

Analisis Mekanisme Sintesis Protein dari Sekuens DNA Menggunakan Prinsip Graf dan Kombinatorial

Bagas Sambega Rosyada - 13522071¹

Program Studi Teknik Informatika

Sekolah Teknik Elektro dan Informatika

Institut Teknologi Bandung, Jl. Ganesha 10 Bandung 40132, Indonesia

¹13522071@mahasiswa.itb.ac.id

Abstrak—Protein merupakan zat kimiawi penting yang diperlukan oleh tubuh manusia. Protein pada umumnya tersusun atas kumpulan asam amino yang bergabung membentuk suatu makromolekul. Mekanisme pembuatan protein disebut juga sebagai sintesis protein. Dalam proses sintesis protein, jenis kodon yang berbeda akan menghasilkan jenis asam amino yang berbeda pula. Makalah ini bertujuan untuk menganalisis bagaimana graf dapat memodelkan proses sintesis protein dan mencari banyak kemungkinan cara membentuk asam amino dari informasi kodon yang tersedia.

Kata kunci—Asam amino, graf, kodon, kombinasi, sintesis protein.

I. PENDAHULUAN

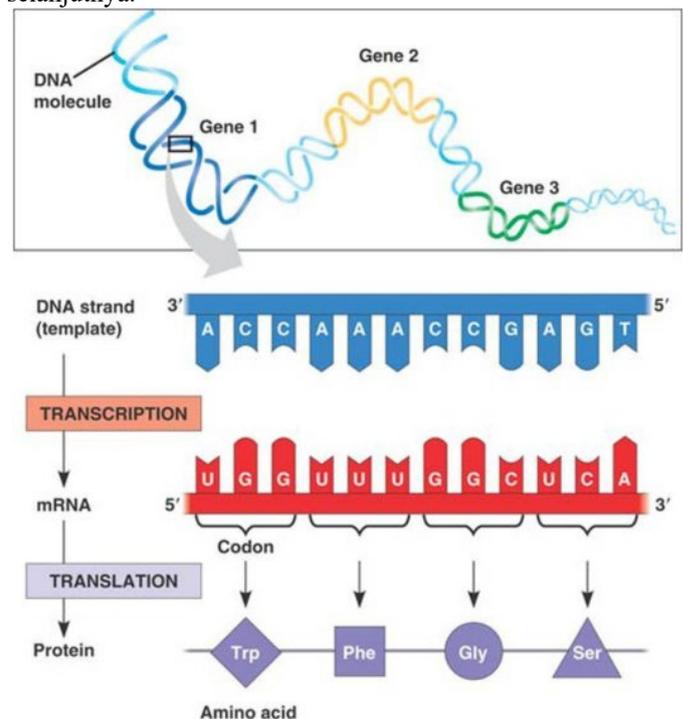
Protein merupakan struktur kimiawi esensial yang diperlukan oleh kebanyakan makhluk hidup bersel banyak (multiseluler), salah satunya manusia. Protein tersusun atas asam amino yang dibentuk oleh susunan basa nitrogen pada RNA yang sebelumnya berasal dari DNA, zat utama penyusun tubuh yang membawa informasi genetik setiap manusia [1].

Fungsi protein sebagai salah satu zat gizi makro yang diperlukan oleh tubuh manusia mendorong perlunya analisis dan penelitian mendalam mengenai cara kerja dan proses pembentukan protein. Mekanisme sintesis protein sangat kompleks bergantung pada jenis protein yang akan dibuat dan juga kondisi lingkungan saat terjadinya sintesis, namun mekanisme sintesis unsur pembentuknya, asam amino, relatif mudah untuk dilacak dan dapat diimplementasikan dalam bentuk graf berarah.

Graf berarah digunakan dalam menentukan pola pembentukan asam amino yang berasal dari sekuens DNA karena proses sintesis asam amino berlangsung hanya satu arah. Sekuens DNA yang sudah dipecah dan diubah menjadi dalam bentuk kodon—kombinasi dari tiga nukleotida dalam bentuk RNA [2]. Kodon-kodon tersebut akan diubah menjadi asam amino berdasarkan klasifikasi kodon dan asam amino yang dihasilkannya.

Terdapat banyak sekali jenis asam amino yang ada, namun hanya 20 jenis yang umum ditemukan dalam protein dalam manusia, yang dihasilkan dari kombinasi nukleotida tertentu dalam kodon. Pada umumnya terdapat 3-4 jenis kodon dapat membentuk satu jenis asam amino yang sama, namun terdapat beberapa jenis asam amino yang dapat dibentuk oleh hanya dua kombinasi kodon saja atau bahkan ada kodon yang tidak dapat membentuk asam amino sama sekali. Kombinasi antarkodon

dalam membentuk asam amino dapat dihitung menggunakan teori kombinatorika yang akan dianalisis lebih dalam di bab selanjutnya.



Gambar 1. Ilustrasi proses sintesis protein, diambil dari [2].

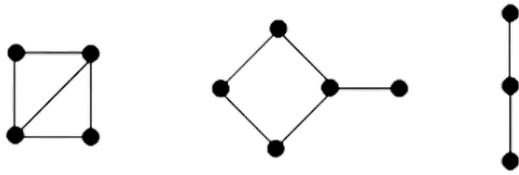
II. LANDASAN TEORI

A. Teori Graf

1. Teori dan Jenis Graf

Graf adalah struktur diskrit yang terdiri atas simpul dan sisi yang menghubungkan sisi lainnya [3]. Graf dinotasikan dalam $G = (V, E)$, yaitu graf tersusun atas sekumpulan set dari *vertices* (simpul) $V = \{v_1, v_2, v_3, \dots, v_n\}$ dan *edges* (sisi) $E = \{e_1, e_2, e_3, \dots, e_n\}$ yang menjadi penghubung antarsimpul.

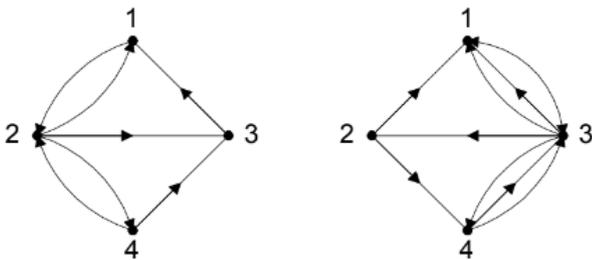
Graf dapat digolongkan menjadi beberapa jenis berdasarkan kriterianya. Berdasarkan ada-tidaknya sisi ganda dan gelang, graf terbagi menjadi graf sederhana dan graf tak-sederhana. Graf sederhana adalah graf yang tidak memiliki sisi ganda ataupun sisi gelang. Graf tak-sederhana dapat dibagi lagi menjadi graf ganda, yaitu graf yang memiliki sisi ganda, dan graf semu, yaitu graf yang memiliki sisi gelang.



Gambar 2. Ilustrasi graf, diambil dari <https://informatika.stei.itb.ac.id/~rinaldi.munir/>.

Berdasarkan ada tidaknya arah dalam graf, graf terbagi menjadi dua, yaitu graf berarah dan graf tak-berarah. Graf berarah adalah graf yang sisi-sisinya diberi orientasi arah (busur). Graf tak-berarah sisinya tidak memiliki orientasi arah.

Pada graf berarah, notasi sisi antara (u, v) dengan (v, u) melambangkan dua sisi yang berbeda. Pada sisi (u, v) , simpul u merupakan simpul asal dan simpul v adalah simpul terminal, sementara pada (v, u) simpul v adalah simpul asal dan simpul u adalah simpul terminalnya [4]. Sementara pada graf tak-berarah, sisi (u, v) dan sisi (v, u) adalah satu sisi yang sama.



Gambar 3. Ilustrasi graf berarah, diambil dari <https://informatika.stei.itb.ac.id/~rinaldi.munir/>.

2. Terminologi Graf

- Ketetanggaan (*adjacent*), dua buah simpul bertetangga jika keduanya berhubungan langsung.
- Berisian (*incidency*), sebuah sisi berisian dengan sebuah simpul jika terdapat sebuah simpul v_1 dan sebuah sisi $e = (v_i, v_j)$.
- Simpul terencil, simpul yang tidak memiliki sisi yang berisian dengan simpul tersebut.
- Graf kosong (*null graph*), graf yang himpunan sisinya adalah himpunan kosong (tidak memiliki sisi sama sekali).
- Derajat (*degree*), jumlah sisi yang berisian dengan suatu simpul.
- Lintasan (*path*), sisi-sisi yang ditempuh dari suatu simpul awal ke simpul tujuan. Panjang lintasan adalah banyak sisi dalam lintasan tersebut.
- Siklus/sirkuit (*cycle/circuit*), lintasan yang memiliki simpul awal dan simpul akhir yang sama.
- Keterhubungan (*connected*), jika setiap simpul pada graf satu dan graf lainnya memiliki lintasan.
- Upagraf (*subgraph*), sebagian graf yang berasal dari graf utamanya. Komplemen upagraf adalah graf lain yang jika digabungkan dengan upagraf akan membentuk graf utamanya.
- Upagraf merentang (*spanning subgraph*), upagraf yang mengandung semua simpul dari graf utamanya.

- Cut-set*, himpunan sisi yang menyebabkan suatu graf menjadi graf tak-terhubung.
- Graf berbobot, graf yang setiap sisinya memiliki nilai.

3. Jenis Graf Khusus

Graf memiliki beberapa jenis graf khusus yang umum ditemukan, di antaranya,

- Graf lengkap (K_n), graf sederhana yang setiap simpulnya memiliki sisi ke semua simpulnya. Jumlah sisi pada graf lengkap dengan n simpul adalah $n(n-1)/2$.
- Graf lingkaran (C_n), graf sederhana yang jumlah sisinya sama dengan jumlah simpulnya. Setiap simpul pada graf berderajat dua.
- Graf teratur, graf yang setiap simpulnya memiliki derajat yang sama. Jika jumlah simpul pada graf adalah n , maka derajat setiap simpul juga adalah n . Jumlah sisi pada graf adalah $e = nr/2$.
- Graf bipartit ($K_{m,n}$), graf yang himpunan simpulnya dapat dipisah menjadi dua himpunan bagian m dan n , sehingga setiap sisi pada K dapat menghubungkan simpul di m dan simpul di n .

B. Kombinatorial

Kombinatorial adalah cabang matematika diskrit yang mempelajari pengaturan objek-objek.

Kombinatorial berkaitan erat dengan perhitungan seluruh kemungkinan yang ada dalam pengaturan peletakan objek. Secara umum, terdapat dua cara/teknik perhitungan dalam kombinatorial, yaitu kaidah perkalian (*rule of product*) dan kaidah penjumlahan (*rule of sum*).

- Kaidah Perkalian (*Rule of Product*). Kaidah perkalian menunjukkan bahwa jika ada n pilihan untuk suatu aksi dan ada m pilihan untuk aksi lain, dan aksi n bisa dilakukan bersama m , maka banyak kemungkinan dihitung sebagai $m \times n$.
- Kaidah Penjumlahan (*Rule of Sum*). Kaidah penjumlahan menunjukkan bahwa jika ada n pilihan aksi dan m pilihan aksi lain, dan antara n dan m tidak bisa dilakukan dalam waktu bersamaan, maka banyak kemungkinan dihitung sebagai $m + n$ [5].

Kaidah perhitungan tersebut memiliki perluasan dalam hal yang lebih spesifik untuk menangani kasus yang lebih khusus.

1. Permutasi

Permutasi adalah kaidah perhitungan yang memperhatikan urutan. Permutasi adalah salah satu perluasan kaidah perkalian [4]. Berdasarkan kaidah perkalian, permutasi n elemen adalah,

$$P_n = n(n-1)(n-2) \dots (2)(1) = n! \quad (1)$$

Jika terdapat n elemen dan akan ditempatkan di sejumlah r tempat, maka permutasi r dari n dirumuskan sebagai,

$$P_r^n = \frac{n!}{(n-r)!} \quad (2)$$

Jika terdapat n buah elemen dengan banyak n_1, n_2, \dots, n_k adalah banyak elemen untuk tiap elemen di dalam n , maka banyak kemungkinan yang dapat dibentuk adalah,

$$P(n; n_1, n_2, \dots, n_k) = \frac{n!}{n_1! n_2! \dots n_k!} \quad (3)$$

2. Kombinasi

Kombinasi adalah kaidah perhitungan yang tidak memperhatikan urutan. Jika terdapat n buah elemen dan akan ditempatkan di r tempat, dengan urutan tidak diperhatikan, maka

$$C_r^n = \frac{n!}{r!(n-r)!} \quad (4)$$

Jika di dalam r bisa ditempatkan lebih dari 1 elemen, maka rumus kombinasi akan dimodifikasi menjadi,

$$C(n+r-1, r) = C(n+r-1, n-1) \quad (5)$$

C. Sintesis Protein

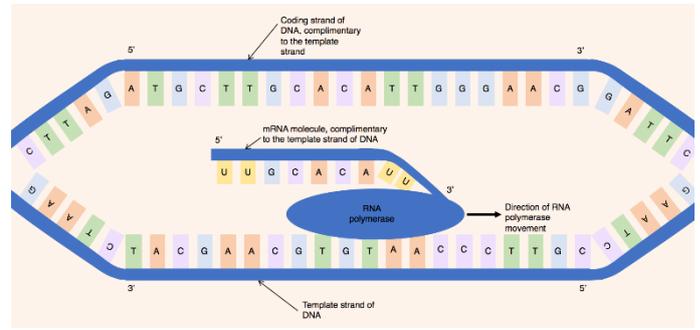
Sintesis protein adalah mekanisme pembentukan protein di dalam suatu organisme. Protein sendiri merupakan molekul biologis fungsional yang terdiri dari satu atau lebih rantai polipeptida menjadi sebuah struktur tiga dimensi [2]. Rangkaian polipeptida tersusun atas monomer (penyusun tunggal) berupa asam amino.

Proses sintesis protein akan mengubah suatu untaian basa dari DNA menjadi asam amino. Terdapat dua tahapan utama pada sintesis protein, yaitu transkripsi dan translasi, seperti ditunjukkan pada gambar (1).

DNA (*deoxyribonucleat acid*, asam deoksiribonukleat) adalah molekul asam nukleat beruntai ganda yang tersusun atas monomer nukleotida dan dapat bereplikasi. DNA membawa informasi genetika dan dapat menentukan struktur pewarisan sifat dari protein-protein suatu sel. DNA memiliki dua untaian, yaitu untaian 3' ke 5', dan untaian 5' ke 3'. Untaian DNA yang akan digunakan untuk sintesis protein adalah untaian dengan arah 3' ke 5' [2].

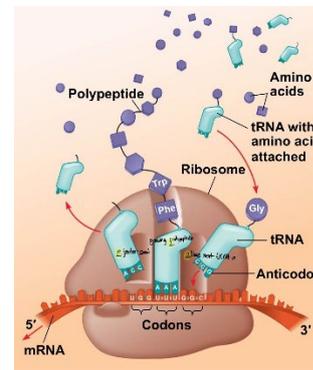
DNA tersusun atas 4 jenis nukleotida, yaitu A (Adenin), T (Timin), G (Guanin), dan C (Sitosin). Setiap basa pada DNA memiliki pasangannya masing-masing, A dengan T dan G dengan C. Hal ini menunjukkan pula bahwa DNA merupakan untaian ganda karena setiap basa di salah satu untaian harus berikatan dengan basa lain. Perbedaan dengan RNA adalah pada basa RNA, tidak ada T (Timin), namun tergantikan oleh U (Urasil) [6].

Proses transkripsi diawali dengan masuknya RNA polimerase yang memisahkan ikatan untaian DNA. Lalu RNA polimerase akan bergerak dari ujung 3' ke 5' sebagai cetakan DNA, dan membaca satu per satu basa di untaian tersebut, untuk membentuk komplemen dari basa DNA tersebut dalam basa RNA dengan arah sebaliknya, yaitu 5' ke 3'. Sehingga jika ditemukan basa A, akan dibentuk basa U, basa T menjadi basa A, basa G menjadi basa C, dan basa C menjadi basa G. Hasil pembacaan dan komplemen DNA tadi disebut sebagai m-RNA (*messenger-RNA*), karena berperan sebagai pembawa pesan dari DNA ke ribosom untuk selanjutnya dibuat protein [7].



Gambar 4. Proses transkripsi DNA, diambil dari https://en.wikipedia.org/wiki/Protein_biosynthesis.

Setelah proses transkripsi selesai, maka akan dilanjutkan dengan proses translasi. Proses translasi adalah proses pembuatan asam amino yang berasal dari m-RNA di dalam ribosom. Ribosom sendiri adalah organel sel yang berperan membentuk protein. Proses translasi menerjemahkan basa-basa yang ada di dalam m-RNA dan membentuk polipeptida sesuai pesan yang terbentuk tersebut. Proses translasi melibatkan lebih banyak makromolekul dan unsur lain, seperti t-RNA, antikodon, dan asam amino [2].



Gambar 5. Ilustrasi proses translasi, diambil dari [2].

Dalam mekanisme proses translasi, akan dihasilkan asam amino-asam amino yang menjadi monomer pembentuk polipeptida. Asam amino yang dihasilkan spesifik dan berasal dari kodon-kodon tertentu saja. Terdapat sebanyak 20 asam amino yang umum ditemukan dan berasal dari proses sintesis protein, dengan tingkat kekerapan di alam yang berbeda-beda.

		1st base					
		U	C	A	G		
2nd base	U	UUU Phenylalanine	UCU Serine	UAU Tyrosine	UGU Cysteine	U	
	UUC Phenylalanine	UCC Serine	UAC Tyrosine	UGC Cysteine	C		
	UUA Leucine	UCA Serine	UAA Stop	UGA Stop	A		
	UUG Leucine	UCG Serine	UAG Stop	UGG Tryptophan	G		
C	CUU Leucine	CCU Proline	CAU Histidine	CGU Arginine	U		
	CUC Leucine	CCC Proline	CAC Histidine	CGC Arginine	C		
	CUA Leucine	CCA Proline	CAA Glutamine	CGA Arginine	A		
	CUG Leucine	CCG Proline	CAG Glutamine	CGG Arginine	G		
A	AUU Isoleucine	ACU Threonine	AAU Asparagine	AGU Serine	U		
	AUC Isoleucine	ACC Threonine	AAC Asparagine	AGC Serine	C		
	AUA Isoleucine	ACA Threonine	AAA Lysine	AGA Arginine	A		
	AUG Methionine (Start)	ACG Threonine	AAG Lysine	AGG Arginine	G		
G	GUU Valine	GCU Alanine	GAU Aspartic Acid	GGU Glycine	U		
	GUC Valine	GCC Alanine	GAC Aspartic Acid	GGC Glycine	C		
	GUA Valine	GCA Alanine	GAA Glutamic Acid	GGA Glycine	A		
	GUG Valine	GCG Alanine	GAG Glutamic Acid	GGG Glycine	G		
		Nonpolar, aliphatic	Polar, uncharged	Aromatic	Positively charged	Negatively charged	

Gambar 6. Tabel kodon dan asam amino yang dihasilkan, diambil dari <https://id.wikipedia.org/wiki/Kodon>.

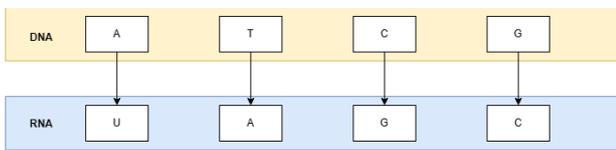
Pembuatan asam amino akan dimulai saat t-RNA membaca sebuah kodon dengan kode AUG, yang berperan menjadi *start* kodon. Kodon adalah bentuk triplet (tiga) basa yang menjadi *template* untuk membentuk asam amino. Setelah kodon AUG selesai dibaca, kodon selanjutnya akan diubah menjadi asam amino, hingga mencapai *stop* kodon dengan kode UAG, UAA, dan UGG. Saat membaca 3 kodon tersebut, proses translasi akan berhenti.

III. IMPLEMENTASI DAN PEMBAHASAN

A. Transkripsi DNA

Proses dimulai dengan mengubah sekuens DNA menjadi cetakan m-RNA. Proses pengubahan yang terjadi akan memiliki kompleksitas waktu linear $O(n)$, hal ini karena proses transkripsi melibatkan pemrosesan secara sekuensial dari karakter nukleotida paling pertama hingga paling akhir secara linear.

Simplifikasi transkripsi dapat direpresentasikan dalam graf berarah dengan satu akar akan memiliki satu anak saja. Hal ini dikarenakan satu nukleotida DNA hanya akan digantikan spesifik oleh satu basa RNA saja. Representasi graf transkripsi DNA ke m-RNA dapat dilihat pada gambar berikut.



Gambar 7. Graf transkripsi DNA ke m-RNA.

Algoritma pengubahan sekuens DNA menjadi sekuens m-RNA dapat dilakukan menggunakan Python. Sekuens DNA yang diterima akan direpresentasikan dalam bentuk *string*. Setiap nukleotida pada sekuens DNA (dilambangkan sebagai *char* di Python) akan diolah menggunakan *for loop* untuk diubah menjadi karakter nukleotida m-RNA.

```

1 import sys
2
3 # Input masukan sekuens DNA dalam bentuk string
4 DNA = input("Masukkan sekuens DNA: ")
5
6 # Pengecekan apakah panjang DNA merupakan kelipatan 3
7 if len(DNA) % 3 != 0:
8     print("Panjang sekuens DNA harus merupakan kelipatan 3!")
9     sys.exit()
10
11 RNA = ""
12 # Algoritma transkripsi, mengubah basa DNA menjadi basa m-RNA
13 for nucleotids in DNA:
14     if nucleotids == 'A':
15         RNA += 'U'
16     elif nucleotids == 'T':
17         RNA += 'A'
18     elif nucleotids == 'G':
19         RNA += 'C'
20     elif nucleotids == 'C':
21         RNA += 'G'
22 # Jika terdapat input yang tidak mengandung basa A, T, G, ataupun C
23 else:
24     print("Nukleotida DNA hanya A, T, G, dan C")
25     sys.exit()
26
27 # Output sekuens m-RNA
28 print("Sekuens m-RNA:", RNA)

```

Gambar 8. Kode Python untuk mengubah basa DNA menjadi basa m-RNA.

B. Pembentukan Kodon

Kodon merupakan kumpulan tiga nukleotida m-RNA. Kodon merepresentasikan asam amino yang akan dibentuk berdasarkan basa-basa yang terdapat di dalam kodon tersebut. Hal ini berarti kodon akan tersusun dari kombinasi tiga basa A, U, G, dan/atau C. Jumlah kombinasi yang mungkin untuk membentuk suatu nukleotida pada kodon adalah,

$$C_3^4 = \frac{4!}{3!(4-3)!} = \frac{4!}{3!} = 4$$

Untuk membentuk suatu kodon, diperlukan tiga basa berurutan. Didapatkan bahwa kombinasi banyak kodon yang dapat terbentuk untuk sebuah kodon *XXX* adalah sebanyak,

$$C_3^4 \times C_3^4 \times C_3^4 = 4 \times 4 \times 4 = 64 \text{ kemungkinan.}$$

Dengan menggunakan hubungan tersebut, jika dalam suatu kodon, hanya diketahui n buah basa saja, maka didapatkan rumus banyak kemungkinan kodon (l) yang mungkin adalah,

$$l = \frac{64}{4^n}, 0 \leq n \leq 3 \quad (6)$$

Untuk sekuens m-RNA sepanjang m , akan didapatkan banyak kemungkinan kodon yang terbentuk adalah,

$$\text{Banyak kodon} = k = \frac{m}{3} \quad (7)$$

$$\text{Kombinasi} = 64^k = 64^{\frac{m}{3}} \quad (8)$$

Algoritma untuk membentuk kodon dapat dibuat dengan memisahkan sekuens basa m-RNA untuk setiap 3 basa. Kumpulan kodon akan disimpan dalam sebuah *array*. Setiap elemen di *array* adalah kodon yang sudah dibentuk.

```

1 # Memisahkan kodon setiap 3 basa ditemukan
2 for i in range(0, len(RNA), 3):
3     codon = ""
4     for j in range(i, i+3):
5         codon += RNA[j]
6     codons.append(codon)
7
8 # Output
9 print(f"Kodon berjumlah {len(codons)} berisi: {codons}")

```

PS D:\IF3\Mat Diskrit> py .\sequence.py
Masukkan sekuens DNA: AGCTAGGAGCCAGTATCAG
Kodon berjumlah 7 berisi: ['UCG', 'AUC', 'CUC', 'CGG', 'UCA', 'UAG', 'UGC']

Gambar 9. Kode Python untuk membentuk suatu kodon.

Graf kodon adalah graf yang setiap daunnya merepresentasikan kodon yang terbentuk, yaitu sebanyak 64 buah [8]. Representasi graf untuk seluruh kemungkinan kodon akan membentuk suatu graf yang sangat kompleks. Representasi kodon dalam bentuk graf dapat dilihat pada gambar (10).

C. Pembentukan Asam Amino

Asam amino yang lazim ditemukan di tubuh manusia adalah sebanyak 20 jenis. Jenis asam amino tersebut antara lain,

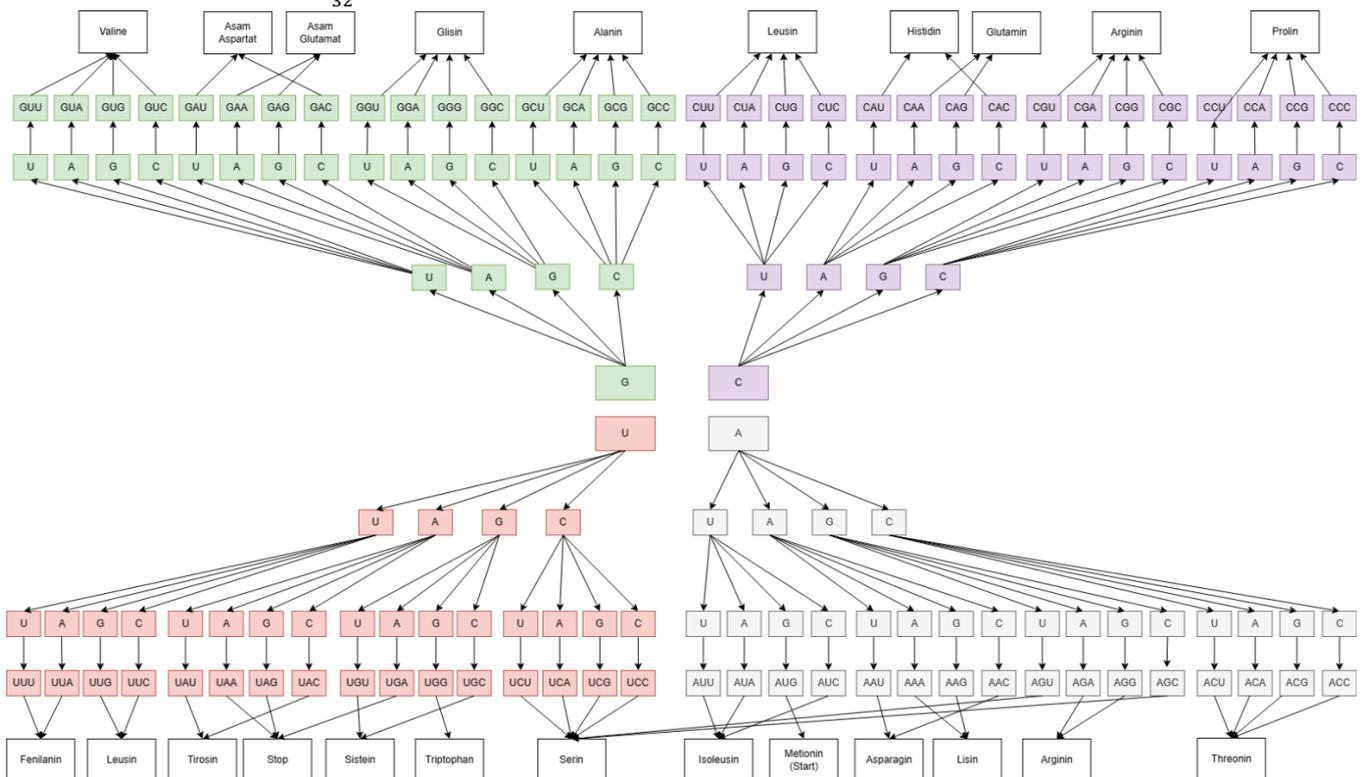
1. Metionin, hanya dapat dibentuk oleh kodon AUG. Berarti hanya ada 1 kombinasi kodon yang dapat membentuk metionin, sehingga peluang terbentuk Metionin adalah $\frac{1}{64}$.
2. Fenilalanin, dibentuk oleh kodon UUU dan UUC, atau 2

- kombinasi kodon dengan peluang $\frac{2}{64} = \frac{1}{32}$ kemungkinan.
- Leusin, dapat dibentuk oleh kodon UUA, UUG, CUU, CUC, CUA, dan CUG. Terdapat 6 kombinasi kodon yang mungkin atau peluangnya $\frac{6}{64} = \frac{3}{32}$.
 - Isoleusin, dibentuk oleh kodon AUU, AUA dan AUC, atau memiliki 3 kombinasi kodon atau dengan peluang $\frac{3}{64}$.
 - Valine, dapat dibentuk oleh GUU, GUC, GUA, GUG, atau 4 kombinasi kodon yang mungkin atau dengan peluang $\frac{4}{64} = \frac{1}{16}$.
 - Serin, dapat dibentuk oleh UCU, UCC, UCA, UCG, AGU, atau AGC, yang berarti terdapat 6 kombinasi kodon atau dengan peluang $\frac{6}{64} = \frac{3}{32}$.
 - Prolin, dapat dibentuk oleh CCU, CCC, CCA, dan CCG, atau 4 kemungkinan kombinasi atau dengan peluang $\frac{4}{64} = \frac{1}{16}$.
 - Threonin, dapat dibentuk oleh ACU, ACC, ACA, ACG, atau 4 kombinasi kodon atau dengan peluang $\frac{4}{64} = \frac{1}{16}$.
 - Alanin, dapat dibentuk oleh kodon GCU, GCC, GCA, atau GCG, berarti ada 4 kombinasi kodon yang memungkinkan atau dengan peluang $\frac{4}{64} = \frac{1}{16}$.
 - Tirosin, dapat dibentuk oleh UAU dan UAC, hanya dua kombinasi kodon atau peluang $\frac{2}{64} = \frac{1}{32}$.
 - Histidin, dibentuk oleh CAU dan CAC, berarti hanya 2 kombinasi kodon saja atau dengan peluang $\frac{2}{64} = \frac{1}{32}$.
 - Glutamin, dapat dibentuk oleh CAA dan CAG saja, atau 2 kombinasi kodon (peluang $\frac{2}{64} = \frac{1}{32}$).

- Asparagin, hanya dapat dibentuk oleh 2 jenis kodon saja, yaitu AAU dan AAC (peluang $\frac{2}{64} = \frac{1}{32}$).
- Lisin, dapat dibentuk oleh AAA dan AAG, atau dua jenis kombinasi kodon saja (peluang $\frac{2}{64} = \frac{1}{32}$).
- Asam aspartat, dapat dibentuk oleh GAU dan GAC, atau 2 kombinasi kodon (peluang $\frac{2}{64} = \frac{1}{32}$).
- Asam glutamat, dapat dibentuk oleh GAA dan GAG, atau 2 kombinasi kodon (peluang $\frac{2}{64} = \frac{1}{32}$).
- Sistein, memiliki 2 kombinasi kodon yaitu UGU dan UGC atau hanya berpeluang $\frac{2}{64} = \frac{1}{32}$.
- Tryptophan hanya dapat dibentuk oleh 1 jenis kombinasi kodon saja, yaitu kodon UGG, hanya $\frac{1}{64}$.
- Arginin memiliki 6 kombinasi kodon, yaitu CGU, CGC, CGA, CGG, AGA, dan AGG, dengan peluang $\frac{6}{64} = \frac{3}{32}$.
- Glisin dapat dibentuk oleh 4 kombinasi kodon, yaitu GGU, GGC, GGA, dan GGG atau dengan peluang $\frac{4}{64} = \frac{1}{16}$.

Berdasarkan data di atas, hanya terdapat 61 kodon yang memungkinkan untuk membentuk asam amino. Dengan kata lain, peluang kodon dapat membentuk asam amino adalah $\frac{61}{64}$.

Dalam bentuk representasi graf berarah, graf untuk membentuk kodon saja akan berbentuk pohon dengan jumlah 64 daun. Namun saat graf tersebut digabungkan dengan graf pembentuk asam amino, graf tidak lagi berbentuk pohon. Graf gabungan yang menunjukkan representasi proses pembentukan kodon dan translasi basa m-RNA menjadi asam amino dapat dilihat pada gambar (10).



Gambar 10. Representasi proses pembentukan kodon dan translasi m-RNA ke asam amino dalam bentuk graf berarah.

Algoritma dalam bahasa Python juga dapat dibuat untuk memudahkan langkah *coding* kodon m-RNA menjadi asam amino. Dalam kode, digunakan tipe data *dictionary* Python yang memiliki *keys* berupa nama asam amino yang akan dibentuk, dan *value* berupa *array* yang berisi kodon apa saja yang dapat membentuk asam amino tersebut.

```

1 aminoAcid = {
2     "Metionin": ['AUG'],
3     "Leusin": ['UUA', 'UUG', 'CUU', 'CUC', 'CUA', 'CUG'],
4     "Fenilalanin": ['UUU', 'UUC'],
5     "Serin": ['UCU', 'UCC', 'UCA', 'UCG'],
6     "Tirosin": ['UAU', 'UAC'],
7     "Sistein": ['UGU', 'UGC'],
8     "Treonin": ['ACU', 'ACC', 'ACA', 'AGG'],
9     "Asparagin": ['AAU', 'AAC'],
10    "Lisin": ['AAA', 'AAG'],
11    "Asam Glutamat": ['GAU', 'GAC'],
12    "Asam Aspartat": ['GAA', 'GAG'],
13    "Prolin": ['CCU', 'CCC', 'CCA', 'CCG'],
14    "Histidin": ['CAU', 'CAC'],
15    "Glisin": ['GGU', 'GGC', 'GGA', 'GGG'],
16    "Arginin": ['CGU', 'CGC', 'CGA', 'CGG', 'AGA', 'AGG'],
17    "Alanin": ['GCU', 'GCC', 'GCA', 'GCG'],
18    "Valin": ['GUU', 'GUC', 'GUA', 'GUG'],
19    "Glutamin": ['CAA', 'CAG'],
20    "Isoleusin": ['AUU', 'AUC', 'AUA'],
21    "Tryptophan": ['UGG']
22 }
23
24 # Translasi sekuens m-RNA menjadi sekuens asam amino
25 protein = ""
26 start = False
27 # Array codons berisi kodon-kodon yang sudah dipisahkan, terdefinisi
28 for codon in codons:
29     if codon == 'AUG':
30         if not start:
31             start = True
32             continue
33     elif codon in ['UAA', 'UAG', 'UGA']:
34         start = False
35     if start:
36         for key, value in aminoAcid.items():
37             if codon in value:
38                 if start:
39                     protein += key + " "
40
41 print("Protein yang dihasilkan: ", protein)

```

Gambar 11. Algoritma dalam bahasa Python untuk mengubah kodon m-RNA menjadi asam amino

Data di atas menunjukkan kombinasi yang mungkin jika diketahui asam aminonya saja. Kombinasi asam amino yang mungkin jika diketahui salah satu kodonnya pun akan hanya memiliki satu asam amino yang dapat dibentuknya saja. Namun pada pemrosesan DNA, dapat terjadi kemungkinan mutasi yang dapat merubah susunan sekuens DNA sehingga dapat mengubah m-RNA dan asam amino yang dihasilkan. Mutasi dapat menghapus, menambah atau mengubah susunan basa pada DNA atau RNA. Oleh karenanya penting untuk mengetahui kemungkinan yang ada jika diketahui hanya beberapa basanya saja.

Kita sudah mengetahui jika suatu kodon diketahui seluruh basanya secara lengkap, maka hanya akan ada 1 kemungkinan asam amino yang dapat dibentuk. Misalkan suatu kodon hanya diketahui satu basa yang membentuknya saja ($X_ _$), maka kemungkinan kodon yang dapat dibentuk adalah sebanyak $4 \times 4 = 16$ kemungkinan. Jika dimisalkan m adalah banyak asam amino yang mungkin dibentuk jika hanya diketahui basa sejumlah n , dan P adalah peluang diketahuinya suatu jenis asam amino yang akan dibentuk, dapat dirumuskan:

$$P = \frac{1}{m_n} \quad (9),$$

1. *AXX*. Jika hanya diketahui satu basa saja yaitu A, maka kemungkinan asam amino yang didapatkan adalah 7 kemungkinan, yaitu Isoleusin, Metionin (*start* kodon), Threonin, Asparagin, Lisin, Serin, dan Arginin. Peluang diketahuinya asam amino yang dibentuk adalah

$$P = \frac{1}{7}$$

- a. Jika diketahui basa kedua adalah A, maka kemungkinan asam amino adalah 2, yaitu Asparagin dengan 2 kemungkinan kodon dan Lisin dengan 2 kemungkinan kodon. Dengan demikian, peluang diketahuinya asam amino yang dapat dibentuk adalah

$$P = \frac{1}{2}$$

- b. Jika diketahui basa kedua adalah U, maka asam amino yang mungkin adalah sebanyak 2, yaitu Metionin dengan kodon yang mungkin hanyalah satu, Metionin, dan asam amino Isoleusin dengan 3 kemungkinan kodon. Dengan demikian peluang diketahuinya asam amino adalah

$$P = \frac{1}{2}$$

- c. Jika diketahui basa kedua adalah G, maka asam amino yang mungkin adalah Serin dengan 2 kemungkinan kodon dan Arginin dengan 2 kemungkinan kodon. Peluang diketahuinya asam amino yang dapat dibentuk adalah

$$P = \frac{1}{2}$$

- d. Jika diketahui basa kedua adalah C, hanya ada 1 asam amino yang mungkin, yaitu Threonin. Peluang diketahuinya asam amino yang terbentuk adalah

$$P = \frac{4}{4} = 1.$$

Artinya apapun basa X nantinya, asam amino yang terbentuk akan selalu Threonin.

2. *CXX*. Jika basa pertama yang diketahui merupakan C, maka terdapat 5 asam amino yang mungkin dibentuk, yaitu Leusin, Prolin, Histidin, Glutamin, dan Arginin. Berarti peluang ditemukannya asam amino adalah,

$$P = \frac{1}{5}$$

- a. Jika diketahui basa kedua adalah A, maka asam amino yang mungkin dibentuk adalah Histidin dan Glutamin. Peluang ditemukannya asam amino yang terbentuk adalah

$$P = \frac{1}{2}$$

- b. Jika diketahui basa kedua adalah U, hanya Leusin asam amino yang dapat terbentuk. Dengan demikian peluang ditemukannya asam amino adalah

$$P = \frac{4}{4} = 1.$$

- c. Basa kedua adalah G, hanya Arginin asam amino yang dapat terbentuk. Peluang ditemukannya asam

amino adalah

$$P = \frac{4}{4} = 1.$$

- d. Basa kedua adalah C, hanya Prolin asam amino yang dapat dibentuk, sehingga peluang ditemukannya jenis asam amino yang terbentuk adalah

$$P = \frac{4}{4} = 1.$$

3. GXX. Jika basa pertama yang diketahui adalah G, maka kemungkinan asam amino yang dapat dibentuk adalah Valine, Alanin, Asam Aspartat, Asam Glutamat, dan Glisin. Kemungkinan ditemukannya suatu jenis asam amino yang dapat dibentuk adalah

$$P = \frac{1}{6}.$$

- a. Jika diketahui basa kedua adalah U, hanya satu asam amino yang dapat dibentuk, yaitu Valine. Apapun X nantinya, Valine akan menjadi asam amino yang terbentuk. Sehingga peluang terbentuknya jenis asam amino adalah,

$$P = \frac{4}{4} = 1.$$

- b. Jika diketahui basa kedua adalah C, asam amino yang dapat dibentuk hanyalah Alanin, sehingga peluang terbentuknya suatu jenis asam amino adalah

$$P = \frac{4}{4} = 1.$$

- c. Jika diketahui basa kedua adalah A, maka asam amino yang dapat terbentuk adalah Asam Aspartat dan Asam Glutamat. Jadi peluang terbentuknya suatu jenis asam amino adalah

- d. Jika diketahui basa kedua adalah G, hanya satu asam amino saja yang terbentuk, yaitu Glisin, sehingga peluang terbentuk satu-satunya asam amino adalah

$$P = \frac{4}{4} = 1.$$

4. UXX. Jika hanya diketahui satu basa saja yaitu U, maka kemungkinan asam amino yang didapatkan adalah Fenilalanin, Leusin, Serin, Tirosin, Sistein, dan Trytophan. Pada basa dengan awalan U, terdapat 3 kodon khusus yaitu UAA, UAG, dan UGA yang merupakan *stop* kodon dan tidak dapat membentuk asam amino. Hal ini menandakan banyak asam amino dan/atau stop kodon yang dapat dibentuk adalah,

$$P = \frac{1}{7}.$$

- a. Jika diketahui basa kedua adalah U, maka asam amino yang mungkin dibentuk adalah Fenilalanin dan Leusin. Maka peluang terbentuknya suatu jenis asam amino adalah

$$P = \frac{1}{2}.$$

- b. Jika diketahui basa kedua adalah A, maka banyak asam amino yang didapatkan adalah 1 saja, yaitu Tirosin, yaitu untuk UAU dan UAC. Jika kodon yang didapat adalah UAA dan UAG, maka tidak akan ada asam amino yang terbentuk. Oleh

karenanya diasumsikan peluang asam amino yang terbentuk yaitu

$$P = \frac{1}{2+1} = \frac{1}{3},$$

dengan satu tambahan merupakan stop kodon.

- c. Jika basa kedua adalah C, hanya satu jenis asam amino yang dapat terbentuk, yaitu Serin, sehingga peluang satu-satunya terbentuk asam amino adalah

$$P = \frac{4}{4} = 1.$$

- d. Jika basa kedua adalah G, maka asam amino yang terbentuk adalah Sistein dan Trytophan. Namun kembali perlu diperhatikan bahwa terdapat salah satu kodon UGA yang menghasilkan *stop* kodon, sehingga seperti kasus 4.b., peluang dimodifikasi menjadi

$$P = \frac{1}{2+1} = \frac{1}{3}.$$

Representasi proses translasi dalam bentuk graf dapat dilihat melalui gambar ().

IV. KESIMPULAN

Proses sintesis protein merupakan proses pembentukan protein yang tersusun dari asam amino, dengan mengubah DNA menjadi m-RNA lalu menjadi asam amino. m-RNA tersusun atas kodon yang merupakan triplet basa dan memiliki kemungkinan tersusun dari basa A, U, C, dan G, sehingga setiap kodonnya memiliki 64 kemungkinan susunan basa. Semakin banyak susunan basa yang diketahui dalam sebuah kodon, maka semakin sedikit asam amino yang mungkin dibentuk.

Sintesis protein juga merupakan suatu proses satu arah, yang dapat direpresentasikan dengan menggunakan graf berarah. Proses transkripsi dapat direpresentasikan dengan sebuah 4 buah graf tunggal karena hanya melibatkan satu proses di setiap basanya. Proses translasi melibatkan graf berarah dalam bentuk graf bipartit berarah yang menghubungkan kodon dan juga asam amino yang dibentuknya.

V. LAMPIRAN

Lampiran A: *code* representasi proses sintesis protein dalam bahasa Python.

Lampiran B: *Link* representasi graf untuk proses sintesis protein secara keseluruhan.

VI. UCAPAN TERIMA KASIH

Rasa syukur penulis haturkan ke hadirat Allah Swt., atas ridha-Nya penulis dapat menyelesaikan makalah ini dalam waktu yang sudah ditentukan. Penulis juga mengucapkan terima kasih kepada Dr. Fariska Zakhralatifa Ruskanda, S.T., M.T., selaku dosen Matematika Diskrit K-02 tahun ajaran 2023/2024. Penulis juga mengucapkan terima kasih kepada keluarga, teman-teman, dan pihak-pihak lain atas seluruh kontribusi yang membantu penulis menyelesaikan makalah ini.

REFERENSI

- [1] D. I. Patt dan G. R. Patt, *An introduction to modern genetics*. Addison-Wesley, 1975.
- [2] N. A. Campbell dan J. B. Reece, *Biology*, 8 ed., vol. 1. Pearson, 2008.
- [3] K. H. Rosen, *Discrete Mathematics and Its Applications*, 8 ed. McGraw Hills, 2019.
- [4] R. Munir, *Matematika Diskrit*. Bandung: Informatika, 2020.
- [5] brilliant.org, "Rule of Sum and Rule of Product Problem Solving." Diakses: 11 Desember 2023. [Daring]. Tersedia pada: <https://brilliant.org/wiki/rule-of-sum-and-rule-of-product-problem-solving/>
- [6] B. Alberts dan A. Johnson, *Molecular Biology of the Cell, 4th edition*, 4 ed. New York: Garland Science, 2002.
- [7] C. Miller, "Protein Synthesis," Thompson Rivers University. Diakses: 11 Desember 2023. [Daring]. Tersedia pada: <https://humanbiology.pressbooks.tru.ca/chapter/5-6-protein-synthesis>
- [8] M. V. José dan G. S. Zamudio, "Symmetrical properties of graph representations of genetic codes: From genotype to phenotype," *Symmetry (Basel)*, vol. 10, no. 9, Sep 2018, doi: 10.3390/sym10090388.

PERNYATAAN

Dengan ini saya menyatakan bahwa makalah yang saya tulis ini adalah tulisan saya sendiri, bukan saduran, atau terjemahan dari makalah orang lain, dan bukan plagiasi.

Bandung, 11 Desember 2023



Bagas Sambega Rosyada
13522071

LAMPIRAN A

```
import sys
DNA = input("Masukkan sekuens DNA: ")
if len(DNA) % 3 != 0:
    print("Panjang sekuens DNA harus merupakan kelipatan 3!")
    sys.exit()
RNA = ""
for nucleotids in DNA:
    if nucleotids == 'A':
        RNA += 'U'
    elif nucleotids == 'T':
        RNA += 'A'
    elif nucleotids == 'G':
        RNA += 'C'
    elif nucleotids == 'C':
        RNA += 'G'
    else:
        print("Nukleotida DNA hanya A, T, G, dan C")
        sys.exit()

codons = []
for i in range(0, len(RNA), 3):
    codon = ""
    for j in range(i, i+3):
        codon += RNA[j]
    codons.append(codon)

aminoAcid = {
    "Metionin": ['AUG'],
    "Leusin": ['UUA', 'UUG', 'CUU', 'CUC', 'CUA', 'CUG'],
    "Fenilalanin": ['UUU', 'UUC'],
    "Serin": ['UCU', 'UCC', 'UCA', 'UCG'],
    "Tirosin": ['UAU', 'UAC'],
    "Sistein": ['UGU', 'UGC'],
    "Treonin": ['ACU', 'ACC', 'ACA', 'ACG'],
    "Asparagin": ['AAU', 'AAC'],
    "Lisin": ['AAA', 'AAG'],
    "Asam Glutamat": ['GAU', 'GAC'],
    "Asam Aspartat": ['GAA', 'GAG'],
    "Prolin": ['CCU', 'CCC', 'CCA', 'CCG'],
    "Histidin": ['CAU', 'CAC'],
    "Glisin": ['GGU', 'GGC', 'GGA', 'GGG'],
    "Arginin": ['CGU', 'CGC', 'CGA', 'CGG', 'AGA', 'AGG'],
    "Alanin": ['GCU', 'GCC', 'GCA', 'GCG'],
    "Valin": ['GUU', 'GUC', 'GUA', 'GUG'],
    "Glutamin": ['CAA', 'CAG'],
    "Isoleusin": ['AUU', 'AUC', 'AUA'],
    "Trypthophan": ['UGG']
}

protein = ""
start = False
for codon in codons:
    if codon == 'AUG':
        if not start:
            start = True
            continue
    elif codon in ['UAA', 'UAG', 'UGA']:
        start = False
    if start:
```

```
for key, value in aminoAcid.items():
    if codon in value:
        if start:
            protein += key + " "
```

LAMPIRAN B

<https://drive.google.com/file/d/1I15FxKcXIQP7GerRGJjsq-1VT4Dlskmk/view?usp=sharing>