dikumpulkan data senyawa ZINC pada sub-himpunan biogenic yang berjumlah 308.035 senyawa. Perhitungan *fingerprint* yang digunakan adalah MACCS dan PubChem, serta dilakukan dengan metode sekuensial CPU dan paralel CPU. Setelah data *fingerprint* terbentuk, dilakukan parameter tuning pada skema paralelisasi GPU yakni penentuan ukuran blok dan penentuan jumlah stream. Setelah ukuran blok dan jumlah stream yang sesuai ditentukan, dilakukanlah proses perbandingan similaritas yang diimplementasikan dengan dua metode, yakni sekuensial GPU dan paralel GPU. Setelah dihitung waktunya, evaluasi kedua metode tersebut dilakukan dengan penghitungan nilai speed-up.

Setelah ditemukan model terbaik, keseluruhan data ZINC yang berjumlah 997.402.117 senyawa serta telah dihitung *fingerprint* PubChemnya diterapkan pada model tersebut untuk menemukan senyawa-senyawa yang mirip dengan senyawa obat Covid-19.



Gambar 4 Diagram alir tahapan penelitian

3.2 Data Penelitian

Data yang digunakan dan dianalisis pada penelitian ini adalah seluruh data SMILES senyawa pada basis data ZINC. Karena memiliki data yang besar, peneliti mengambil sebagian data ZINC yang berada dalam subset biogenic yang berjumlah 308.035 senyawa sebagai dasar pengembangan algoritme. Data senyawa tersebut dicari tingkat kemiripannya terhadap delapan senyawa obat Covid-19 yang sudah disetujui atau dianggap sebagai kandidat kuat dalam dunia kedokteran. Untuk memudahkan pencarian, daftar senyawa obat yang dijadikan acuan dapat dilihat dalam Tabel 1.

Tabel 1 Daftar senyawa obat yang digunakan sebagai acuan

Nama Senyawa	Cara Kerja	Sumber
Remdesivir	Menghentikan replikasi RNA ^a coronavirus di dalam sel inang.	(Cao <i>et al.</i> 2020)

Favipiravir	Inhibitor RNA-dependen RNA-polimerase pada virus flu secara umum.	(Furuta, Y., Komeno, T., & Nakamuba 2017)
Lopinavir	Agen antiretrovirus, inhibitor protease.	(Cvetkovic dan Goa 2003)
Hydroxichloroquine	Menyebabkan alkalisasi dalam sel, mencegah asamisasi yang diperlukan virus dalam proses replikasi.	(Lei <i>et al.</i> 2020)
Chloroquine	Menyebabkan alkalisasi dalam sel, mencegah asamisasi yang diperlukan virus dalam proses replikasi.	(Lei <i>et al.</i> 2020)
Nitazoxanide	Menekan inflamasi saat badai sitokin.	(Mahmoud <i>et al.</i> 2020)
Ivermectin	Agen anti-parasit yang diakui FDA ^b , menunjukkan sifat anti-viral pada HIV ^c , VEEV ^d , dan Zika. Saat ini sedang diteliti untuk pengobatan Covid.	(Caly <i>et al.</i> 2020)
Oseltamivir	Inhibitor pada protein 3CLpro virus corona	(Tan <i>et al.</i> 2020)

^aRibonucleic Acid, pembawa informasi genetik pada virus.

^bFood and Drug Administration, lembaga pemerintah Amerika Serikat yang mengurus izin makanan dan obat.

^cHuman Immuno-deficiency Virus, penyebab penyakit Acquired Immuno-deficiency Syndrome (AIDS)

^dVenezuelan Equine Encephalitis Virus, virus penyebab ensepalitis pada kuda, keledai, dan zebra. Virus ini menyebar melalui nyamuk

3.3 Pengolahan *Fingerprint*

Seperti yang terlihat pada Gambar 5, data yang diperoleh dari basis data ZINC terdiri atas `zinc_id`, yang merupakan kode unik untuk mengindeks tiap senyawa, dan SMILES. SMILES adalah bentuk penulisan senyawa dalam bentuk string sehingga dapat direpresentasikan secara ringkas dan jelas. Namun representasi ini tidak dapat digunakan sebagai acuan bahwa suatu senyawa mirip atau tidak, karena SMILES hanyalah bentuk string sederhana dari suatu senyawa.

Representasi SMILES ini perlu dirubah ke dalam bentuk *fingerprint* yang merupakan sekumpulan representasi senyawa dalam bentuk biner yang digunakan untuk menentukan tingkat kemiripan suatu senyawa dengan senyawa yang lainnya. Penggunaan *fingerprint* sebagai data representasi suatu senyawa membuat hasil perbandingan similaritas lebih akurat karena kita menilai dari sifat-sifat yang dibawa struktur senyawa itu sendiri.

Penelitian ini menggunakan dua jenis SMILES *fingerprint* yang sudah umum digunakan, yakni *fingerprint* MACCS dan PubChem. MACCS mengkodekan suatu senyawa ke dalam 166 bit biner, sedangkan PubChem *fingerprint* akan mengkodekan senyawa ke dalam 881 bit biner. Setiap bit biner yang ada dalam

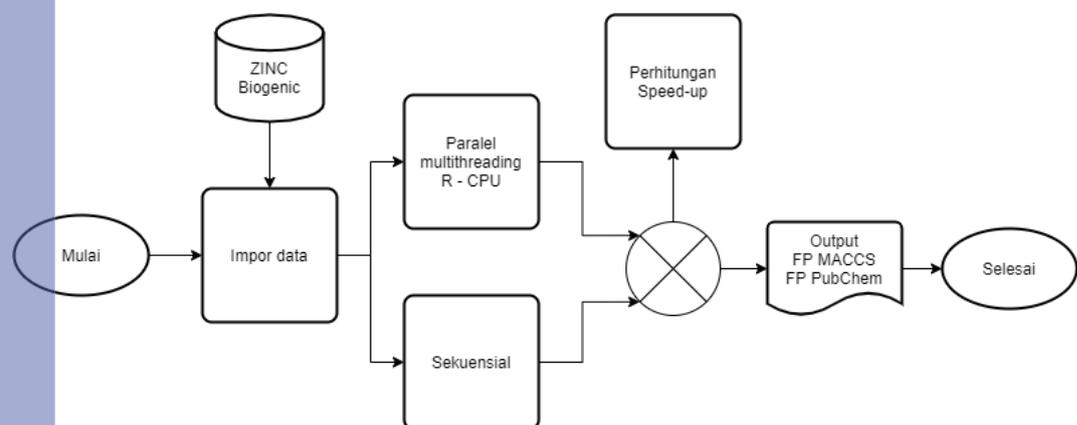
fingerprint MACCS dan PubChem merupakan penciri molekuler yang biasa digunakan kimiawan untuk melakukan pengelompokan senyawa. Beberapa contoh penciri yang digunakan adalah jumlah atom C dalam senyawa tersebut berjumlah lebih dari 4, serta keberadaan struktur cis/trans.

```
@Hak cipta milik IPB University
> head(df1)
  zinc_id smiles
1 ZINC000017653949 COC(=O)[C@H]1[C@H](OC)[C@@H](OC(=O)C2CC(OC)C(OC)C2)C[C@@H]2CN3CCC4C([NH]C5CC(OC)CC45)[C@H]3C[C@H]21
2 ZINC000031502642 COC(=O)C1=CO[C@@H](O)[C@@H]2O[C@H](CO)[C@@H](O)[C@H](O)[C@H]2O)[C@H]2[C@@H]1C=C[C@]2(O)CO
3 ZINC000006018601 C[C@@H]1CC[C@@]2(C)[C@H](C)/C(=C\COC(=O)/C=C/C3CC(O)C(O)C3)CC[C@]12C
4 ZINC000017654726 CO1CC2C(C(OC)C1OC)-C1C(CC3C(C1OC)OC3)C[C@H](C)[C@@](C)(O)[C@@H]2OC(=O)C1CCCC1
5 ZINC000001721658 CC1(C)[C@H]2CCC(=O)[C@@H]1C2
6 ZINC000032296250 CC1=CC(=O)[C@@]2(C(C)C)C(=O)C3=C([C@H](O)[C@H]4OC(=O)[C@@]35CCCC(C)C)[C@H]45)[C@@H]12
```

Gambar 5 Data yang diunduh dari basis data ZINC

Secara garis besar, diagram alir dari proses pengolahan *fingerprint* digambarkan seperti Gambar 6. Proses ini dimulai dengan mengimpor subset data ZINC ke perangkat lunak RStudio, lalu melakukan dua proses pengolahan *fingerprint*, yakni secara sekuensial CPU dan paralel CPU. Proses sekuensial akan mengolah *fingerprint* MACCS dan PubChem secara satu per satu. Proses paralel akan membagi tugas pengolahan *fingerprint* ini ke dalam semua core CPU yang tersedia, yakni 32 core. Proses paralelisasi dilakukan dengan CPU, karena lebih praktis dan sudah mendukung library yang dibutuhkan, yakni rcdk.

Setelah proses pengolahan fingerprint selesai, selanjutnya dihitung berapa nilai *speed-up* yang dihasilkan dari waktu pemrosesan secara sekuensial dan paralel. Luaran yang dihasilkan adalah *fingerprint* MACCS sepanjang 166 bit dan *fingerprint* PubChem sepanjang 881 fingerprint. Luaran ini akan digunakan pada proses selanjutnya, yakni penghitungan kemiripan atau similaritas menggunakan CUDA. Contoh fingerprint yang dihasilkan dapat dilihat pada Gambar 7.



Gambar 6 Diagram alir proses pengolahan *fingerprint*

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.


```
FOR setiap nilai iterasi dari 1 sampai 7
  FOR setiap ukuran blok yang bernilai 32 hingga 1024 dengan
    kelipatan 32
    DO kalkulasi waktu perbandingan similaritas pada data
      biogenik fingerprint PubChem dengan 7 obat covid
      fingerprint PubChem
    ENDFOR
  ENDFOR
CALCULATE nilai rata-rata setiap ukuran blok pada 7 iterasi
DETERMINE ukuran blok dengan waktu rata-rata tercepat
```

Gambar 10 Pseudocode untuk algoritme penentuan ukuran blok

3.5 Penentuan Jumlah Aliran (*Streamsize*)

Penentuan *streamsize* diperlukan untuk mencari tahu jumlah *stream* optimal yang dapat digunakan saat proses paralelisasi GPU. Gambar 11 menjelaskan proses penentuan jumlah *stream* yang optimal. Jumlah *stream* dicari dalam rentang satu hingga seratus dengan kelipatan satu.

```
FOR setiap nilai iterasi dari 1 sampai 7
  FOR setiap jumlah stream yang bernilai 1 hingga 100 dengan
    kelipatan 1
    DO kalkulasi waktu perbandingan similaritas pada data
      biogenik fingerprint PubChem dengan 7 obat covid
      fingerprint PubChem
    ENDFOR
  ENDFOR
CALCULATE nilai rata-rata setiap jumlah stream pada 7 iterasi
DETERMINE jumlah stream dengan waktu rata-rata tercepat
```

Gambar 11 Pseudocode untuk algoritme paralel CPU pengolahan *fingerprint*

3.6 Perbandingan Similaritas

3.6.1 Similaritas Tanimoto

Algoritme similaritas Tanimoto akan membandingkan setiap bit pada lokasi yang sama. Hal ini dilakukan mengingat bahwa *fingerprint* pada posisi yang sama menunjukkan penciri senyawa yang sama pula, sehingga kita bisa melihat mirip atau tidaknya suatu pasangan senyawa dari penciri tersebut.

Sebagai contoh, pada Gambar 12 terdapat dua pasang *fingerprint* senyawa yang dibandingkan. Similaritas Tanimoto dihitung dengan cara menghitung jumlah semua bit pada suatu posisi di mana kedua senyawa memiliki kesamaan, yakni bernilai satu, lalu dibagi dengan jumlah semua posisi bit di mana salah satu dari kedua senyawa tersebut bernilai satu. Pada contoh di atas, terdapat dua posisi di mana kedua senyawa memiliki kesamaan bernilai satu, yakni posisi keempat dan kelima. Serta terdapat enam posisi di mana salah satu dari kedua senyawa tersebut bernilai satu, yakni posisi satu, tiga, empat, lima, enam, dan tujuh. Sehingga nilai similaritas dari kedua senyawa tersebut adalah dua per enam, atau dengan kata lain satu per tiga.

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Fingerprint Senyawa Kandidat

0	0	1	1	1	0	1	0
---	---	---	---	---	---	---	---

Fingerprint Obat

1	0	0	1	1	1	0	0
---	---	---	---	---	---	---	---

Gambar 12 Ilustrasi penghitungan similaritas Tanimoto

3.6.2 Algoritme Sekuensial

Algoritme sekuensial mengambil satu data senyawa kandidat yang diwakilkan dengan satu baris data, lalu menghitung similaritasnya dengan beberapa senyawa obat yang sudah dipilih, dan menyimpan hasilnya dalam suatu tabel yang berukuran sesuai dengan jumlah senyawa kandidat dan senyawa obat.

Pada ilustrasi yang terdapat pada Gambar 13, dan juga pseudokode pada Gambar 14, kita ambil satu data *fingerprint* pada baris teratas, lalu dibandingkan dengan tiga *fingerprint* obat yang sudah ada. Hasil dari penghitungan similaritas Tanimoto dituliskan dalam tabel Hasil Similaritas baris teratas, sehingga setiap baris dari tabel hasil mewakili senyawa kandidat, dan setiap kolomnya mewakili senyawa obat yang dibandingkan. Tabel Hasil Similaritas pada baris pertama dan kolom ketiga menunjukkan nilai similaritas Tanimoto antara senyawa kandidat pertama dengan senyawa obat ketiga.

Sesuai dengan namanya, algoritme ini dilakukan secara sekuensial dalam suatu *thread* CPU, sehingga waktu pengerjaannya bergantung kepada kemampuan CPU untuk mengolah data. Implementasi ini ditulis dalam suatu program berbahasa Python dengan menggunakan pustaka *pandas* untuk melakukan impor data, ekspor data, dan manajemen dataframe, serta pustaka *NumPy* untuk melakukan operasi pada array dua dimensi.

Data Senyawa

1	0	1	...
0	1	0	...
0	0	1	...
0	0	0	...
1	0	0	...

Data Obat

0	0	1	...
1	0	0	...
0	1	1	...

Hasil Similaritas

0.3	0.24	0.1
0.4	0.4	0.9
0.34	0.55	0.97
0.8	0.95	0.1
0.05	0.4	0.6

Gambar 13 Representasi proses penjesajaran secara sekuensial

```
FOR setiap baris representasi fingerprint pada senyawa
kandidat obat
  FOR setiap baris representasi fingerprint pada senyawa
  obat Covid-19
    DO perbandingan similaritas
  ENDFOR
ENDFOR
```

Gambar 14 Pseudocode untuk algoritme sekuensial CPU perbandingan similaritas

3.6.3 Algoritme Paralel

Paralelisasi proses perbandingan similaritas dilakukan dengan memanfaatkan General Purpose GPU (GPGPU) Nvidia CUDA sebagai prosesor. Setiap *thread* yang ada dalam GPU dimanfaatkan untuk memproses satu senyawa kandidat dengan kedelapan senyawa obat. Setelah dihitung nilai similaritasnya, GPU ini menuliskan nilai tersebut ke tabel hasil, layaknya proses sekuensial di atas. Istilah ini dikenal dengan *single instruction, multiple data* (SIMD).

Pada Gambar 15 juga ditunjukkan bahwa selain membagi setiap senyawa ke dalam satu *thread*, pada proses ini juga turut dilakukan *multistreaming*, yakni membagi proses paralel ke dalam beberapa aliran (*stream*) yang berbeda. Hal ini dicapai dengan cara membagi data senyawa sama rata ke setiap aliran, serta menyalin data obat yang sama ke setiap aliran yang akan dipakai, sehingga setiap aliran tidak perlu melakukan komunikasi dengan aliran lain dan bisa berjalan secara optimal.

Mekanisme ini dapat mempercepat proses paralelisasi karena setiap aliran memiliki clock-nya sendiri, sehingga memungkinkan proses-proses tersebut berjalan secara tidak sinkron (*asynchronous*). Tentu saja proses ini hanya terjadi sampai penulisan data ke dalam tabel similaritas. Di akhir proses, setiap aliran mengumpulkan tabel hasilnya dan menuliskan kembali dalam satu tabel hasil similaritas secara berurutan, mulai dari aliran pertama sampai terakhir. Penulisan hasil ini dilakukan dengan sinkronisasi, sehingga tidak terjadi tumpang tindih, dan urutan tabel data yang dihasilkan sesuai dengan urutan senyawa yang dipakai. Algoritme ini dijabarkan secara singkat pada pseudokode yang ada di Gambar 16.

Pada abstraksi paralelisasi GPU, selain dikenal istilah *thread* yang melakukan kerja nyata, ada juga istilah blok yang merupakan sekumpulan thread yang berada dalam wadah yang sama. Pemrogram perlu menentukan ukuran blok yang sesuai sehingga dapat mencapai peningkatan kinerja secara maksimal. Maka dari itu, pada penelitian ini digunakan berbagai skema paralelisasi dengan ukuran blok yang berbeda, lalu diulang sebanyak tujuh kali, sehingga dapat menemukan ukuran blok yang dapat memproses data latih secara optimal. Nilai ukuran blok yang digunakan berada pada rentang 32, 64, 96, dan seterusnya hingga 1024. Hal ini dikarenakan jumlah maksimal *thread* yang berada dalam satu blok pada arsitektur CUDA hanya sampai 1024 *thread* per blok, dan ukuran warp yang ada pada arsitektur CUDA bernilai 32, sehingga untuk mencapai efisiensi optimal diperlukan ukuran blok yang memiliki kelipatan 32.

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Seperti yang disebutkan di atas, mekanisme *multistreaming* dilakukan agar setiap proses paralel berjalan secara tidak sinkron, sehingga dapat berjalan lebih cepat. Penelitian ini juga melakukan penghitungan similaritas dengan jumlah aliran yang berbeda, lalu diulang sebanyak tujuh kali, sehingga peneliti dapat menentukan jumlah aliran terbaik untuk digunakan dalam paralelisasi keseluruhan basis data ZINC. Nilai jumlah aliran yang digunakan berada pada rentang satu hingga seratus aliran.

Proses perbandingan similaritas secara paralel ini ditulis dalam bahasa pemrograman Python sebagai dasar, dengan bantuan pustaka PyCUDA yang digunakan sebagai kode pemrograman yang terjadi pada GPU. Pada setiap *thread* GPU, kode program ditulis menggunakan bahasa C++, sesuai dengan kebutuhan pustaka CUDA agar dapat mengeksekusi perintahnya.

	Data Senyawa	Data Obat	Hasil Similaritas																										
Stream 1	<table border="1"> <tr><td>1</td><td>0</td><td>1</td><td>...</td></tr> <tr><td>0</td><td>1</td><td>0</td><td>...</td></tr> </table>	1	0	1	...	0	1	0	...	<table border="1"> <tr><td>0</td><td>0</td><td>1</td><td>...</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>...</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>...</td></tr> </table>	0	0	1	...	1	0	0	...	0	1	1	...	<table border="1"> <tr><td>0.3</td><td>0.24</td><td>0.1</td></tr> <tr><td>0.4</td><td>0.4</td><td>0.9</td></tr> </table>	0.3	0.24	0.1	0.4	0.4	0.9
1	0	1	...																										
0	1	0	...																										
0	0	1	...																										
1	0	0	...																										
0	1	1	...																										
0.3	0.24	0.1																											
0.4	0.4	0.9																											
Stream 2	<table border="1"> <tr><td>0</td><td>0</td><td>1</td><td>...</td></tr> <tr><td>0</td><td>0</td><td>0</td><td>...</td></tr> </table>	0	0	1	...	0	0	0	...	<table border="1"> <tr><td>0</td><td>0</td><td>1</td><td>...</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>...</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>...</td></tr> </table>	0	0	1	...	1	0	0	...	0	1	1	...	<table border="1"> <tr><td>0.34</td><td>0.55</td><td>0.97</td></tr> <tr><td>0.8</td><td>0.95</td><td>0.1</td></tr> </table>	0.34	0.55	0.97	0.8	0.95	0.1
0	0	1	...																										
0	0	0	...																										
0	0	1	...																										
1	0	0	...																										
0	1	1	...																										
0.34	0.55	0.97																											
0.8	0.95	0.1																											
Stream 3	<table border="1"> <tr><td>1</td><td>0</td><td>0</td><td>...</td></tr> <tr><td>1</td><td>1</td><td>1</td><td>...</td></tr> </table>	1	0	0	...	1	1	1	...	<table border="1"> <tr><td>0</td><td>0</td><td>1</td><td>...</td></tr> <tr><td>1</td><td>0</td><td>0</td><td>...</td></tr> <tr><td>0</td><td>1</td><td>1</td><td>...</td></tr> </table>	0	0	1	...	1	0	0	...	0	1	1	...	<table border="1"> <tr><td>0.05</td><td>0.4</td><td>0.6</td></tr> <tr><td>0.03</td><td>0.41</td><td>0.23</td></tr> </table>	0.05	0.4	0.6	0.03	0.41	0.23
1	0	0	...																										
1	1	1	...																										
0	0	1	...																										
1	0	0	...																										
0	1	1	...																										
0.05	0.4	0.6																											
0.03	0.41	0.23																											

Gambar 15 Representasi proses perbandingan secara paralel

```

DO pembuatan stream sesuai jumlah stream yang optimal
DO duplikasi data fingerprint obat covid-19
DO pembagian data senyawa kandidat obat ke setiap stream sesuai
dengan jumlah stream yang ada
FOR setiap stream yang ada di GPU
    FOR setiap thread GPU yang menghitung nilai similaritas
    baris representasi fingerprint pada senyawa obat Covid-19
        FOR setiap baris representasi fingerprint pada obat
        covid-19
            DO perbandingan similaritas
        ENDFOR
    ENDFOR
DO konkatenasi hasil similaritas sesuai urutan
    
```

Gambar 16 Pseudocode untuk algoritme paralel CPU perbandingan similaritas

3.7 Evaluasi Algoritme

Algoritme paralel dan algoritme sekuensial dibandingkan dengan cara menghitung nilai *speed-up*. Nilai *speed-up* adalah perbandingan dari waktu yang dibutuhkan untuk menjalankan algoritme sekuensial dengan waktu yang dibutuhkan untuk menjalankan algoritme paralel. Sebagai contoh, apabila algoritme sekuensial membutuhkan waktu 9 detik dan algoritme paralel memakan waktu 5 detik, maka nilai *speed-up* yang didapat adalah 1,8.

$$speed - up = \frac{t_{sekuensial}}{t_{paralel}} \quad (1)$$

Persamaan (1) merupakan rumus umum penghitungan nilai *speed-up*, di mana $t_{sekuensial}$ adalah waktu yang diperlukan untuk melakukan sesuatu secara sekuensial, dan $t_{paralel}$ adalah waktu yang diperlukan untuk melakukan sesuatu secara paralel. Semakin besar perbedaan waktu sekuensial dan paralel, maka nilai *speed-up* akan semakin tinggi, dengan catatan waktu sekuensial lebih lama daripada waktu paralel.

3.8 Aplikasi pada Basis Data ZINC

Setelah mendapat model paralelisasi yang baik dari data latih ZINC subset biogenic, model tersebut diaplikasikan pada seluruh dataset yang ada pada basis data ZINC. Proses paralelisasi ini memakan waktu cukup lama sehingga dengan diterapkannya paralelisasi, waktu pengolahan data dapat dipercepat. Hasil dari pengolahan seluruh basis data ini menunjukkan informasi tentang beberapa senyawa potensial yang mirip dengan senyawa obat Covid-19. Penelitian ini menggunakan *fingerprint* PubChem untuk mengolah seluruh basis data ZINC karena memiliki deskriptor lebih banyak, sehingga dapat mengidentifikasi kemiripan senyawa dengan lebih akurat. Sebagai perbandingan, data latih yang digunakan sebesar 308.035 senyawa, sedangkan keseluruhan basis data ZINC berjumlah 997.402.117 senyawa.

3.9 Lingkungan Pengembangan

Spesifikasi perangkat lunak dan perangkat keras yang digunakan dalam penelitian ini adalah sebagai berikut:

- Tipe perangkat : Desktop PC
- Sistem operasi : Xubuntu 18.04
- Prosesor : Intel Xeon Silver 4110 2,2 GHz

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Memori	: 64 GB
GPU	: NVIDIA RTX 2080
Penyimpanan	: SSD 512 GB, HDD 7 TB
Bahasa pemrograman	: C++, Python, dan R
Dukungan library	: PyCUDA dan rcdk
Perangkat lunak	: RStudio dan Visual Studio Code

@Hak cipta milik IPB University

IPB University



IV HASIL DAN PEMBAHASAN

4.1 Pengolahan *Fingerprint*

Gambar 17 dan 19 menunjukkan perbedaan waktu pemrosesan dari algoritme sekuensial dan paralel, baik pada *fingerprint* MACCS maupun PubChem. Grafik yang ditunjukkan pada Gambar 18 dan Gambar 20 menunjukkan nilai *speed-up* yang didapat untuk proses konversi dari representasi SMILES ke *fingerprint* MACCS dan PubChem. Secara keseluruhan, nilai *speed-up* yang didapat cukup besar dan signifikan. Hal ini menunjukkan bahwa penggunaan pemrosesan paralel dengan *multithread* CPU untuk mengolah *fingerprint* dapat memberi dampak yang signifikan terhadap waktu pengerjaan.

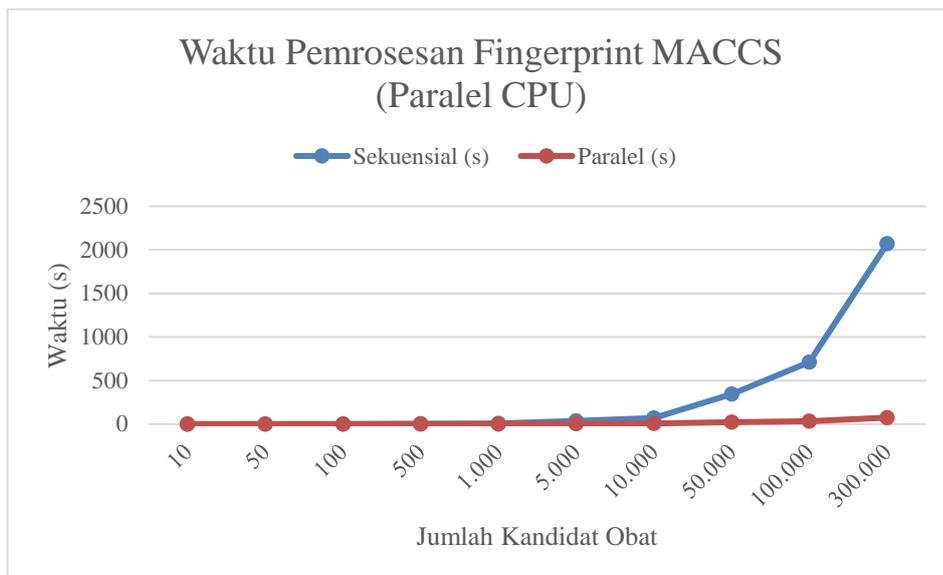
Pada Gambar 18, terjadi peningkatan yang sangat signifikan dari nilai *speed-up* terhadap jumlah kandidat obat. Pada jumlah kandidat senyawa obat yang berjumlah seribu ke bawah, terlihat peningkatan yang tidak terlalu signifikan. Kenaikan nilai *speed-up* mulai besar dan terlihat pada jumlah data sebanyak lima ribu ke atas. Dengan melihat grafik ini, terlihat bahwa potensi skalabilitas dari pengolahan *fingerprint* MACCS masih sangat tinggi, atau dengan kata lain, nilai *speed-up* masih dapat meningkat lagi seiring bertambahnya jumlah senyawa kandidat yang diproses. Grafik tersebut juga menunjukkan bahwa diperlukannya penelitian lebih lanjut mengenai batasan potensi pengolahan paralel terhadap pengolahan *fingerprint* MACCS untuk mendapat nilai *speed-up* optimal. Sejauh ini, nilai *speed-up* tertinggi sebesar 27,73 didapat dari jumlah kandidat obat sebanyak 300.000. Dengan kata lain, jika kita melakukan proses paralel pengolahan *fingerprint* MACCS pada data senyawa berjumlah 300.000, maka akan menjadi 27 kali lebih cepat dibandingkan jika kita mengolahnya secara sekuensial. Dengan memperhitungkan jumlah *thread* prosesor yang berjumlah 32, maka efisiensi dari proses paralelisasi MACCS dengan 300.000 kandidat obat adalah sebesar 86,66%.

Pada Gambar 20 terlihat bahwa nilai *speed-up* yang terjadi pada proses pengolahan *fingerprint* PubChem sebesar 15,27 cenderung lebih kecil apabila dibandingkan dengan *fingerprint* MACCS yang dapat mencapai nilai 27. Dengan memperhitungkan jumlah *thread*, maka diperoleh juga nilai efisiensi sebesar 47,72%. Terlihat peningkatan yang signifikan dari jumlah kandidat sepuluh, lima puluh, seratus, dan lima ratus. Untuk jumlah kandidat seribu ke atas, nilai *speed-up* terlihat semakin landai, hingga pada nilai 50.000, 100.000, dan 300.000 terlihat mengalami naik dan turun, sehingga cukup memungkinkan bahwa sudah mencapai batas (*threshold*) pada area ini. Perbedaan *speed-up* antara pengolahan *fingerprint* MACCS dan PubChem kemungkinan disebabkan oleh jumlah bit yang ada pada MACCS dan PubChem, yakni MACCS sebesar 166 bit dan PubChem sebesar 881 bit. Tentu saja hal ini akan mempengaruhi pekerjaan setiap *thread* yang memproses *fingerprint* tersebut, sehingga pekerjaan yang dilakukan suatu *thread* untuk mengolah *fingerprint* PubChem akan lebih besar daripada *fingerprint* MACCS. Hal ini juga turut menjelaskan nilai *speed-up* yang cenderung tidak mengalami peningkatan pada pengolahan *fingerprint* MACCS untuk jumlah kandidat senyawa sepuluh, lima puluh, seratus, lima ratus, dan seribu, karena waktu yang digunakan untuk mengolah *fingerprint* lebih singkat daripada waktu yang digunakan untuk membagi data mentah dan mengumpulkan data hasil pengolahan dari setiap *thread* yang bekerja.

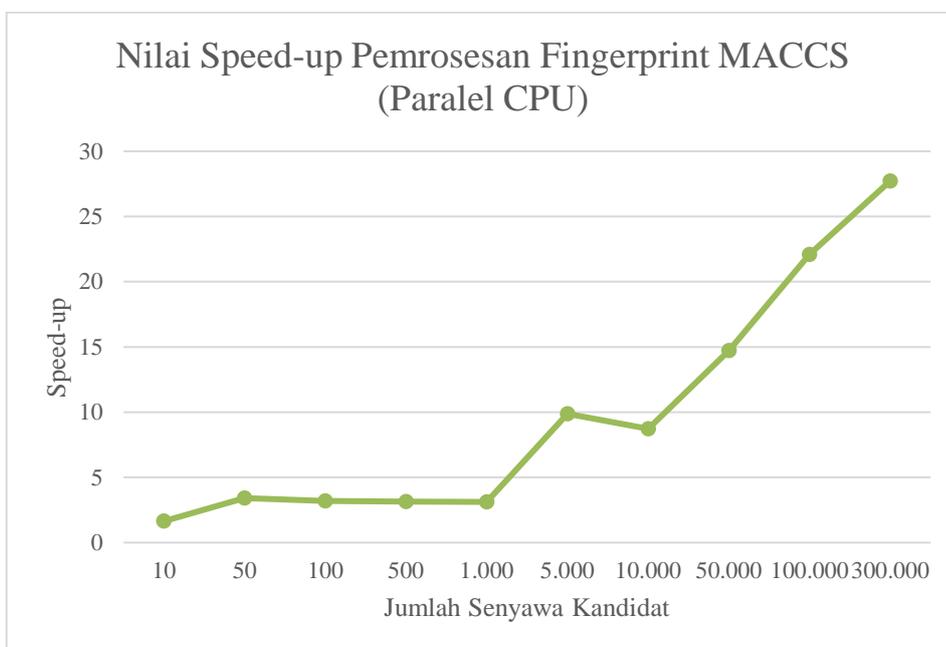
@auipzsmwvka - IPB University

IPB University

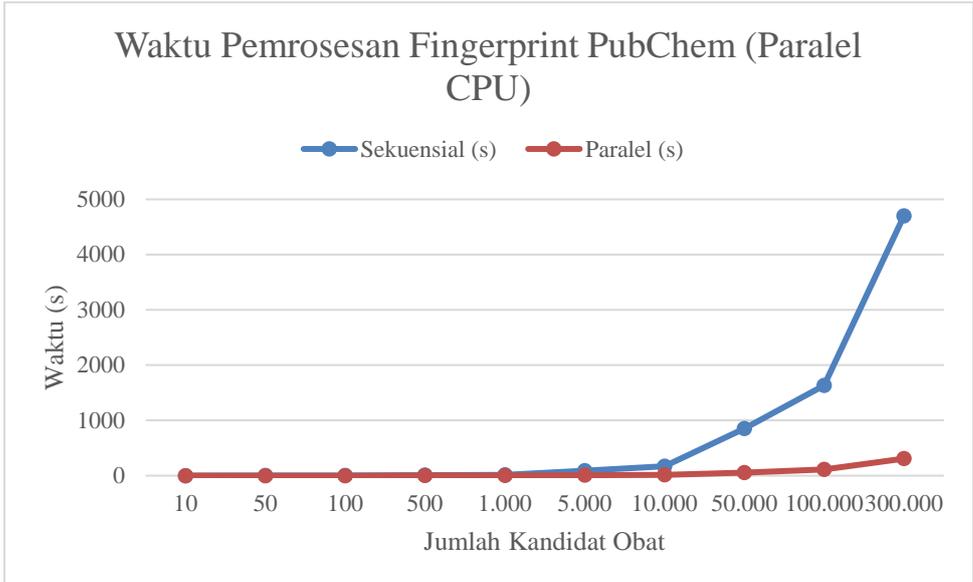
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



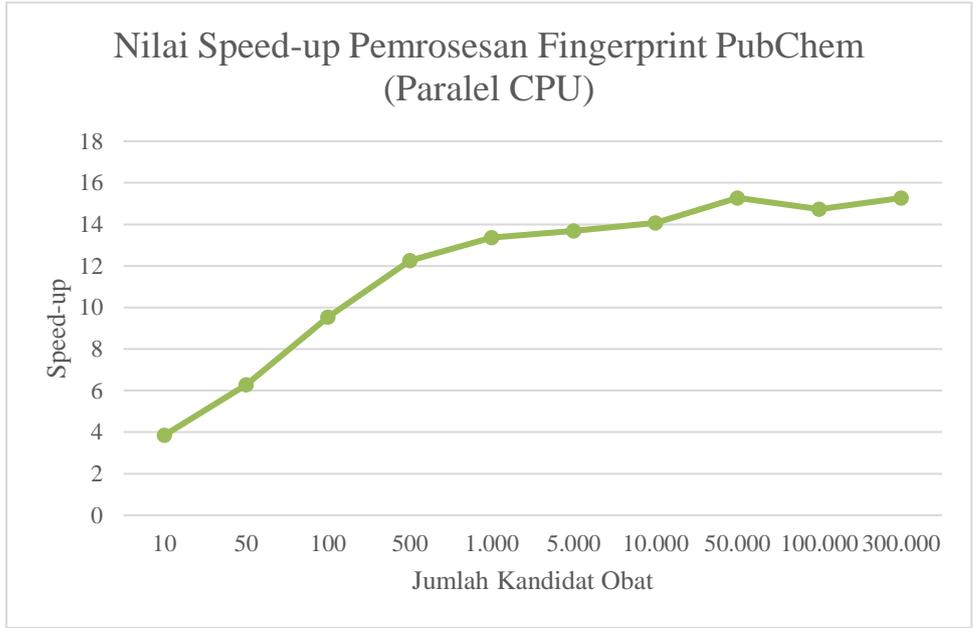
Gambar 17 Grafik waktu pengolahan *fingerprint* MACCS (paralel CPU)



Gambar 18 Grafik nilai *speed-up* pengolahan *fingerprint* MACCS (paralel CPU)



Gambar 19 Grafik waktu pengolahan *fingerprint* PubChem (paralel CPU)

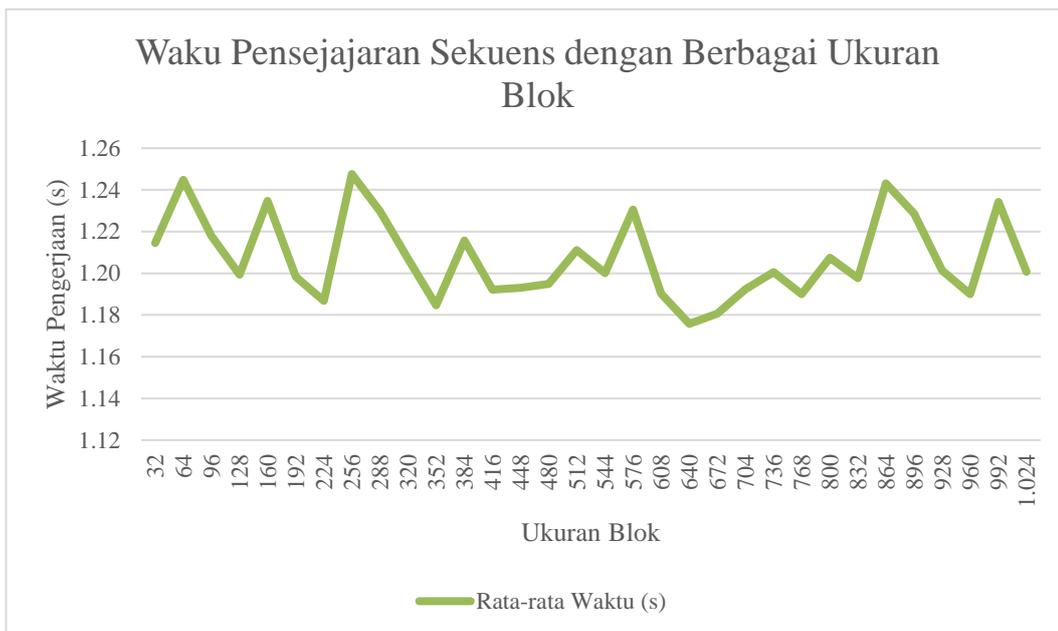


Gambar 20 Grafik nilai *speed-up* pengolahan *fingerprint* PubChem (paralel CPU)

4.2 Penentuan Ukuran Blok

Gambar 21 menunjukkan bahwa waktu pengerjaan untuk berbagai ukuran blok terlihat dinamis dan lebih memiliki nilai minimal global di sekitar angka 600 *thread* per blok. Garis kuning tua menunjukkan rata-rata waktu pengerjaan dari tujuh ulangan. Setelah mencari nilai rata-rata yang paling sedikit, didapatkanlah ukuran blok sebesar 640 *thread* per blok dengan waktu rata-rata 1,176 detik.

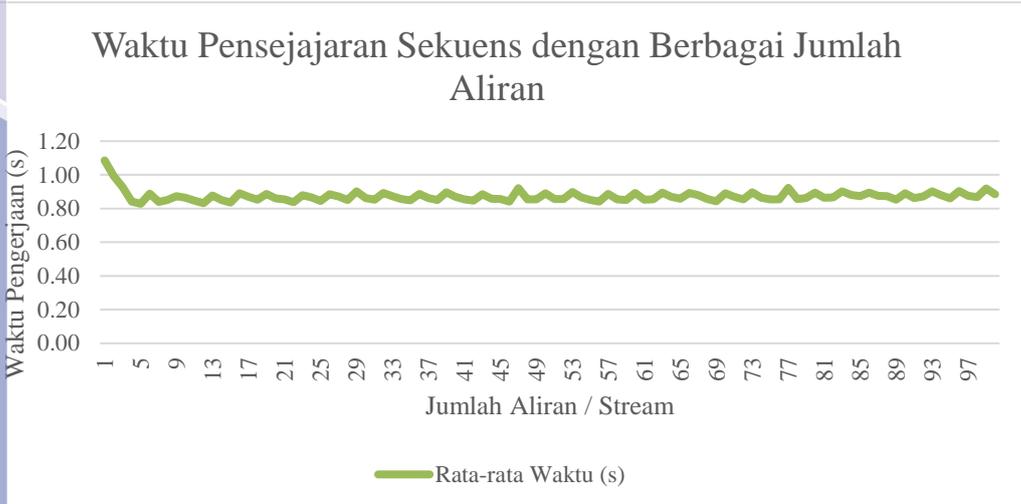
Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Gambar 21 Grafik nilai ukuran blok dengan waktu pengerjaan algoritme perbandingan similaritas

4.3 Penentuan Jumlah Aliran (*Streamsize*)

Nilai rata-rata yang ditunjukkan dengan garis berwarna kuning tua pada Gambar 22 terlihat menunjukkan osilasi atau perulangan seperti gelombang. Bisa dilihat pada jumlah aliran satu sampai lima, waktu pengerjaan cenderung menurun, lalu dilanjutkan pada jumlah aliran di atas lima yang secara perlahan menunjukkan peningkatan. Nilai terendah menunjuk pada jumlah aliran sebesar lima dengan waktu pengerjaan 0,828 detik. Hal ini menunjukkan, penggunaan jumlah aliran kurang dari lima termasuk ke dalam keadaan tidak optimal, karena jumlah aliran belum sebanding dengan banyaknya data yang diproses. Pada nilai lebih dari lima, yang menunjukkan peningkatan waktu pengerjaan lama, mengindikasikan bahwa jumlah aliran yang terlalu banyak terhadap ukuran data yang diproses, sehingga komputer melakukan pekerjaan yang tidak perlu yakni membuat dan mengalokasikan pembagian data ke dalam aliran tambahan tersebut.

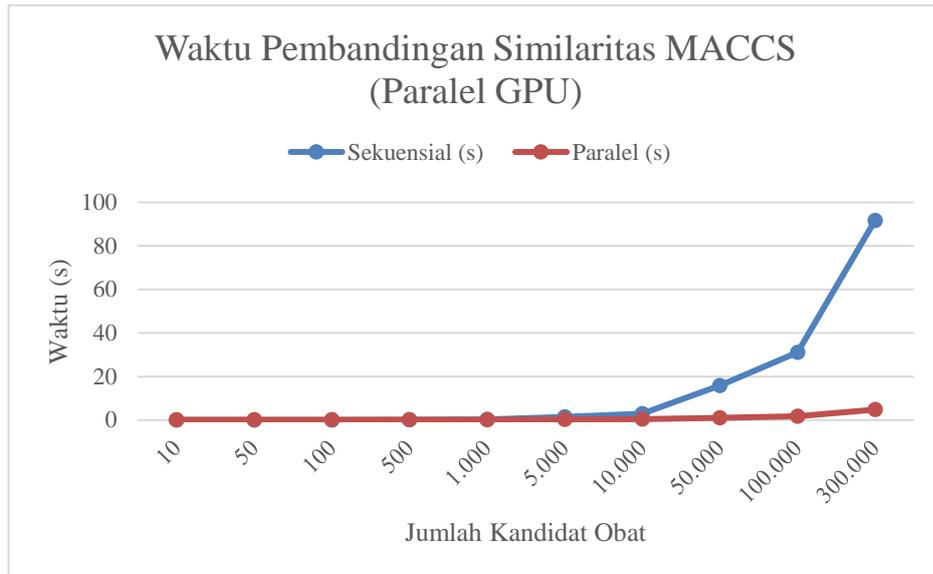


Gambar 22 Grafik nilai jumlah aliran terhadap waktu pengerjaan algoritme perbandingan similaritas

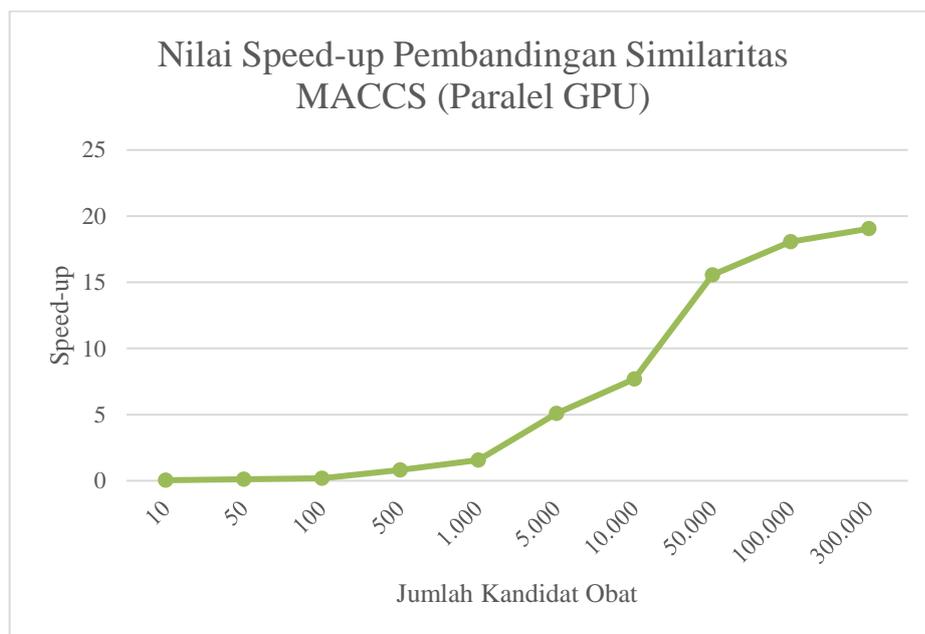
4.4 Perbandingan Similaritas

Gambar 23 dan Gambar 25 menunjukkan perbedaan waktu proses perbandingan similaritas yang dilakukan secara sekuensial dan paralel, baik pada fingerprint MACCS maupun PubChem. Grafik yang diperlihatkan oleh Gambar 24 dan Gambar 26 menunjukkan bahwa perbandingan algoritme sekuensial dan paralel pada algoritme perbandingan similaritas, baik untuk *fingerprint* MACCS ataupun PubChem, menghasilkan nilai *speed-up* yang cukup signifikan. Untuk *fingerprint* MACCS, nilai *speed-up* tertinggi berada pada nilai 19,05 yang ada pada jumlah kandidat obat 300.000, sedangkan untuk *fingerprint* PubChem, nilai *speed-up* tertinggi berada pada nilai 55,51 yang ada pada jumlah kandidat obat 100.000. Perbedaan nilai *speed-up* ini menunjukkan bahwa tugas-tugas yang dapat diparalelasi pada *fingerprint* PubChem cenderung lebih banyak, mengingat jumlah bitnya yang mencapai 881.

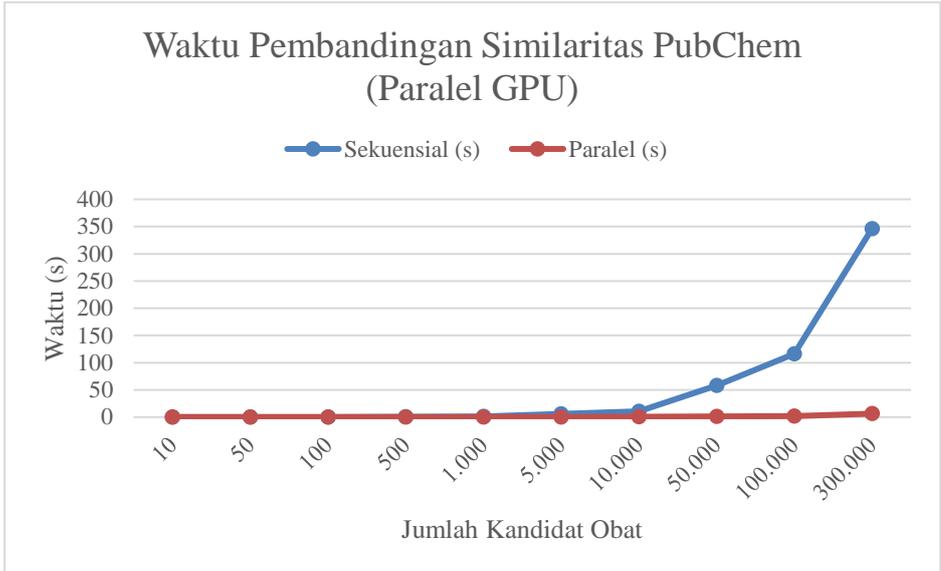
Hak Cipta Dilindungi Undang-undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



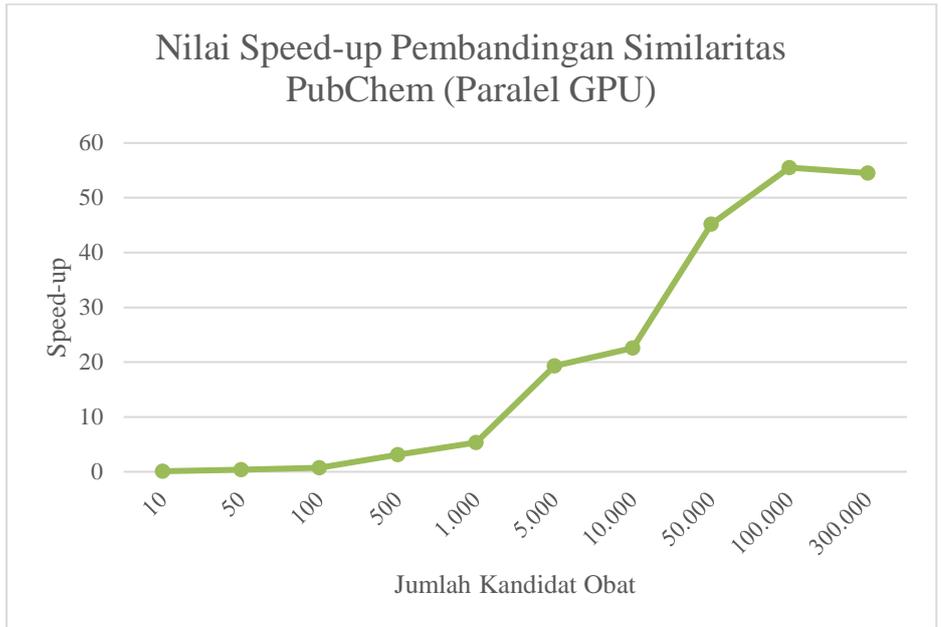
Gambar 23 Grafik waktu proses pembeding similaritas dengan *fingerprint* MACCS (paralel GPU)



Gambar 24 Grafik nilai *speed-up* pembeding similaritas dengan *fingerprint* MACCS (paralel GPU)



Gambar 25 Grafik waktu proses pembandingan similaritas dengan *fingerprint* PubChem (paralel GPU)



Gambar 26 Grafik nilai *speed-up* pembandingan similaritas dengan *fingerprint* PubChem (paralel GPU)

4.5 Aplikasi pada Basis Data ZINC

Tabel 2 Hasil teratas pembandingan similaritas pada seluruh basis data zinc yang dapat diidentifikasi

Nama Obat	Nilai Similaritas	ID Kandidat Obat	Nama lain	Sumber Herbal
Remdesivir	0.7799	ZINC29134440	Brivanib	-
	0.7567	ZINC13684256	Alaninate	-

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

	0,7391	ZINC53147179	Puromycin	Bakteri <i>Streptomyces alboniger</i> (de Koning-Ward <i>et al.</i> 2001)
Favipiravir	0,7500	ZINC5116994	5-hydroxypyrazinamide	-
	0,7478	ZINC1081066	2-carboxypyrazine	-
	0,6850	ZINC500059	Pyrazine Carboxylic Acid Hydrazide	-
Lopinavir	0,8214	ZINC49889244	Gliotide	Jamur <i>Gliocladium sp.</i> (Lang <i>et al.</i> 2006)
	0,8214	ZINC95607016	Paecilodepsi peptide C	Jamur <i>Paecilomyces cinnamomeus</i> (Isaka <i>et al.</i> 2007)
	0,8083	ZINC230078061	Adouetine Y	Kulit batang <i>Discaria Americana</i> (Giacomelli <i>et al.</i> 2001)
Hydroxi-chloroquine	1,0000	ZINC1530654	Hydroxy-chloroquine	Kulit batang <i>Peruvia cinchona</i> (Tolkushin <i>et al.</i> 2020)
	0,9934	ZINC1843038	Cletoquine	-
	0,9334	ZINC19144231	Chloroquine	-
Chloroquine	1,0000	ZINC19144231	Chloroquine	Kulit batang <i>Peruvia cinchona</i> (Tolkushin <i>et al.</i> 2020)
	0,9929	ZINC1873617	Desethyl-chloroquine	-
	0,9858	ZINC6036375	Bidesethyl-chloroquine	-
Nitazoxanide	0,8971	ZINC5924265	Tizoxanide	-
	0,8839	ZINC29124339	Tizoxanide Glucuronide	-

	0,7273	ZINC2257	Zolamine	-
	1,0000	ZINC169300893	Avermectin B1 Mono- saccharide	Fermentasi <i>Streptomyces avermitilis</i> (Kitani <i>et al.</i> 2009)
Ivermectin	0,9604	ZINC169292347	Nemadectin	Bakteri <i>Streptomyces cyaneogriseus</i> (Li <i>et al.</i> 2019)
	0,9510	ZINC169339191	Milbemycin Alpha25	Bakteri <i>Streptomyces hygroscopicus</i> subsp. <i>aureo- lacrimosus</i> (Nonaka <i>et al.</i> 2000)
	0,7593	ZINC13370140	Antillatoxin	Sianobakter <i>Lyngbya majuscula</i> (Goto <i>et al.</i> 2011)
Osetamivir	0,7265	ZINC169367255	Aspochala- sin J	Jamur <i>Aspergillus flavipes</i> (Zhou <i>et al.</i> 2004)
	0,7248	ZINC169290233	Cespitulac- tam J	Koral <i>Cespitularia taeniata</i> (Shen <i>et al.</i> 2007)

Setelah mengaplikasikan algoritme perbandingan similaritas secara paralel pada seluruh basis data ZINC, ditemukan beberapa senyawa yang berpotensi menjadi kandidat obat. Sebagian besar data senyawa yang memiliki nilai similaritas tinggi masuk ke dalam kategori *unnamed compound*, sehingga tidak bisa dilakukan identifikasi lanjutan. Secara ringkas, hasil dari pencarian similaritas dapat dilihat pada Tabel 2.

Mulai dari atas, kandidat senyawa obat yang mirip remdesivir adalah puromycin, namun hal ini tidak begitu layak untuk dilakukan penelitian lanjutan karena nilai similaritas yang rendah yakni sebesar 0,7391. Dua senyawa lainnya tidak bisa ditemukan sumber alamnya, dan produksinya hanya dapat dilakukan dalam laboratorium.

Obat yang kedua yakni favipiravir tidak memiliki senyawa yang berasal dari alam pada hasil pencarian teratasnya. Dapat dilihat juga bahwa yang paling mirip dengan favipiravir adalah 5-hydroxypyrazinamide yang memiliki tingkat similaritas 0,75, cukup rendah untuk dapat dikatakan mirip.

Obat yang ketiga yakni lopinavir memiliki beberapa kandidat senyawa yang cukup mirip, yakni gliotide, paecilopdepsipeptice C, dan adouetine Y. Ketiga senyawa ini secara berurutan dapat ditemukan di jamur *Gliocladium sp.*, jamur *Paecilomyces cinnamomeus*, dan kulit batang *Discaria americana*. Perlu dicatat juga bahwa nilai similaritas dari ketiga kandidat senyawa ini tidak terlalu tinggi, yakni berkisar pada rentang 0,80 – 0,82.

Untuk kandidat obat yang keempat dan kelima, yakni hydroxichloroquine dan chloroquine, keduanya dapat ditemukan dalam kulit batang tumbuhan *Peruvia cinchona*. Hal yang perlu dicatat adalah tumbuhan ini bukan berasal dari Indonesia, jadi akan sulit untuk dapat menemukannya.

Obat keenam yakni nitazoxanide tidak memiliki senyawa yang berasal dari alam pada hasil pencarian teratasnya. Senyawa seperti tizoxanide, tizoxanide glucuronide, dan zolamine hanya dapat diproduksi dalam laboratorium.

Obat ketujuh, yakni ivermectin, mendapat perhatian khusus karena selain terdapat senyawa yang mirip di alam, tingkat similaritasnya juga tinggi, yakni di atas 0,95. Yang pertama adalah senyawa ivermectin sendiri yang dapat diperoleh dari hasil fermentasi bakteri *Streptomyces avermitilis*. Senyawa kandidat selanjutnya adalah nemadectin dengan nilai similaritas 0,9604 yang dapat diproduksi oleh bakteri *Streptomyces cyaneogriseus*. Senyawa kandidat yang terakhir, yakni milbemycin alpha25 dengan tingkat kemiripan 0,9510, dapat ditemukan pada bakteri *Streptomyces hygroscopicus subsp. lacrimosus*. Namun perlu dicatat bahwa penetapan ivermectin sebagai obat coronavirus sampai saat ini masih dalam tahap pengembangan di laboratorium dan dibuktikan pada beberapa penelitian *in vitro*.

Obat yang kedelapan, yakni oseltamivir, memiliki kesamaan pada sumber bahan alam pada rentang similarita 0,70 hingga 0,75. Similaritas 75% terdapat pada senyawa antillatoxin, yakni senyawa toksin yang didapat dari sianobakter laut *Lyngbya majuscula*. Similaritas 72% terdapat pada senyawa aspochalasin J dan *Cespitulactam B*, yang secara berurutan terdapat pada jamur *Aspergillus flavipes* dan koral *Cespitularia taeniata*.

Terakhir, karena batasan dari penelitian ini adalah mencari bahan herbal, yakni bahan alam yang berasal dari tanaman, maka dari pencarian ini didapatkan *Gliocladium sp*, *Paecylomyces cinnamomeus*, *Discaria americana*, serta *Peruvia cinchona*.

V SIMPULAN DAN SARAN

5.1 Simpulan

Algoritme perbandingan similaritas dapat diterapkan untuk mencari nilai similaritas pada data senyawa molekuler berupa SMILES dengan cara mencari deskriptornya berupa *fingerprint*, kemudian membandingkan kedua *fingerprint* senyawa dengan similaritas tanimoto. Penggunaan komputasi GPU dapat meningkatkan kinerja proses ini hingga 55 kali lebih cepat.

Terdapat beberapa kandidat tanaman herbal, dan bahkan senyawa obat coronavirus itu sendiri dapat ditemukan di alam dan menjadi kandidat senyawa obat Covid-19. Beberapa bahan herbalnya sulit ditemukan di Indonesia karena memang tempat asalnya dari luar negeri.

5.2 Saran

Meskipun penelitian ini berfokus pada obat herbal, namun tidak menutup kemungkinan bahwa senyawa non-herbal yang didapat dari penapisan basis data ZINC ini dapat menjadi alternatif yang lebih praktis untuk menjadi obat Covid-19. Peneliti berharap bahwa hasil dari penelitian ini dapat menjadi saran, masukan, dan motivasi untuk melakukan analisis lebih lanjut di dalam laboratorium untuk mengidentifikasi bahan herbal yang dapat digunakan untuk mengobati virus corona.

Selain itu, peneliti juga berharap bahwa hasil penelitian kandungan senyawa herbal yang ada dalam tanaman dan organisme asal Indonesia dapat dikumpulkan menjadi satu basis data agar dapat dilakukan analisis *in silico* secara komprehensif. Saran ini muncul dari kesulitan peneliti dalam mengidentifikasi lebih lanjut apakah suatu senyawa terdapat pada tanaman herbal tertentu di Indonesia.

DAFTAR PUSTAKA

- Ashburn TT, Thor KB. 2004. Drug repositioning: Identifying and developing new uses for existing drugs. *Nat Rev Drug Discov.* 3(8):673–683. doi:10.1038/nrd1468.
- Athanasopoulos A, Dimou A, Mezaris V, Kompatsiaris I, Wen Z, Shi J, Li Q, He B, Chen J, Nakahara H, *et al.* 2017. Performance of deep learning computation with TensorFlow software library in GPU-capable multi-core computing platforms. *Int Work Image Anal Multimed Interact Serv.* 19 i:240–242. doi:10.1109/FCCM.2012.47.
- Bajusz D, Racz A, Heberger K. 2015. Why is Tanimoto index an appropriate choice for fingerprint-based similarity calculations?. *J Cheminformatics.* 7(1):1-13. doi:10.1186/s13321-015-0069-3.
- Basnet P, Skalko-Basnet N. 2011. Curcumin: An anti-inflammatory molecule from a curry spice on the path to cancer treatment. *Molecules.* 16(6):4567–4598. doi:10.3390/molecules16064567.
- Bologa CG, Ursu O, Oprea TI. 2019. How to prepare a compound collection prior to virtual screening. *Methods Mol Biol.* 1939:119–138. doi:10.1007/978-1-4939-9089-4_7.
- Caly L, Druce JD, Catton MG, Jans DA, Wagstaff KM. 2020. The FDA-approved drug ivermectin inhibits the replication of SARS-CoV-2 in vitro. *Antiviral Res.* 178 April:3–6. doi:10.1016/j.antiviral.2020.104787.
- Cao YC, Deng QX, Dai SX. 2020. Remdesivir for severe acute respiratory syndrome coronavirus 2 causing COVID-19: An evaluation of the evidence. *Travel Med Infect Dis.* 35 April:101647. doi:10.1016/j.tmaid.2020.101647.
- Corpet F. 1984. Multiple sequence alignment with hierarchical clustering. *Nucleic Acid Res.* 16(22):10881–10890. doi:10.1093/nar/16.22.10881.
- Cvetkovic RS, Goa KL. 2003. Lopinavir/Ritonavir. *Drugs.* 63(8):769–802. doi:10.2165/11204950-000000000-00000.
- de Koning-Ward TF, Waters AP, Brendan S. 2001. Puromycin-N-acetyltransferase as a selectable marker for use in *Plasmodium falciparum*. *Mol Biochem Parasitol.* 117(2):155-160. doi:10.1016/S0166-6851(01)00344-9.
- Furuta Y, Komeno T, Nakamura T. 2017. Favipiravir (T-705), a broad spectrum inhibitor of viral RNA polymerase. *Proc Jpn Acad Ser B Phys Biol Sci.* 93(7):449–463. doi: 10.2183/pjab.93.027
- Giacomelli SR, Missau FC, Mostardeiro MA, da Silva UF, Dalcol II, Zanatta N, Morel A. 2001. Cyclopeptides from the bark of *Disaria a americana*. *J Nat Prod.* 64(7):997-999. doi:10.1021/np000552v.
- Goto R, Okura K, Sakazaki H, Sugawara T, Matsuoka S, Inoue M. 2011. Synthesis and biological evaluation of triazole analogues of antillatoxin. *Tetrahedron.* 67(35):6659-6672. doi:10.1016/j.tet.2011.05.012.
- Imbernon B, Serrano A, Bueno-Crespo A, Abellan JL, Perez-Sanchez H, Cecilia JM. 2021. METADOCK 2: a high-throughput parallel metaheuristic scheme for molecular docking. *Bioinformatics.* 37(11):1515-1520. doi:10.1093/bioinformatics/btz958.
- Irwin JJ, Sterling T, Mysinger MM, Bolstad ES, Coleman RG. 2012. ZINC: a free tool to discover chemistry for biology. *Chem Inf Model.* 52(7):1757–1768. doi:

10.1021/ci3001277.

Isaka M, Palasarn S, Lapanun S, Sriklung K. 2007. Paecilodepsipeptide A, an antimalarial and antitumor cyclohexadepsipeptide from the insect pathogenic fungus *Paecilomyces cinnamomeus* BCC 9616. *J Nat Prod.* 70(4):675-678. doi:10.1021/np060602h.

Kamboj VP. 2000. Herbal medicine. *Current Science.* 78(1):35-39.

Kitani S, Ikeda H, Sakamoto T, Noguchi S, Nihira T. 2009. Characterization of a regulatory gene, *aveR*, for the biosynthesis of avermectin in *Streptomyces avermitilis*. *Appl Microbiol Biotechnol.* 82(6):1089-1096. doi:10.1007/s00253-008-1850-2.

Lang G, Mitova MI, Ellis G, van der Sar S, Phipps RK, Blunt JW, Cummings NJ, Cole AL, Munro MH. 2006. Bioactivity profiling using HPLC/microtiter-plate analysis: application to a New Zealand marine alga-derived fungus, *GlIOCladium sp.* *J Nat Prod.* 69(4):621-624. doi:10.1021/np0504917.

Lei ZN, Wu ZX, Dong S, Yang D-H, Zhang L, Ke Z, Zou C, Chen Z-S. 2020. Chloroquine and hydroxychloroquine in the treatment of malaria and repurposing in treating COVID-19. *Pharmacology Ther.* 216(1):107672. doi:10.1016/j.pharmthera.2020.107672.

Li C, He H, Wang J, Liu H, Wang H, Zhu Y, Wang X, Zhang Y, Xiang W. 2019. Characterization of a LAL-type regulator *NemR* in nemadectin biosynthesis and its application for increasing nemadectin production in *Streptomyces cyaneogriseus*. *Sci China Life Sci.* 62(3):394-405. doi:10.1007/s11427-018-9442-9

Li F, Ye Y, Tian Z, Zhang X. 2019. CPU versus GPU: which can perform matrix computation faster—performance comparison for basic linear algebra subprograms. *Neural Comput Appl.* 31(8):4353–4365. doi:10.1007/s00521-018-3354-z.

Luebke D, Harris M, Govindaraju N, Lefohn A, Houston M, Owens J, Segal M, Papakipos M, Buck I. 2006. GPGPU: general-purpose computation on graphics hardware. *2006 ACM/IEEE Conference on Supercomputing.* 208-es. doi:10.1145/1188455.1188672.

Mahmoud DB, Shitu Z, Mostafa A. 2020. Drug repurposing of nitazoxanide: can it be an effective therapy for COVID-19? *J Genet Eng Biotechnol.* 18(35):1–10. doi:10.1186/s43141-020-00055-5.

McGovern SL, Caselli E, Grigorieff N, Shoichet BK. 2002. A common mechanism underlying promiscuous inhibitors from virtual and high-throughput screening. *J Med Chem.* 45(8):1712–1722. doi:10.1021/jm010533y.

Medail F, Quezel P. 1999. Biodiversity hotspots in the Mediterranean Basin: Setting global conservation priorities. *Conserv Biol.* 13(6):1510–1513. doi:10.1046/j.1523-1739.1999.98467.x.

Nonaka K, Tsukiyama T, Okamoio Y, Sato K, Kumasaka C, Yamamoto T, Maruyama F, Yoshikawa H. 2000. New milbemycins from *Streptomyces hygroscopicus* subsp. *aureolacrimosus*: fermentation, isolation and structure elucidation. *J Antibiot.* 53(7):694-704. doi:10.7164/antibiotics.53.694.

NVIDIA Corporation. 2021. *CUDA Toolkit Documentation - v11.4.0.* Santa Clara: NVIDIA Corporation.

Oprea TI, Bologna CG, Edwards BS, Prossnitz ER, Sklar LA. 2005. Post-high-throughput screening analysis: An empirical compound prioritization scheme.

Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

- J Biomol Screen.* 10(5):419–426. doi:10.1177/1087057104272660.
- Pal SK, Shukla Y. 2003. Herbal medicine: current status and the future. *Asian Pacific J Cancer Prev.* 4(4):281–288.
- Reinaldo A, Kusuma WA, Rahmawan H, Herdiyeni Y. 2020. Implementation of breadth-first search parallel to predict drug-target interaction in plant-disease graph. *International Conference on Computer Science and Its Application in Agriculture (ICOSICA) 2020.* 1-5. doi:10.1109/ICOSICA49951.2020.9243216.
- Rishton GM. 1997. Reactive compounds and in vitro false positives in HTS. *Drug Discov Today.* 2(9):382–384. doi:10.1016/S1359-6446(97)01083-0.
- Rishton GM. 2003. Nonleadlikeness and leadlikeness in biochemical screening. *Drug Discov Today.* 8(2):86–96. doi:10.1016/S1359644602025722.
- Sanchez-Linares I, Perez-Sanchez H, Cecilia JM, Garcia JM. 2012. High-throughput parallel blink virtual screening using BINDSURF. *BMC Bioinformatics.* 13. doi:10.1186/1471-2105-13-S14-S13.
- Sethi A, Sanam S, Munagalasetty S, Jayanthi S, Alvala M. 2020. Understanding the role of galectin inhibitors as potential candidates for SARS-CoV-2 spike protein: In silico studies. *RSC Adv.* 10(50):29873–29884. doi:10.1039/d0ra04795c.
- Shen YC, Cheng YB, Kobayashi J, Kubota T, Takahashi Y, Mikami Y, Ito J, Lin YS. 2007. Nitrogen-containing verticillene diterpenoids from the taiwanese soft coral *Cespitularia taeniata*. *J Nat Prod.* 70(12):1961–1965. doi:10.1021/np078011u.
- Tan A, Duan L, Ma Y, Huang Q, Mao K, Xiao W, Xia H, Zhang S, Zhou E, Ma P, *et al.* 2020. Is oseltamivir suitable for fighting against COVID-19: In silico assessment, in vitro and retrospective study. *Bioorg Chem.* 104:104257. doi:10.1016/j.bioorg.2020.104257.
- Tolkushin AG, Luchinin EA, Kholovnya-Voloskova ME, Zavyalov AA. 2020. History of aminoquinoline preparations: from cinchona bark to chloroquine and hydroxychloroquinon. *Problemy Sotsial'noi Gigieny, Zdravookhraneniia i Istorii Meditsiny.* 28(Special Issue):1118–1122. doi:10.32687/0869-866x-2020-28-s2-1118-1122.
- [WHO Indonesia] World Health Organization Indonesia. 2021. Weekly Epidemiological Update on COVID-19. Jakarta (ID). [diakses 23 Juli 2021].
- Zhou GX, Wijeratne EK, Bigelow D, Pierson LS, VanEtten HD, Gunatilaka AL. 2004. Aspochalasin I, J, and K: three new cytotoxic cytochalasins of *Aspergillus flavipes* from the rhizosphere of *Ericameria laricifolia* of the Sonoran Desert. *J Nat Prod.* 67(3):328–332. doi:10.1021/np030353m.



@Hak cipta milik IPB University

LAMPIRAN

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 1 Hasil nilai *speed-up* pengolahan *fingerprint* MACCS

Jumlah Data	Waktu (detik)		<i>Speed-up</i>
	Sekuensial	Paralel	
10	0,110	0,067	1,642
50	0,402	0,118	3,407
100	0,730	0,229	3,188
500	3,980	1,270	3,134
1.000	6,745	2,164	3,117
5.000	35,982	3,643	9,877
10.000	68,594	7,861	8,726
50.000	346,762	23,550	14,723
100.000	711,037	32,199	22,083
300.000	2.070,507	74,681	23,725

@Hak cipta milik IPB University

IPB University



Lampiran 2 Hasil nilai *speed-up* pengolahan *fingerprint* PubChem

Jumlah Data	Waktu (detik)		<i>Speed-up</i>
	Sekuensial	Paralel	
10	0,262	0,068	3,853
50	0,911	0,145	6,283
100	2,097	0,220	9,532
500	8,436	0,688	12,262
1.000	16,500	1,235	13,360
5.000	87,179	6,367	13,692
10.000	169,011	12,014	14,068
50.000	849,909	55,638	15,276
100.000	1.634,857	110,966	14,733
300.000	4.701,407	307,838	15,272

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 3 Hasil waktu perbandingan similaritas dengan berbagai ukuran blok

Ukuran Blok	Rata-rata Waktu (detik)
32	1.2144
64	1.2449
96	1.2180
128	1.1993
160	1.2348
192	1.1982
224	1.1867
256	1.2476
288	1.2296
320	1.2068
352	1.1846
384	1.2158
416	1.1922
448	1.1931
480	1.1949
512	1.2111
544	1.2001
576	1.2305
608	1.1903
640	1.1758
672	1.1806
704	1.1923
736	1.2006
768	1.1900
800	1.2075
832	1.1976
864	1.2431
896	1.2286
928	1.2014
960	1.1899
992	1.2343
1.024	1.2008



Lampiran 4 Hasil waktu perbandingan similaritas dengan berbagai jumlah aliran (*stream*)

Jumlah Aliran	Rata-rata Waktu (detik)
1	1.0857
2	0.9925
3	0.9267
4	0.8419
5	0.8283
6	0.8887
7	0.8391
8	0.8501
9	0.8729
10	0.8646
11	0.8466
12	0.8327
13	0.8778
14	0.8506
15	0.8361
16	0.8914
17	0.8703
18	0.8526
19	0.8876
20	0.8599
21	0.8538
22	0.8383
23	0.8796
24	0.8655
25	0.8442
26	0.8842
27	0.8707
28	0.8516
29	0.9023
30	0.8619
31	0.8523
32	0.8933
33	0.8730
34	0.8569
35	0.8495
36	0.8871
37	0.8620
38	0.8518
39	0.8969
40	0.8701
41	0.8552
42	0.8476
43	0.8856
44	0.8583

45	0.8565
46	0.8421
47	0.9209
48	0.8547
49	0.8554
50	0.8915
51	0.8559
52	0.8562
53	0.8991
54	0.8653
55	0.8501
56	0.8409
57	0.8865
58	0.8551
59	0.8514
60	0.8934
61	0.8533
62	0.8543
63	0.8939
64	0.8691
65	0.8594
66	0.8919
67	0.8802
68	0.8556
69	0.8437
70	0.8897
71	0.8695
72	0.8546
73	0.8963
74	0.8634
75	0.8549
76	0.8544
77	0.9230
78	0.8566
79	0.8629
80	0.8953
81	0.8645
82	0.8653
83	0.9017
84	0.8819
85	0.8737
86	0.8952
87	0.8755
88	0.8735
89	0.8525
90	0.8911
91	0.8623
92	0.8713

@Hak cipta milik IPB University

IPB University



Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

93	0.9020
94	0.8789
95	0.8603
96	0.9039
97	0.8757
98	0.8688
99	0.9189
100	0.8857

@Hak cipta milik IPB University

IPB University



IPB University

Bogor Indonesia

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Perpustakaan IPB University

Lampiran 5 Hasil nilai *speed-up* algoritme perbandingan similaritas untuk *fingerprint* MACCS

Jumlah Data	Waktu (detik)		<i>Speed-up</i>
	Sekuensial	Paralel	
10	0,008	0,192	0,041
50	0,022	0,182	0,121
100	0,037	0,193	0,190
500	0,166	0,207	0,803
1.000	0,307	0,196	1,562
5.000	1,511	0,297	5,081
10.000	3,010	0,392	7,687
50.000	15,922	1,023	15,570
100.000	31,126	1,723	18,062
300.000	91,684	4,813	19,048

@Hak cipta milik IPB University

IPB University



Lampiran 6 Hasil nilai *speed-up* algoritme pembandingan similaritas untuk *fingerprint* PubChem

Jumlah Data	Waktu (detik)		<i>Speed-up</i>
	Sekuensial	Paralel	
10	0,016	0,167	0,094
50	0,077	0,192	0,402
100	0,160	0,216	0,742
500	0,577	0,186	3,101
1.000	1,118	0,209	5,349
5.000	5,891	0,305	19,291
10.000	10,826	0,480	22,551
50.000	57,994	1,283	45,209
100.000	116,079	2,091	55,505
300.000	346,162	6,349	54,523

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 7 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap remdesivir

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000029134440	0.77992278	Brivanib Alaninate
ZINC000013684256	0.75675678	Brivanib
ZINC000044351074	0.73966944	Diastereomeric Mix
ZINC000053147179	0.73913043	Puromycin
ZINC000049771447	0.73828125	Pad
ZINC000053684280	0.73828125	Cpad
ZINC000053683811	0.73170733	Tyrosyladenylate
ZINC000016051386	0.72972971	Tryptophanyl-5'amp
ZINC000003927238	0.72908365	Bamirastine
ZINC000029571783	0.72384935	Deaminofomycin
ZINC000003924220	0.72314047	Soraprazan
ZINC000053684179	0.72156864	Mant-Adp
ZINC000049775188	0.72118961	Nikkomycin X
ZINC000049946796	0.71984434	Benzamideadeninedinucleoside

@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 8 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap favipiravir

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000005116994	0.7500000000	5-Hydroxypyrazinamide
ZINC000001081066	0.7478260994	2-Carboxypyrazine
ZINC000000500059	0.7264957428	Pyrazine Carboxylic Acid Hydrazide
ZINC000001481960	0.7226890922	Acipimox
ZINC000021673450	0.6821705699	Morinamide
ZINC000000002005	0.6725663543	Pza
ZINC000014448311	0.6465517282	2-Methoxy-3,5-Dimethylpyrimidine
ZINC000000405269	0.6465517282	Methoxymethylpyrazine
ZINC000012660678	0.6375839114	Xanthopterin
ZINC000001840969	0.6347826123	2-Methoxy-5-Methylpyrazine
ZINC000026673320	0.6347826123	2-Methoxy-6-Methylpyrazine
ZINC000000164508	0.6240000129	2-Acetyl-3-Methylpyrazine
ZINC000014491514	0.6209150553	1h,3h
ZINC000002575506	0.6198347211	2-Ethoxy-3-Methylpyrazine

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang
 1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 9 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap lopinavir

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000096903141	0.827380955	Scutianine D
ZINC000028637921	0.821428597	Paecilodepsipeptide A
ZINC000049889244	0.821428597	Gliotide
ZINC000095607016	0.821428597	Paecilodepsipeptide C
ZINC000095607009	0.815476179	Paecilodepsipeptide B
ZINC000100822468	0.808383226	Vignatic Acid A
ZINC000230078061	0.808383226	Adouetine Y
ZINC000100822463	0.808383226	Vignatic Acid A
ZINC000085908661	0.808383226	Adouetine Y
ZINC000095561271	0.808139562	Pullularin C
ZINC000095563846	0.808139562	Pullularin A
ZINC000044387723	0.803571403	Sanjoinine G1
ZINC000096903019	0.802395224	Scutianine B
ZINC000004214064	0.801242232	Oxiramide

@Hak cipta milik IPB University

IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.



Lampiran 10 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap hydroxychloroquine

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000001530652	1.000000000	Hydroxychloroquine
ZINC000001843038	0.993377506	Cletoquine
ZINC0000019144231	0.933774828	Chloroquine
ZINC000001873617	0.927152335	Desethylchloroquine
ZINC000006036375	0.920529782	Bidesethyl Chloroquine
ZINC000001666887	0.915584445	Chloroquine- Pyrolidinyl
ZINC000005751278	0.909677445	Chloroquine Mustard
ZINC000045357215	0.891719758	Phenylequine
ZINC000003882239	0.853503167	Aminoquinol
ZINC000040380706	0.847561001	Dehydroxy Tebaquine
ZINC000103292547	0.832335352	Dehydroxyfluoro Tebaquine
ZINC000028470572	0.830303013	Metaquine
ZINC000027309019	0.827380955	Flourotebuquine
ZINC000033505730	0.823170721	Piclopastine

@Hak cipta milik IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 11 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap chloroquine

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000019144226	1.0000000000	Chloroquine
ZINC000001873617	0.9929078221	Desethylchloroquine
ZINC000006036375	0.9858155847	Bidesethyl Chloroquine
ZINC000001666887	0.9791666865	Chloroquine- Pyrolidinyl
ZINC000005751278	0.9724137783	Chloroquine Mustard
ZINC000045357215	0.9523809552	Phenylequine
ZINC000001530654	0.9337748289	Hydroxychloroquine
ZINC000001843038	0.9271523356	Cletoquine
ZINC000003882239	0.9115646482	Aminoquinol
ZINC000040380706	0.9025974274	Dehydroxy Tebaquine
ZINC000103292547	0.8853503466	Dehydroxyfluoro Tebaquine
ZINC000028470572	0.8838709593	Metaquine
ZINC000027309019	0.8797468543	Flourotebuquine
ZINC000002018210	0.8750000000	Iomethin

@Hak cipta milik IPB University

IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 12 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap nitazoxanide

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000005924265	0.897058844	Tizoxanide
ZINC000029124339	0.883928597	Tizoxanide Glucuronide
ZINC000000002257	0.727272748	Zolamine
ZINC000004217523	0.693779885	Tenonitrozole
ZINC000013889772	0.691588759	Neticonazole
ZINC000004194760	0.684873938	Salazosulfathiazole
ZINC000040917958	0.684210538	Gw804482x
ZINC000004217102	0.679487168	Nifuralide
ZINC000053294263	0.679012358	Bitopertin
ZINC000049889260	0.678571403	Lyngbyapeptin A
ZINC000001846293	0.671361506	Zoliprofen
ZINC000000007363	0.671171188	Irampanel
ZINC000001846054	0.671052634	Tioxamast
ZINC000000968328	0.669527888	Rosiglitazone

@Hak cipta milik IPB University

- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumunkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

Lampiran 13 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap ivermectin

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000169300893	1.000000000	Avermectin B1 Monosaccharide
ZINC000169366782	0.980392158	Avermectin B2 Monosaccharide
ZINC000169302726	0.980198025	27-Betamethoxy Milbemycin A4
ZINC000169302721	0.980198025	27-Alphamethoxy Milbemycin A4
ZINC000169302690	0.980198025	27-Hydroxy Milbemycin A4
ZINC000169302679	0.970297038	27-Hydroxy Milbemycin A3
ZINC000169361193	0.960396051	Ivm Aglycone
ZINC000169292347	0.960396051	Nemadectin
ZINC000169339191	0.950980365	Milbemycin Alpha25
ZINC000169302778	0.950495064	Milbemycin A4
ZINC000169345978	0.940594077	Milbemycin Beta9
ZINC000169302773	0.940594077	Milbemycin A3
ZINC000169339190	0.932692289	Milbemycin Alpha11
ZINC000169302674	0.923809528	27-Oxomilbemycin A4

@Hak cipta milik IPB University

IPB University

Lampiran 14 Daftar senyawa yang memiliki nilai similaritas tinggi terhadap oseltamivir

ID Kandidat Obat	Nilai Similaritas	Nama Lain
ZINC000003929509	0.9494949579	oseltamivir carboxylic acid
ZINC000013370140	0.7592592835	Antillatoxin
ZINC000150548865	0.7500000000	Deoxyaurilide
ZINC000169367251	0.7304347754	Bistratene B
ZINC000169290233	0.7264957428	Aspochalasin J
ZINC000028971343	0.7247706652	Cespitulactam E
ZINC000169345618	0.7155172229	Thailandepsin B
ZINC000169371244	0.7142857313	Aspochalasin I
ZINC000095615787	0.7105262876	Aurilide
ZINC000169291254	0.7079645991	Spiruchostatin A
ZINC000169296555	0.7068965435	Spiruchostatin B
ZINC000040424526	0.7058823705	Pyriperine B
ZINC000044032035	0.7039999962	Renieramycin G
ZINC000169306815	0.7027027011	Zampanolide

@Hak cipta milik IPB University

Hak Cipta Dilindungi Undang-undang

1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.

RIWAYAT HIDUP

Penulis dilahirkan di Kota Blitar, Jawa Timur pada tanggal 4 Oktober 1998 sebagai anak pertama dari pasangan bapak Satwiko Santo dan ibu Wiesda Lusina. Pendidikan sekolah menengah atas (SMA) ditempuh di SMA Negeri 1 Blitar, dan lulus pada tahun 2017. Pada tahun yang sama, penulis diterima sebagai mahasiswa program sarjana (S-1) di Program Studi Ilmu Komputer, Fakultas MIPA di IPB.

Selama mengikuti program S-1, penulis aktif menjadi bagian dari tim tanggap darurat mahasiswa, panitia kegiatan, dan relawan kemanusiaan. Penulis juga pernah mengikuti PKM Kemasyarakatan dan sampai pada tahap didanai dan mengikuti konferensi nasional.

@Hak cipta milik IPB University

IPB University



- Hak Cipta Dilindungi Undang-undang
1. Dilarang mengutip sebagian atau seluruh karya tulis ini tanpa mencantumkan dan menyebutkan sumber :
 - a. Pengutipan hanya untuk kepentingan pendidikan, penelitian, penulisan karya ilmiah, penyusunan laporan, penulisan kritik atau tinjauan suatu masalah
 - b. Pengutipan tidak merugikan kepentingan yang wajar IPB University.
 2. Dilarang mengumumkan dan memperbanyak sebagian atau seluruh karya tulis ini dalam bentuk apapun tanpa izin IPB University.