

**PENCARIAN *OPEN READING FRAME* PADA RANGKAIAN
NUKLEOTIDA DNA MENGGUNAKAN METODE PEMROGRAMAN
DINAMIS**

SKRIPSI



Oleh:
YAYU APRILIKA YUNUS
H13114502

**PROGRAM STUDI ILMU KOMPUTER DEPARTEMEN MATEMATIKA
FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM
UNIVERSITAS HASANUDDIN
JANUARI 2019**



**PENCARIAN *OPEN READING FRAME* PADA RANGKAIAN
NUKLEOTIDA DNA MENGGUNAKAN METODE PEMROGRAMAN
DINAMIS**

SKRIPSI

Diajukan sebagai salah satu syarat untuk memperoleh gelar Sarjana Komputer pada Program Studi Ilmu Komputer Departemen Matematika Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin Makassar

UNIVERSITAS HASANUDDIN

YAYU APRILIKA YUNUS

H13114502

PROGRAM STUDI ILMU KOMPUTER DEPARTEMEN MATEMATIKA

FAKULTAS MATEMATIKA DAN ILMU PENGETAHUAN ALAM

UNIVERSITAS HASANUDDIN

JANUARI 2019



LEMBAR PERNYATAAN KEOTENTIKAN

Saya yang bertanda tangan di bawah ini menyatakan dengan sungguh-sungguh bahwa skripsi yang saya buat dengan judul:

**PENCARIAN *OPEN READING FRAME* PADA RANGKAIAN
NUKLEOTIDA DNA MENGGUNAKAN METODE PEMROGRAMAN
DINAMIS**

adalah benar hasil karya saya sendiri bukan hasil plagiat dan belum pernah dipublikasikan dalam bentuk apapun

Makassar, 24 Januari 2019

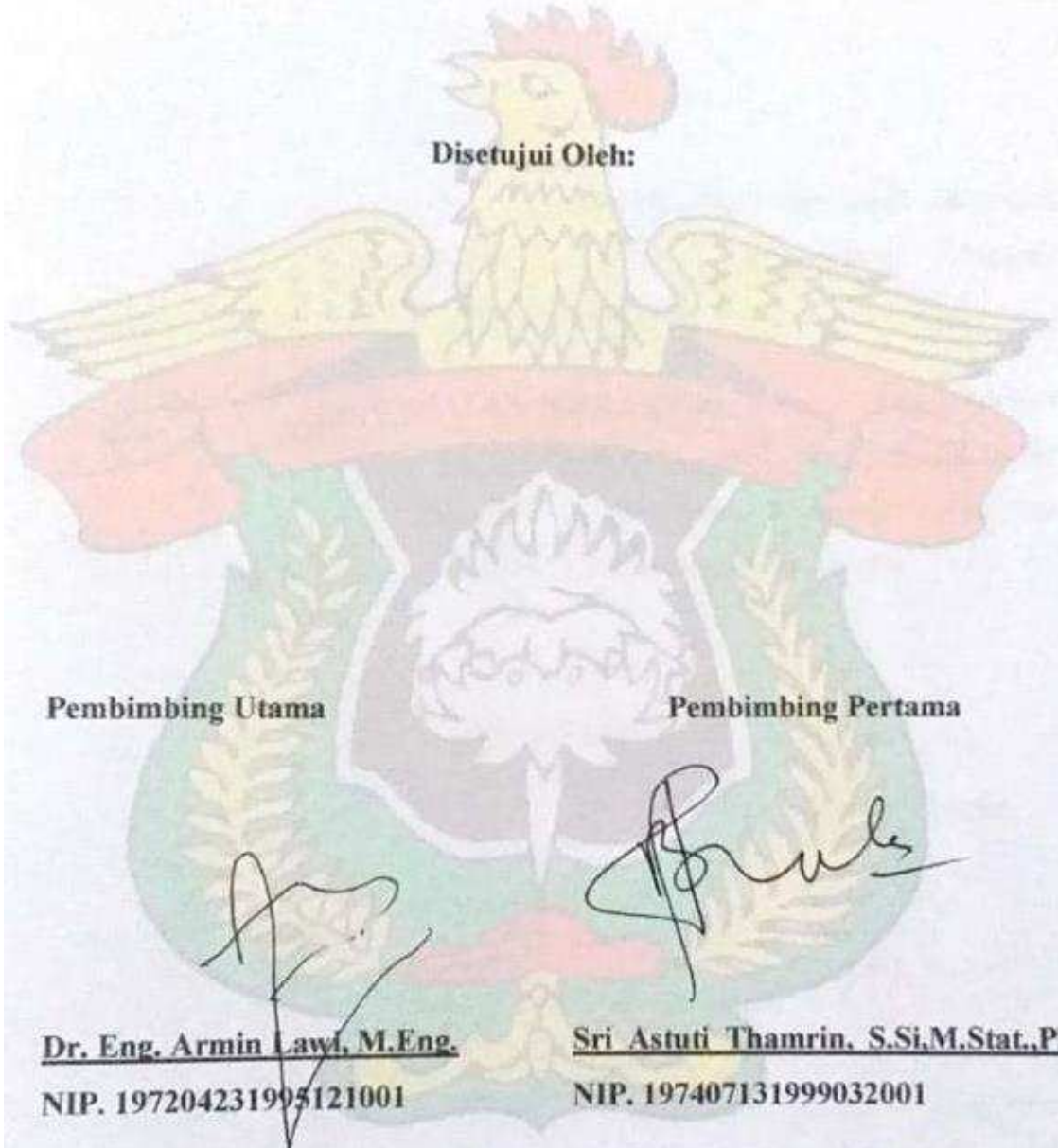


Yayu Aprilika Yunus

NIM. H13114502

**PENCARIAN *OPEN READING FRAME* PADA RANGKAIAN
NUKLEOTIDA DNA MENGGUNAKAN METODE PEMROGRAMAN
DINAMIS**

Disetujui Oleh:



Pembimbing Utama

Pembimbing Pertama

Dr. Eng. Armin Lawi, M.Eng.
NIP. 197204231995121001

Sri Astuti Thamrin, S.Si,M.Stat.,Ph.D
NIP. 197407131999032001

Pada Tanggal : 24 Januari 2019



HALAMAN PENGESAHAN

Skripsi ini diajukan oleh :

Nama : Yuyu Aprilika Yunus

NIM : H13114502

Program Studi : Ilmu Komputer

Judul Skripsi : Pencarian *Open Reading Frame* Pada Rangkaian Nukleotida DNA Menggunakan Metode Pemrograman Dinamis

Telah berhasil dipertahankan dihadapan dewan penguji dan diterima sebagai bagian persyaratan yang diperlukan untuk memperoleh gelar Sarjana Komputer pada Program Studi Ilmu Komputer Fakultas Matematika dan Ilmu Pengetahuan Alam Universitas Hasanuddin.

DEWAN PENGUJI

Tanda Tangan

1. Ketua : Dr. Eng. Armin Lawi, M.Eng (.....)
2. Sekretaris : Sri Astuti Thamrin, S.Si, M.Stat. PhD. (.....)
3. Anggota : Drs. H. Muhammad Hasbi, M.Sc (.....)
4. Anggota : Dr. Nurdin, S.Si., M.Si (.....)

Disahkan di : Makassar

Tanggal : 24 Januari 2019



KATA PENGANTAR

Bismillahirrahmanirrahim, segala puji dan syukur kehadirat Allah SWT atas segala rahmat, berkah dan ridho-Nya. Serta, salam dan shalawat senantiasa tucurahkan kepada Rasulullah Muhammad SAW. Alhamdulillah, dengan segala nikmat yang dilimpahkan kepada penulis sehingga dapat menyelesaikan tugas akhir yang berjudul **Pencarian *Open Reading Frame* Pada Rangkaian Nukleotida DNA Menggunakan Metode Pemrograman Dinamis.**

Penghargaan yang tak terhingga dan segenap rasa terima kasih yang tulus kepada kedua orang tua tercinta, Ayahanda **Yunus Yusuf** dan Ibunda **JB Nurhaeda Jaya** yang tak pernah berhenti mencurahkan segenap cinta dan kasih sayang dalam mengasuh, mendidik, membimbing, dan mengiringi perjalanan hidup penulis serta senantiasa memberi dukungan dan mendoakan segala yang terbaik bagi penulis dalam menggapai impiannya. Kepada Adik-adik penulis, **Rezki** dan **Yuni** serta untuk seluruh keluarga besar, terima kasih atas seluruh doa dan dukungannya. Semoga Allah Yang Maha Pengasih, senantiasa memberikan rahmat-Nya atas kalian.

Penulis menyadari bahwa penulisan ini dapat terselesaikan karena kehendak-Nya melalui bantuan dari berbagai pihak. Oleh karena itu, penulis juga ingin menyampaikan ucapan terima kasih yang tulus kepada :

1. Ibu **Rektor Universitas Hasanuddin** beserta jajarannya, Bapak **Dekan Fakultas Matematika dan Ilmu Pengetahuan Alam** beserta jajarannya, dan seluruh pihak birokrasi atas pengetahuan dan kemudahan-kemudahan yang diberikan, baik dalam bidang akademik maupun bidang kemahasiswaan.
2. Bapak **Prof. Dr. Amir Kamal Amir, M.Sc.**, selaku Ketua Jurusan Matematika, dan Bapak **Dr. Amran, S.Si., M.Si.**, selaku Sekretaris Jurusan, Bapak **Dr. Diaraya, M.Ak.**, selaku Kepala Program Studi Ilmu Komputer yang telah memberikan banyak bantuan selama penulis menjalani



pendidikan. Terima kasih juga untuk segenap jajaran Pegawai Akademik Jurusan Matematika atas bantuannya dalam pengurusan akademik selama ini.

3. Bapak **Dr. Eng. Armin Lawi, M.Eng**, selaku pembimbing utama, atas segala ilmu, nasehat, dan kesabaran dalam membimbing penulis serta meluangkan waktu di sela-sela rutinitas yang begitu padat hingga skripsi ini dirampungkan, dan Ibu **Sri Astuti Thamrin, S.Si, M.Stat. PhD.**, selaku pembimbing pertama, untuk segala ilmu, nasehat, dan kesabaran dalam membimbing dan mengarahkan penulis, serta bersedia meluangkan waktunya untuk mendampingi penulis sejak awal penyusunan hingga akhir perampungan skripsi ini.
4. Bapak **Drs. H. Muhammad Hasbi, M.Sc**, selaku Penasehat Akademik (PA) penulis untuk segala ilmu, nasehat, saran dan motivasi yang diberikan kepada penulis mulai dari perkuliahan hingga penyusunan skripsi ini dan sekaligus menjadi tim penguji bersama bapak **Dr. Nurdin, S.Si, M.Si**, yang telah bersedia meluangkan waktu untuk memberikan saran dan arahan kepada penulis dalam penyusunan skripsi ini.
5. **Ibu Hilal, kak Supri dan kak Edi** yang telah memberikan saran selama proses perkuliahan serta **bapak Wilem Musu** yang telah meluangkan waktu untuk membantu penulis di tengah rutinitas yang padat.
6. **Alight (Winanda, Nur Aisyah)** dan **An-nisa (Pitti, Tasya, Zahra, Hani)** yang menemani penulis sejak masa putih abu-abu, saling mendoakan dalam kebaikan, memberikan saran, tempat berbagi cerita suka-duka dan kebersamaan selama menuntut ilmu hingga menyelesaikan tugas akhir. Juga kepada **Jokka(Nuhi, Mukrimah, Murni, Anita)** yang selalu memberikan doa, dan menyemangati penulis.
7. Teman-teman seperjuangan **Ilmu Komputer 2014 (Firda, Iyam, Jo, Fuad, Luki, Hikma, Sukma, Ola, Mamet, Aspar, Nadya, Khalil, Agus, Budi, Titi, Niar, Tio, Nura, Firman, Ochi, Syam, Darul, Ij`Lal, A.Aryani , Nurfadhila, Ayu, Nurul, Yaumil, Yusri, Sarwan, Fajar, Farid, Nawir, Isman, Hajar, dll)** yang membantu dan memberi support penulis dalam penyusunan skripsi ini.

Adik Ilmu Komputer 2015, 2016, 2017, dan 2018.



9. Rekan – rekan **KKN UNHAS Gelombang 96 Jepang** yang telah memberikan moment menyenangkan dan tak terlupakan selama proses KKN yang serta memberi peluang untuk melanjutkan pendidikan.
10. Seluruh pihak yang telah berpartisipasi, baik secara langsung maupun tidak langsung dalam penyusunan skripsi ini yang tak sempat penulis sebutkan satu per satu.

Semoga seluruh bantuan yang diberikan dengan tulus mendapatkan balasan dari Allah SWT. Penulis menyadari bahwa dalam penyusunan skripsi ini masih terdapat kekurangan, sehingga dengan segala kerendahan hati penulis memohon maaf. Mudah – mudahan tulisan ini bermanfaat bagi segala pihak yang membutuhkan dan terutama untuk penulis.

Makassar, 24 Januari 2019

Yayu Aprilika Yunus



**PERNYATAAN PERSETUJUAN PUBLIKASI TUGAS AKHIR UNTUK
KEPENTINGAN AKADEMIK**

Sebagai civitas akademik Universitas Hasanuddin saya yang bertanda tangan di bawah ini:

Nama : Yuyu Aprilika Yunus

NIM : H 131 14 502

Program Studi : Ilmu Komputer

Departemen : Matematika

Fakultas : Matematika dan Ilmu Pengetahuan Alam

Jenis Karya : Skripsi

Demi pengembangan ilmu pengetahuan, menyetujui untuk memberikan kepada Universitas Hasanuddin **Hak Prediktor Royalti Non-eksklusif (*Non-exclusive Royalty- Free Right*)** atas tugas akhir saya yang berjudul:

**“PENCARIAN *OPEN READING FRAME* PADA RANGKAIAN
NUKLEOTIDA DNA MENGGUNAKAN METODE PEMROGRAMAN
DINAMIS”**

Beserta perangkat yang ada (jika diperlukan). Terkait dengan hal di atas, maka pihak universitas berhak menyimpan, mengalih-media/format-kan, mengelola dalam bentuk pangkalan data (*database*), merawat, dan memublikasikan tugas akhir saya selama tetap mencantumkan nama saya sebagai penulis/pencipta dan sebagai pemilik Hak Cipta.

Demikian pernyataan ini saya buat dengan sebenarnya.

Dibuat di Makassar pada tanggal, 24 Januari 2019

Yang menyatakan,



rilika Yunus

ABSTRAK

Selain sebagai pembawa informasi tentang sifat organisme atau sel, DNA juga berperan penting dalam pembentukan asam amino. Sebelum mengkode urutan basa nitrogen dalam DNA menjadi asam amino, terdapat proses pencarian *Open Reading Frame* (ORF) yang akan menghasilkan bagian-bagian dari sekuens DNA yang berpotensi menjadi asam amino. Salah satu masalah yang muncul dalam pencarian ORF adalah bahwa masih kurangnya program komputasi untuk mendapatkan ORF. Oleh karena itu, penelitian ini membuat program pencarian ORF dengan mengimplementasikan pemrograman dinamis yang banyak digunakan di bidang bioinformatika menggunakan aplikasi *python*. Selain itu, pentingnya menampilkan semua kemungkinan ORF dalam urutan DNA adalah dapat membantu penelitian di bidang analisis genetika.

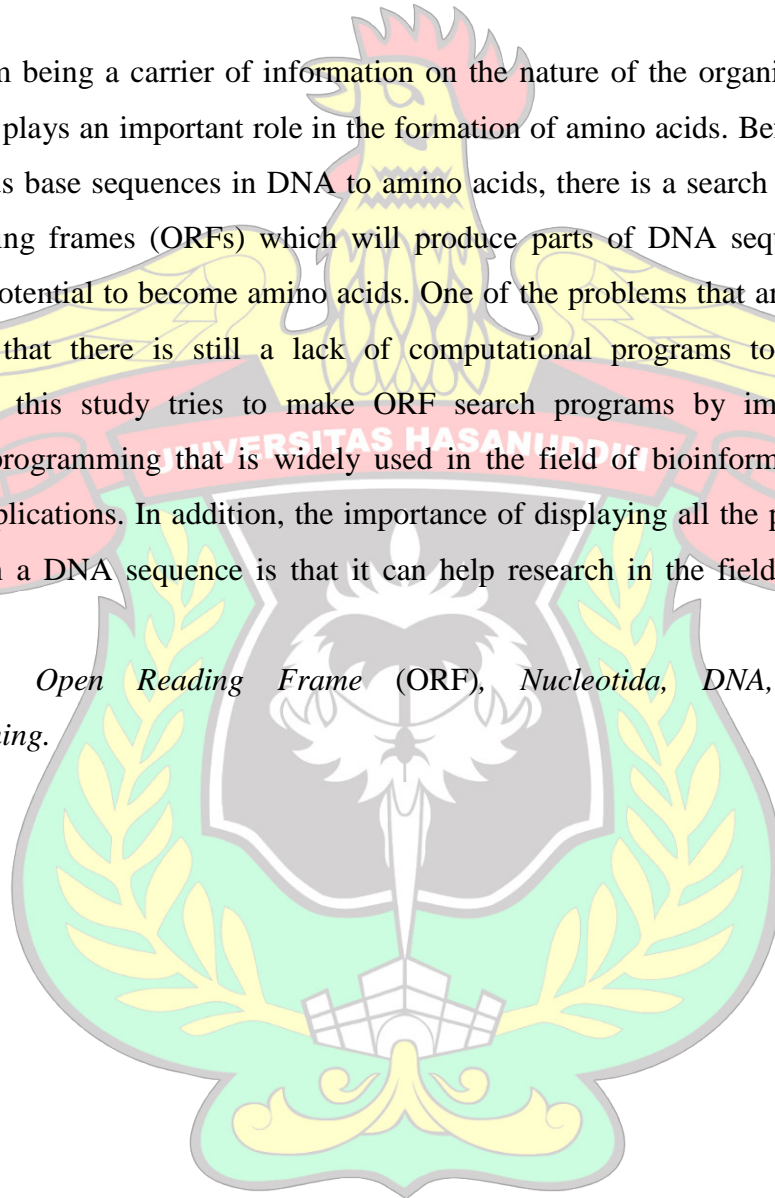
Kata Kunci : *Open Reading Frame* (ORF), *Nucleotida*, *DNA*, *Dynamic Programming*.



ABSTRACT

Aside from being a carrier of information on the nature of the organism or cell, DNA also plays an important role in the formation of amino acids. Before coding nitrogenous base sequences in DNA to amino acids, there is a search process for open reading frames (ORFs) which will produce parts of DNA sequences that have the potential to become amino acids. One of the problems that arise in ORF search is that there is still a lack of computational programs to get ORF. Therefore, this study tries to make ORF search programs by implementing dynamic programming that is widely used in the field of bioinformatics using python applications. In addition, the importance of displaying all the possibilities of ORF in a DNA sequence is that it can help research in the field of genetic analysis.

Keywords: *Open Reading Frame (ORF), Nucleotida, DNA, Dynamic Programming.*



DAFTAR ISI

LEMBAR PERNYATAAN KEOTENTIKAN	iii
HALAMAN PENGESAHAN	Error! Bookmark not defined.
KATA PENGANTAR.....	vi
PERSETUJUAN PUBLIKASI KARYA ILMIAH	ix
ABSTRAK.....	x
ABSTRACT	xi
DAFTAR ISI	xii
DAFTAR GAMBAR.....	xiv
BAB I PENDAHULUAN	1
1.1 Latar Belakang	1
1.2 Rumusan Masalah	3
1.3 Batasan Masalah.....	3
1.4 Tujuan Penelitian.....	3
1.5 Manfaat Penelitian.....	3
BAB II TINJAUAN PUSTAKA	4
2.1 Landasan Teori	4
2.1.1 <i>Deoxyribonucleic Acid</i>	4
2.1.2 <i>Central Dogma</i>	6
2.1.3 Transkripsi	7
2.1.4 <i>Open Reading Frame</i>	8
2.1.5 Pemrograman Dinamis.....	10
2.1.6 <i>Biopython</i>	12
2.2 Kerangka Konseptual	13
BAB III METODE PENELITIAN	14
Tahapan Penelitian	14
Waktu dan Lokasi Penelitian.....	15
Rancangan Sistem	15
	xii



3.4	Sumber Data	15
3.5	Instrumen Penelitian	15
BAB IV HASIL DAN PEMBAHASAN		16
4.1	Proses Penginputan Data	16
4.2	Penerapan Metode Pemrograman Dinamis	16
4.2.1	Pencarian <i>Start Codon</i>	17
4.2.2	Pencarian <i>Stop Codon</i>	19
4.3	Penentuan <i>Open Reading Frame</i>	20
BAB V KESIMPULAN DAN SARAN		24
5.1	Kesimpulan.....	24
5.2	Saran	24
DAFTAR PUSTAKA.....		25
LAMPIRAN		27
Lampiran 1		28
Lampiran 2		29
Lampiran 3		31



DAFTAR GAMBAR

Gambar 1	<i>Deoxyribonucleic acid</i>	4
Gambar 2	Nukleotida pada untaian tunggal.....	5
Gambar 3	Proses membentuk template baru dari parent DNA.....	5
Gambar 4	Central Dogma, dari DNA hingga menjadi asam amino	6
Gambar 5	Kerangka Konseptual	13
Gambar 6	Tahapan Penelitian	14
Gambar 7	Rancangan Sistem	15
Gambar 8	Proses Input Data	16
Gambar 9	Aturan posisi nukleotida.....	17
Gambar 10	Pseudocode Pencarian <i>Start Codon</i>	17
Gambar 11	Ilustrasi Perulangan	18
Gambar 12	Pengkondisian <i>start codon</i>	18
Gambar 13	Pseudocode Pencarian <i>StopCodon</i>	19
Gambar 14	Multi-dimensional lists.....	19
Gambar 15	Pengkondisian <i>Stop Codon</i>	19
Gambar 16	Pseudocode ORF.....	20
Gambar 17	Pengkondisian tidak terpenuhi	21
Gambar 18	Proses perulangan.....	21
Gambar 19	$i = 0$ dan $j = 6$	22
Gambar 20	Pengkondisian kedua.....	22
Gambar 21	Pengkondisian terpenuhi	23



BAB I

PENDAHULUAN

1.1 Latar Belakang

Gen adalah unit pembawa informasi yang mempengaruhi karakteristik suatu organisme yang diturunkan dari generasi ke generasi. Materi gen meliputi deoksiribonukleat atau *deoxyribonucleic acid (DNA)* dan asam ribonukleat atau *ribonucleic acid (RNA)*. DNA merupakan suatu polimer yang terdiri dari empat jenis monomer yang berbeda yang dinamakan nukleotida atau *nucleotide* (Campbell & Reece, 2003).

Selain sebagai pembawa informasi sifat terhadap organisme atau sel, DNA juga berperan penting dalam proses pembentukan protein. Proses perubahan setiap protein tersusun dari sejumlah asam amino dengan urutan tertentu dan setiap pembentukan asam amino dikode oleh urutan basa nitrogen yang terdapat di dalam DNA.

Sebelum mengkode urutan basa nitrogen pada DNA, terdapat proses pencarian rangka baca terbuka atau *Open Reading Frame (ORF)* yang nantinya akan menghasilkan bagian dari rangkaian DNA yang berpotensi menjadi asam amino. Terdapat beberapa website yang menyediakan halaman ORF Finder namun tidak menampilkan hasil ORF secara keseluruhan terhadap rangkaian DNA yang telah diinput.

Salah satu persoalan yang muncul pada pencarian ORF adalah masih kurangnya program komputasi untuk mendapatkan ORF dari rangkaian DNA. Oleh karena itu, pada penelitian ini mencoba untuk membuat program pencarian ORF dengan menerapkan dynamic programming yang banyak digunakan dalam bidang bioinformatika. Selain itu, pentingnya menampilkan seluruh kemungkinan ORF pada suatu rangkaian DNA ialah dapat membantu penelitian pada bidang analisis genetika.

Penelitian mengenai ORF telah dibahas pada jurnal *Open Reading Frame*

Methods for Identification of Genic Sequences (Al-Thani, dkk, 2018)

yang mengidentifikasi ORF pada gen bakteri *Escherichia coli (E.coli)*

yang berkaitan dengan resistensi ampisilin terhadap enzim β -lactamase. Selain itu, di



penelitian ini juga membandingkan beberapa metode yang berbeda dari penelitian lainnya. Salah satunya adalah penelitian *Selecting Open Reading Frames from DNA* (Zacchi, dkk, 2003) yang menerapkan cara *phage display*. Metode *phage display* adalah teknik laboratorium untuk mempelajari interaksi protein dengan DNA yang menggunakan proses virus menginfeksi bakteri untuk menghubungkan protein dengan informasi genetik yang mengkode bakteri tersebut. Di penelitian ini, metode yang digunakan menghasilkan 50% ORF dari seluruh rangkaian DNA yang digunakan. Pada jurnal *A Systematic Approach to Dynamic Programming in Bioinformatics* (Giegerich, 2000) Pemrograman dinamis salah satu metode yang sering digunakan dalam biologi molekuler komputasi. Data rangkaian seperti asam nukleat dan protein ditentukan pada skala industri saat ini sehingga keinginan untuk memberi makna pada data molekuler meningkatkan jumlah tugas analisis rangkaian. Pemrograman dinamis digunakan untuk mengumpulkan data sekuens DNA dari fragmen yang dikirim oleh mesin pengurutan otomatis dan untuk menentukan struktur intron/ekson gen *eucaroytic* serta digunakan untuk memprediksi struktur sekunder gen. Sedangkan, pada penelitian (Muda & Kadir, 2014) menggunakan metode *Dynamic Programming* dengan algoritma *Needleman-Wunsch* pada 2 data rangkaian protein yang berbeda.

Berdasarkan latar belakang yang ada maka hasil penelitian akan dirampungkan dalam bentuk tulisan (skripsi) dengan judul : **“Pencarian *Open Reading Frame* Pada Rangkaian Nukleotida DNA Menggunakan Metode Pemrograman Dinamis”**



1.2 Rumusan Masalah

Berdasarkan latar belakang diatas, dapat dirumuskan beberapa masalah berikut :

1. Bagaimana proses penginputan rangkaian DNA pada python?
2. Bagaimana menerapkan metode pemrograman dinamis pada data rangkaian DNA?
3. Bagaimana menentukan ORF berdasarkan posisi *start codon* dan *stop codon* ?

1.3 Batasan Masalah

1. Data yang digunakan dalam penelitian ini terbatas pada rangkaian DNA
2. ORF masih dalam bentuk rangkaian nukleotida DNA *frame 1*

1.4 Tujuan Penelitian

Tujuan dari penelitian ini adalah sebagai berikut :

1. Menerapkan metode Pemrograman Dinamis pada pencarian ORF
2. Menampilkan banyak rangkaian nukleotida yang termasuk ORF

1.5 Manfaat Penelitian

Hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut :

1. Program pencarian ORF yang lebih efisien
2. Sebagai penelitian awal dan bahan acuan untuk penelitian selanjutnya. Misalnya, Mengukur peningkatan *GC Content* berdasarkan ORF atau translasi ORF menjadi asam amino.



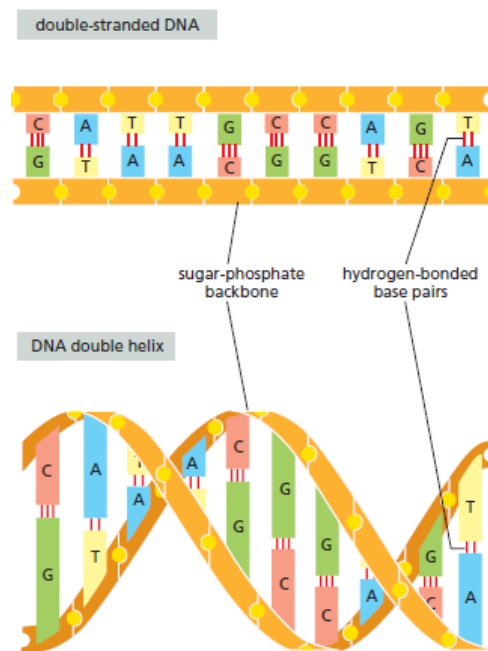
BAB II

TINJAUAN PUSTAKA

2.1 Landasan Teori

2.1.1 *Deoxyribonucleic Acid*

Deoxyribonucleic acid (DNA) adalah polimer asam nukleat yang tersusun secara sistematis dan merupakan pembawa informasi genetik yang diturunkan kepada keturunannya (Campbell & Reece, 2003). Seperti pada Gambar 1, bentuk molekul DNA berupa untaian ganda berupa rantai polimer yang panjang, tidak bercabang, berpasangan, selalu terbentuk dari empat jenis monomer yang sama berupa senyawa kimia yang dikenal sebagai nukleotida, dan keduanya dirangkai dalam urutan linear panjang yang mengkodekan informasi genetik.



Gambar 1 *Deoxyribonucleic acid*

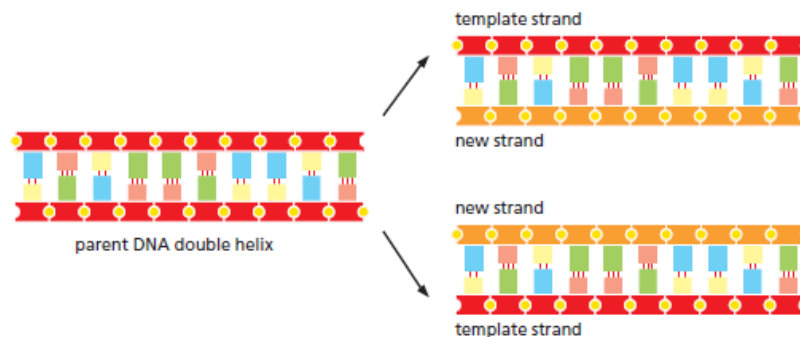
Nukleotida merupakan subunit sederhana yang menyusun DNA terdiri dari gula (deoksiribosa) dengan gugus fosfat yang melekat padanya, dan basa berupa adenin, guanin, sitosin, dan timin. Sesuai dengan empat jenis nitrogen yang berbeda, maka nukleotida diberi label A, G, C, dan T seperti Gambar 2 (Albert & Johnson, 2014).





Gambar 2 Nukleotida pada untaian tunggal

Di dalam sel yang masih hidup, untaian DNA tidak bersintesis dalam bentuk untaian bebas tetapi membentuk template baru yang berasal dari untaian *parent DNA double helix* yang telah ada sebelumnya seperti pada Gambar 3. Basa nitrogen yang menonjol dari untaian DNA double helix yang telah terpisah membentuk template baru dengan menerapkan aturan ikatan basa nitrogen yaitu A berikatan dengan T, dan C berikatan dengan G. Dengan cara ini, struktur double helix dibuat, yang terdiri dari urutan A, C, T, dan G dimana kedua untai memutar satu sama lain, membentuk DNA double helix. Template DNA menjadi mekanisme yang digunakan sel untuk melakukan replikasi urutan nukleotida. Proses replikasi selain digunakan pada dogma pusat, juga membantu pertumbuhan sel pada saat terjadi kerusakan sel ataupun terdapat luka.



Gambar 3 Proses membentuk template baru dari parent DNA

Selain berperan dalam pewarisan sifat, informasi genetik yang ada pada DNA dapat digunakan untuk mengarahkan proses sintesis protein yang melibatkan *ribonucleic acid* atau RNA. Sebuah proses yang menjelaskan tentang perubahan gen dari DNA menjadi RNA kemudian menjadi protein disebut *central*

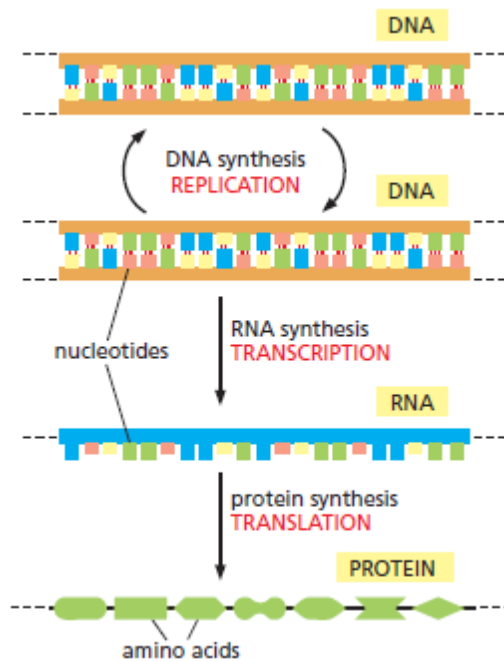
al ini merupakan konsep dasar bagi biologi molekuler yang melibatkan el, dari bakteri hingga manusia.



2.1.2 Central Dogma

Central dogma adalah konsep dasar yang menjelaskan proses perubahan DNA menjadi protein yang melibatkan RNA sebagai perantara. Proses ini dimulai dengan polimerisasi template yang disebut transkripsi, di mana segmen dari urutan DNA digunakan sebagai template untuk sintesis molekul dari polimer ribonukleat asam, atau RNA. Kemudian, dalam proses penerjemahan yang lebih kompleks, molekul RNA ini mengarahkan ke proses sintesis polimer yang nantinya akan menjadi asam amino (Albert & Johson, 2014).

Pada RNA, kandungan gula yang digunakan adalah ribose berbeda dengan DNA yang mengandung deoksiribosa. Selain itu, satu dari empat basa nitrogen RNA yang digunakan berbeda dengan DNA yaitu tidak mengandung timin namun urasil. Sehingga ikatan basa A berkaitan dengan U.



Gambar 4 Central Dogma, dari DNA hingga menjadi asam amino

Gambar 4 menjelaskan proses dari central dogma yang diawali oleh DNA yang direplikasi hingga menjadi protein (Albert & Johson, 2014). Replikasi proses awal yaitu penyalinan pasangan basa yang menghasilkan satu DNA baru yang sama dengan molekul DNA yang menjadi acuan. Terjadi replikasi DNA melalui tiga model, yaitu



semikonservatif, konservatif dan dispersive. Kemudian, transkripsi menjadi tahap selanjutnya yang biasa disebut sebagai sintesis RNA menggunakan hasil replikasi dari DNA. Pembentukan RNA pada tahap ini melibatkan enzim polymerase. Informasi dalam RNA disalin ke bentuk yang berbeda namun tetap dalam penulisan urutan nukleotida.

Langkah selanjutnya adalah translasi yang berlangsung di dalam sitoplasma dan ribosom dimana sintesis polipeptida atau protein yang diarahkan oleh RNA menginterpretasikan suatu pesan genetik dan membentuk protein yang sesuai dengan pesan tersebut. Pesan tersebut berupa kode genetik (*codon*) disepanjang molekul mRNA yang interpreturnya RNA transfer (tRNA). Fungsi tRNA adalah mentransfer asam-asam amino dari sitoplasmanya ke ribosom tempat molekul protein dibentuk sewaktu protein disintesis (Yuwono, 2012).

2.1.3 Transkripsi

Transkripsi adalah sintesis RNA yang menggunakan informasi dalam DNA. Asam nukleat ditulis dalam berbagai bentuk dengan bahasa yang sama, dan informasinya hanya ditranskripsi, atau dengan kata lain penulisan ulang dari DNA ke RNA. Untuk gen pengkode protein, molekul RNA yang dihasilkan adalah transkrip dari instruksi pembuatan protein gen. Molekul RNA jenis ini disebut messenger RNA (mRNA) karena membawa pesan genetik dari DNA ke pemrosesan sintesis protein sel. Transkripsi adalah istilah umum untuk sintesis segala jenis RNA pada template DNA (Campbell & Reece, 2003).

Jika DNA berbentuk untaian ganda, pada RNA berlaku sebaliknya yaitu hanya dalam bentuk untaian tunggal. RNA dalam sel dibuat oleh transkripsi DNA, suatu proses yang memiliki kesamaan tertentu dengan proses replikasi DNA yang dibahas pada landasan teori central dogma. Transkripsi dimulai dengan pembukaan dan penguraian sebagian kecil dari DNA double helix untuk mengekspos basa pada untaian DNA. Salah satu dari dua helai heliks ganda DNA kemudian bertindak sebagai template untuk sintesis molekul RNA. Seperti pada replikasi DNA, urutan nukleotida rantai RNA ditentukan oleh pasangan basa komplementer antara nukleotida yang masuk dan template DNA.

Transkripsi dapat dibagi menjadi tiga tahap yaitu: Inisiasi (permulaan), Elongasi (pemanjangan), dan terminasi (pengakhiran). Enzim yang bertanggung



jawab atas transkripsi adalah RNA polimerase, yang bergerak di sepanjang gen mulai dari promotor hingga persis di belakang terminator.

1. *Inisiasi*. transkripsi diawali oleh promotor yaitu daerah DNA tempat RNA polymerase melekat. Fungsi promotor menentukan dimana transkripsi dimulai dan juga menentukan DNA yang digunakan sebagai cetakan.
2. *Elongasi*. Pada saat RNA polimerase bergerak di sepanjang DNA, enzim ini akan terus membuka pilinan untai ganda tersebut, memperlihatkan kira-kira 10-20 basa DNA sekaligus untuk berpasangan dengan nukleotida RNA. Enzim ini menambahkan nukleotida ke ujung 3' dari molekul RNA yang sedang tumbuh. Pada saat sintesis RNA berlangsung, untai ganda DNA terbentuk kembali dan molekul RNA baru akan lepas dari cetakan DNA-nya.
3. *Terminator*. Transkripsi berlangsung sampai RNA polimerase selesai menyalin urutan DNA yang disebut terminator. Terminator merupakan suatu urutan RNA yang berfungsi sebagai sinyal pengakhiran transkripsi.

Ketika promotor mengawali dimana transkripsi dimulai, maka pada tahap ini pula inisiasi ORF memulai pencariannya terhadap untaian DNA yang memiliki potensi protein dan nantinya akan diterjemahkan ke bentuk asam amino ketika memasuki proses translasi.

2.1.4 *Open Reading Frame*

Open Reading Frame (ORF) adalah urutan triplet nukleotida yang dibaca sebagai *codon* yang menentukan asam amino, satu untaian DNA memiliki tiga frame pembacaan yang memungkinkan. Panjang dari ORF dapat mengindikasikan daerah pengkodean protein kandidat dalam urutan DNA (Nature.com). Bentuk ORF dapat berupa DNA atau mRNA, hal ini bergantung pada kebutuhan peneliti. Suatu ORF dapat dikatakan ORF jika satu rangkaian dengan rangkaian lainnya tidak menggunakan *startcodon* yang sama. Kemudian, wilayah DNA yang menyandikan protein ditranskripsikan ke RNA pembawa pesan dan diterjemahkan

protein. Dengan memeriksa urutan DNA maka dapat menentukan urutan inian asam amino yang akan muncul pada protein akhir.



Codon adalah sejenis kode genetik, yang merupakan seperangkat aturan, dengan mana informasi tertentu dikodekan dalam materi genetik berupa urutan DNA atau mRNA, yang mana *codon* ini diterjemahkan ke dalam protein. Protein ini terdiri dari asam amino yang dirangkai dalam urutan tertentu. Setiap perubahan dalam urutan ini menandakan perubahan dalam pengkodean, biasanya terlihat pada kasus mutasi genetik. Setiap *codon* terdiri dari tiga basa, yang merupakan kode untuk asam amino tunggal, dan mereka membentuk pemetaan yang dikodekan dalam tRNA (Nature.com). Secara keseluruhan, ada empat basa yang ada dalam DNA yaitu adenin, sitosin, guanin, dan timin. Ciri – ciri dari suatu ORF adalah sebagai berikut :

1. Kodon inisiasi (*start codon*) yaitu ATG pada DNA atau AUG pada mRNA.

Sebuah *start codon* dalam DNA memulai penerjemahan asam amino pertama dalam rantai polipeptida. Tiga basa pertama dari urutan pengkodean mRNA yang akan diterjemahkan ke dalam protein, adalah tempat kodon inisiasi berada. Letak ini merupakan wilayah di mana transkripsi dimulai dan berakhir tepat sebelum kodon memulai wilayah pengkodean asam amino.

ATG adalah *start codon* yang paling umum di DNA dan kode untuk *metionin* asam amino (Met) dalam *eukariota* dan *formil metionin (fMet)* dalam *prokariota*. Selama sintesis protein, ATG yang diubah menjadi AUG pada mRNA yang kemudian dikenali tRNA sebagai *start codon* dengan bantuan beberapa faktor inisiasi dan memulai translasi mRNA (Staff, 2018).

2. Terdapat serangkaian urutan basa nukleotida yang menyusun tiap ORF yang ditemukan.
3. Kodon terminasi (*Stop codon*) yaitu TAA, TAG, TGA pada DNA atau UAA, UAG, UGA pada RNA. Sebuah *stop codon* sebenarnya adalah tiga basa nukleotida dalam RNA pembawa pesan yang menandai penghentian proses penerjemahan. Sebagian besar codon pada mRNA berhubungan dengan penambahan asam amino spesifik ke rantai protein yang sedang



tumbuh, dalam urutan tertentu. Pembentukan rantai asam amino akan berhenti pada titik ini (BiologyWiseStaff, 2018).

Dalam kode genetik standar, ada tiga jenis *stop codon* yang berbeda, yaitu, TAA (timin - adenin - adenin), TAG (timin - adenin - guanin), dan TGA (timin - guanin - adenin). Ketika kodon ini menjalani transkripsi DNA, mereka berubah menjadi UAA (urasil - adenin - adenin), UAG (urasil - adenin - guanin), dan UGA (urasil - guanin - adenin). Ini adalah *stop codon* yang biasanya ditemukan dalam molekul RNA.

Posisi *stop codon* tidak tetap, karena fakta bahwa panjang molekul protein, dan jumlah asam amino yang ditetapkan bervariasi. Dengan demikian, *stop codon* pada DNA dan RNA dapat ditemukan pada berbagai interval dalam keseluruhan panjang rantai. *Stop codon* dapat dengan mudah diidentifikasi ketika molekul DNA diurutkan, dan dapat digunakan untuk mengidentifikasi lokasi dalam kode genetik, yang secara khusus sesuai dengan jenis protein tertentu.

2.1.5 Pemrograman Dinamis

Dynamic Programming atau pemrograman dinamis, seperti metode *divide-and-conquer* yang memecahkan masalah dengan menggabungkan solusi ke sub-masalah. Algoritma *divide-and-conquer* menguraikan masalah menjadi sub-masalah, memecahkan sub-masalah secara rekursif, dan kemudian menggabungkan setiap solusi untuk memecahkan masalah asli. Namun, pemrograman dinamis berlaku ketika sub-masalah menjadi *overlap* (kondisi dimana sub-masalah terbagi menjadi beberapa sub-masalah yang sama). Algoritma pemrograman dinamis menyelesaikan setiap sub-masalah hanya sekali dan kemudian menyimpan jawabannya dalam sebuah tempat, dengan menghindari pekerjaan mengkomputasi ulang jawaban setiap kali menyelesaikan setiap bagian sub-masalah (Cormen, dkk, 2009). Selain konsep rekursif, pemrograman dinamis juga dikenal dengan istilah memoisasi atau pencatatan yang merupakan penggambaran teknik pengoptimalan dengan menyimpan hasil hitung ke dalam

memo atau dengan kata lain disimpan pada sebuah variable kemudian hasil panggil kembali ketika akan digunakan (rahmadya, 2018).



Pemrograman dinamis adalah metode optimasi matematis, dan metode pemrograman komputer. Dalam konteks ini mengacu pada penyederhanaan masalah yang rumit dengan memecahnya menjadi subproblem sederhana dengan cara rekursif. Submasalah ini kemudian ditangani satu per satu, sehingga jawaban untuk masalah kecil digunakan untuk menyelesaikan yang lebih besar (Muda & Kadir, 2014).

Dynamic Programming biasanya diterapkan untuk masalah pengoptimalan. yang dapat memiliki banyak solusi yang memungkinkan. Setiap solusi memiliki nilai, dan ingin ditemukan solusi dengan nilai optimal (minimum atau maksimum). Saat mengembangkan algoritma pemrograman dinamis, kami mengikuti urutan empat langkah:

1. Karakterisasi struktur pada solusi optimasi.
2. Mendefinisikan nilai solusi optimal secara rekursif.
3. Menghitung nilai solusi optimal pada model *bottom-up*.
4. Menyusun solusi optimal dari informasi hasil perhitungan

Langkah 1 sampai langkah 3 adalah dasar dynamic – programming dalam menemukan solusi untuk suatu problem, langkah ke- 4 dapat dilakukan jika nilai solusinya optimal diperlukan (Cormen, 2009).

Pemrograman dinamis bergantung pada kebutuhan waktu suatu algoritma terhadap ukuran masukan n yang dikenal dengan istilah kompleksitas algoritma. Ada dua macam kompleksitas algoritma yaitu kompleksitas waktu dan kompleksitas ruang. Kompleksitas waktu diukur dari jumlah tahapan komputasi yang dibutuhkan untuk menjalankan program sebagai fungsi dari ukuran masukan n dimana n adalah jumlah data yang diproses oleh sebuah program (Munir, 2015). Jumlah tahapan komputasi dihitung dari berapa banyak suatu operasi dilaksanakan di dalam sebuah algoritma. Operasi yang mendasari suatu algoritma, yaitu :

1. Pencarian di dalam array atau list
2. Pengurutan atau *sorting*
3. Penjumlahan atau perkalian 2 buah matriks.



2.1.6 *Biopython*

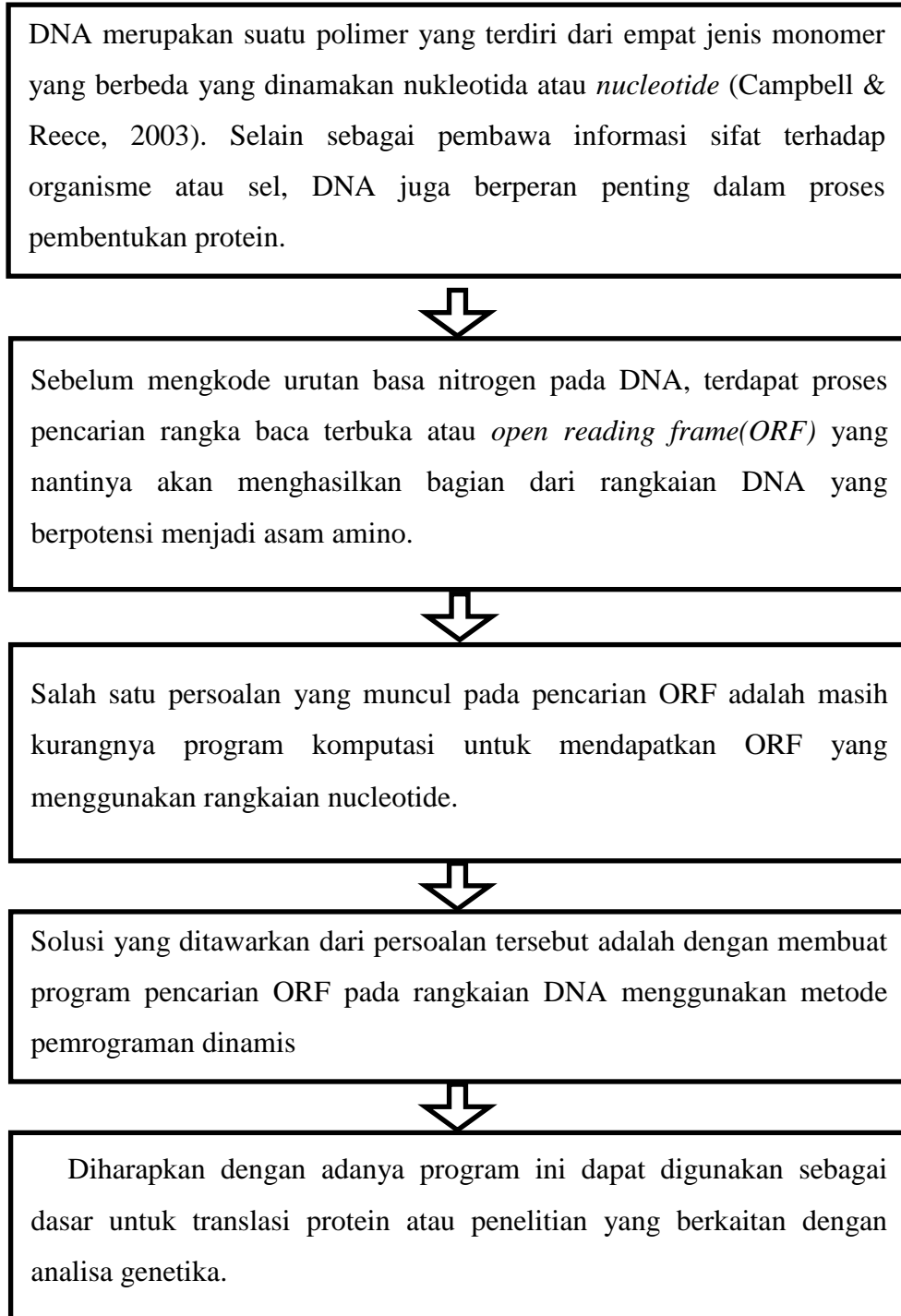
Biopython adalah seperangkat pustaka yang ada pada perangkat lunak python untuk komputasi pada bidang bioinformatika. Python adalah bahasa pemrograman yang lebih menekankan pada keterbacaan kode agar lebih mudah untuk memahami sintaks. Bahasa ini muncul pertama kali pada tahun 1991, dirancang oleh seorang bernama Guido van Rossum yang hingga saat ini masih dikembangkan oleh Python Software Foundation. Situs web *biopython* menyediakan sumber daya online untuk modul, skrip, dan tautan web untuk penggunaan dan penelitian bioinformatika (Chang, dkk, 2018).

Sequence atau rangkaian biologis dapat dikatakan sebagai objek sentral dalam Bioinformatika, dan dalam pustaka ini terdapat beberapa mekanisme untuk menangani sekuens, yaitu package *Sequence Object*, *Sequence Annotation Object* dan *Sequence Input/Output* dan sebagainya. Package *Sequence Input/Output* yang menyediakan interface sederhana untuk proses *input output* terhadap input data rangkaian dengan berbagai format data, salah satunya *fasta* yang menjadi format data di dalam penelitian ini. Package ini ditandai dengan kode program *SeqIO*.



2.2 Kerangka Konseptual

Berdasarkan uraian yang telah di jelaskan sebelumnya maka dapat digambarkan kerangka konseptual pada Gambar 5 :



Gambar 5 Kerangka Konseptual

