

Analisis *Epitope* Sel T pada SARS-Cov2 dengan Pendekatan Bioinformatika

(Analysis of T-cell Epitopes of SARS-Cov2 with Bioinformatics Approach)

Miftahurrahma Rosyda^{1*}, Faisal Fajri Rahani²

Abstract—At the beginning of 2020, the world was shocked by the spread of global outbreaks that attacked respiratory like the SARS outbreak in 2003, namely COVID-19. The virus that causes the outbreak called SARS-Cov2. It turns out to have a similarity of ~ 87.5% with SARS-Cov. This similarity can be used to develop drugs and vaccines that are compatible with the current virus. In this case, the bioinformatics approach can be carried out as an initial stage of vaccine development. One way to develop vaccines is epitope-based vaccines. Biological data available and submitted to the public regarding T cell epitopes and protein sequences in viruses can be processed with several bioinformatics tools available online. This study compared the calculation of physicochemical characters between the SARS-Cov epitope and the SARS-Cov2 protein sequence at the same location. The characters being compared are molecular weight, point of isoelectric, aliphatic index, GRAVY, instability index, and antigenicity. Data processing is evaluated by correlation matrix. The results of the processing show that the physicochemical character between SARS-Cov and SARS-Cov2 has a strong relationship.

Intisari—Pada awal tahun 2020 dunia digemparkan dengan tersebar wabah secara global yang menyerang pernapasan seperti wabah SARS pada tahun 2003, yaitu COVID-19. Virus penyebab wabah yang bernama SARS-Cov2 ini ternyata memiliki kemiripan ~87.5% dengan SARS-Cov. Kemiripan ini dapat digunakan untuk mengembangkan obat dan vaksin yang sesuai dengan virus yang sekarang. Dalam hal ini, pendekatan secara bioinformatika dapat dilakukan sebagai tahap awal pengembangan vaksin. Salah satu cara pengembangan vaksin adalah vaksin berbasis *epitope*. Data biologi yang tersedia dan diserahkan kepada publik terkait *epitope* Sel T dan sekuen protein pada virus dapat diolah dengan beberapa *tool* bioinformatika yang tersedia secara *online*. Makalah ini membandingkan perhitungan karakter *physicochemical* antara *epitope* SARS-Cov dan sekuen protein SARS-Cov2 pada lokasi yang sama. Karakter yang dibandingkan yaitu berat molekuler, titik *isoelectric*, indeks alifatik, GRAVY, indeks kestabilan, dan antigenisitas. Pengolahan data dilakukan dengan matriks korelasi. Dari hasil pengolahan tersebut diketahui bahwa karakter *physicochemical* antara SARS-Cov dan SARS-Cov2 memiliki hubungan yang kuat.

Kata Kunci— Bioinformatika, SARS-Cov2, COVID-19, *Epitope*, Immunoinformatika, *Biomolecular*, *Bio-datamining*, *Coronavirus*.

^{1*,2} Teknik Informatika Fakultas Teknologi Industri Universitas Ahmad Dahlan, Jl Ring Road Selatan, Tamanan Banguntapan, Bantul, Yogyakarta, 55166 (telp: 0274-563515; fax: 0274-564604; e-mail: miftahurrahma.rosyda@tif.uad.ac.id, faisal.fajri@tif.uad.ac.id)

I. PENDAHULUAN

Dalam dua dekade terakhir ini, virus corona telah menyebabkan dua pandemi berskala besar, yaitu *Severe Acute Respiratory Syndrome* (SARS) dan *Middle East Respiratory Syndrome* (MERS). Kedua virus ini berasal dari kelelawar dan dapat menginfeksi manusia [1].

Di akhir tahun 2019 muncul varian virus corona baru yang berawal dari Wuhan, China. Persebaran virus ini sangat cepat hingga WHO mengumumkan darurat kesehatan global pada tanggal 30 Januari 2020 [2]. Virus ini dikenal dengan sebutan SARS-Cov2 dan penyakitnya disebut *Coronaviruses Disease-2019* (Covid-19). Kini lebih dari enam juta kasus dan 400.000 kematian telah dilaporkan dari seluruh dunia [3].

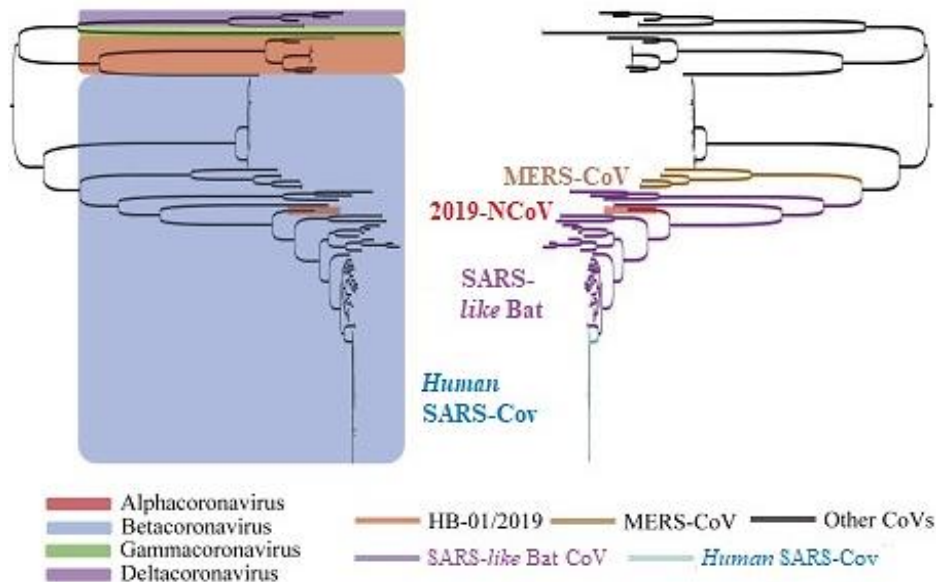
Baik SARS-Cov, MERS, maupun SARS-Cov2, ketiganya masuk dalam subfamili *beta-coronavirus*. Secara umum, virus ini terdiri atas rantai RNA tunggal sebagai material nukleotida dan diselubungi oleh lemak [4]. Virus ini berukuran ~30 kb dan terbentuk dari beberapa struktur protein besar, yaitu protein *surface* (S), protein *envelope* (E), protein *membrane* (M), dan protein *nucleocapsid* (N) [5], [6].

Gbr. 1 menunjukkan pohon filogenetik yang menjelaskan hubungan 2019-nCov (SARS-Cov2) dengan virus corona lain dalam satu famili [7], sedangkan pada Tabel I diperlihatkan perbandingan kemiripan struktur protein pada SARS-Cov2 dengan struktur protein pada SARS-Cov, Bat-Cov, dan MERS-Cov [8]. Ternyata SARS-Cov2 memiliki kemiripan > 90% dengan Bat-Cov, tetapi yang menyebabkan pandemi global adalah SARS-Cov, dengan tingkat kemiripan ~87.5%, dan MERS-Cov dengan tingkat kemiripan ~40%.

Hingga saat ini belum ditemukan vaksin yang efektif untuk mencegah dan memerangi SARS-Cov2, padahal vaksin sangat penting untuk mengurangi penyebaran virus dan memperkecil angka kematian. Salah satu cara untuk mendesain vaksin adalah dengan mempelajari *epitope*-nya. *Epitope* merupakan bagian dari protein yang dapat memicu aktifnya sistem kekebalan tubuh dan juga merupakan bagian yang dapat berikatan dengan antibodi [9].

Makalah ini bertujuan membantu proses awal desain vaksin berbasis *epitope* dengan menganalisis karakter *physicochemical*-nya. Ada banyak karakter *physicochemical* yang dimiliki oleh protein. Dengan menganalisis karakter yang paling berpengaruh dalam pembentukan *epitope*, dapat dilakukan reduksi dimensi data untuk masuk dalam proses prediksi selanjutnya.

Epitope terdiri atas *epitope* sel T dan *Epitope* sel B. Sel B bertugas khusus membentuk antibodi yang sesuai, sedangkan sel T bertugas membunuh sel yang menghasilkan *epitope*.



Gbr. 1 Perbandingan 2019-nCov (SARS-Cov2) dengan virus corona lain.

TABEL I
TINGKAT KEMIRIPAN STRUKTUR PROTEIN SARS-COV2 DENGAN
BETACORONAVIRUS LAINNYA

Virus	Struktur Protein			
	S	M	E	N
Bat-Cov	80%	98%	100%	94%
SARS-Cov	76%	90%	94%	90%
MERS-Cov	35%	42%	36%	48%

Berdasarkan penelitian pada SARS-Cov, diketahui bahwa sel T menyerang secara dominan pada protein S dan N [6]. Oleh karena itu, makalah ini fokus untuk menganalisis *epitope* sel T pada protein S dan N.

Bioinformatika memadukan ilmu biologi molekuler, matematika, dan informatika [10]. Pendekatan dengan ilmu bioinformatika berperan untuk mengurangi waktu penelitian pada laboratorium basah biologi. Salah satu topik yang berperan dalam bidang bioinformatika adalah analisis sekuen [11]. Analisis sekuen SARS-Cov2 sangat penting mengingat virus corona dari subfamili *betacoronavirus* pernah menyebabkan wabah global dengan tingkat kematian yang cukup tinggi. Makalah ini mengkaji hubungan antara karakter *epitope* SARS-Cov dengan SARS-Cov2 dengan memanfaatkan *tool* bioinformatika yang dapat diakses publik di internet.

II. DATASET

Prediksi sekuen protein yang menjadi *epitope* sel T SARS-Cov2 diperoleh dari *Immune Epitope Database* (IEDB). IEDB merupakan *website* yang menyediakan *tool* untuk memprediksi *epitope* dengan metode komputasi [12], [13]. Dari *tool* ini dapat diprediksi *epitope* SARS-Cov2 dari *epitope* SARS-Cov berdasarkan kemiripannya.

Informasi yang terkumpul terkait SARS-Cov2 masih sangat sedikit dan data *epitope* juga belum tersedia. Meskipun demikian, informasi umum terkait virus corona seperti SARS-Cov dan MERS-Cov dapat diperoleh [14].

TABEL II
PERBANDINGAN JUMLAH *EPITOPE* SEL T PADA *CORONAVIRUS*

	SARS-Cov	MERS-Cov	Lainnya	Total
<i>Epitope</i> Sel T	164	25	54	243

TABEL III
JUMLAH *EPITOPE* SEL T PADA PROTEIN SARS-COV

Protein SARS-Cov	Jumlah <i>Epitope</i> Sel T
<i>Surface/Spike Glycoprotein</i>	48
<i>NucleoProtein</i>	33
<i>Replicase Polyprotein 1ab</i>	9
Protein 3a	7
<i>Membrane Protein</i>	4

Pada Tabel II dapat dilihat jumlah *epitope* sel T yang tersedia pada situs IEDB. Selain memiliki tingkat kemiripan yang tinggi dengan SARS-Cov2, ternyata SARS-Cov juga memiliki data *epitope* yang terbanyak. Hal ini sangat mendukung untuk memprediksi *epitope* SARS-Cov2 dengan pemetaan sekuen pada lokasi yang sama. Letak *epitope* pada masing-masing protein SARS-Cov diperlihatkan pada Tabel III [8]. Jumlah *epitope* pada Tabel I dan Tabel II berbeda karena pada Tabel II hanya diambil *epitope* SARS-Cov yang berenang pada manusia.

III. METODOLOGI

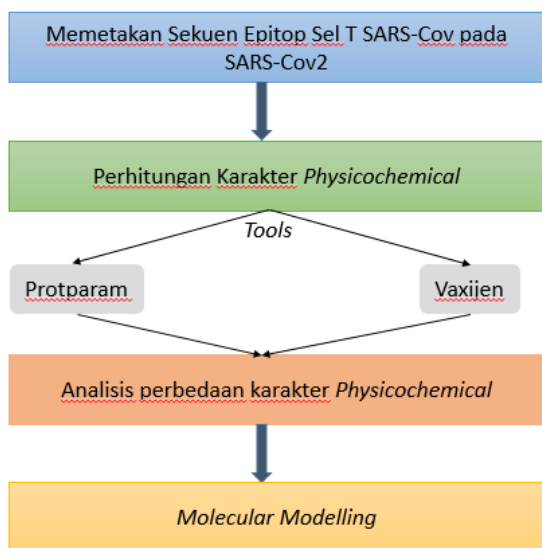
Metode yang dilakukan pada makalah ini ditunjukkan pada Gbr. 2. Setelah mendapatkan nilai karakteristik, masing-masing sekuen dibandingkan, kemudian diambil beberapa sekuen yang memiliki nilai antigenisitas tertinggi, lalu dilakukan *molecular modeling* menggunakan aplikasi PyMOL.

A. Pemetaan Sekuen *Epitope* SARS-Cov pada SARS-Cov2

Metode pemetaan ini dilakukan dengan membuat daftar *epitope* SARS-Cov dari IEDB yang paling dominan, kemudian menyimpan lokasi sekuen asam amino tersebut. Lalu, akan dibandingkan sekuen asam amino pada protein SARS-Cov2

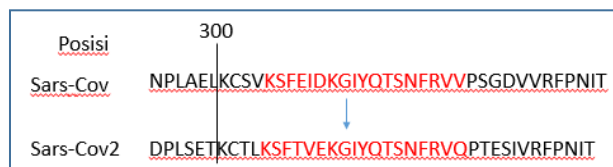
TABEL IV
PEMETAAN SEKUEN SARS-COV PADA SARS-COV2

Kode Sekuen	SARS-COV	SARS-Cov2	Protein	Pemetaan Lokasi	Kemiripan (%)
S1	KSFEIDKGIYQTSNFRVV	KSFTVEKGIYQTSNFRVQ	S	304-321	78
S2	STFFSTFKCYGVSATKL	SASFSTFKCYGVSPTKL	S	371-387	82
S3	KLPDDFMGCV	KLPDDFTGCV	S	424-433	90
S4	KSIVAYTMSLGADESSIAAY	QSIAYTMSLGAENSVAY	S	690-707	72
S5	SIVAYTMSL	SIIAYTMSL	S	691-699	89
S6	TECANLLLQYGSFCTQL	TECSNLLLQYGSFCTQL	S	747-763	94
S7	VKQMYKTPTLKYFGGGFNF	VKQIYKTPPIKDFGGFNF	S	785-802	78
S8	ESLTTTSTALGKLQDVV	DSLSTASALGKLQDVV	S	936-952	71
S9	ALNTLVKQL	ALNTLVKQL	S	958-966	100
S10	VLNDILSRL	VLNDILSR	S	976-984	100
S11	LITGRLQSL	LITGRLQSL	S	996-1004	100
S12	QLIRAAEIRASANLAATK	QLIRAAEIRASANLAATK	S	1011-1028	100
S13	SWFITQRNFFSPQII	HWFVTQRNFYEPQI	S	1101-1115	73
S14	RLNEVAKNL	RLNEVAKNL	S	1185-1193	100
S15	NLNESLIDL	NLNESLIDL	S	1192-1200	100
S16	FIAGLIAIV	FIAGLIAIV	S	1220-1228	100
M1	GLMWLSYFV	GLMWLSYFI	M	61-70	89
M2	TLACFVLAHV	TLACFVLAHV	M	148-156	100
M3	HLRMAGHSL	HLRIAGHHL	M	89-97	78
N1	ALNTPKDHI	ALNTPKDHI	N	138-146	100
N2	LQLPQGTTL	LQLPQGTTL	N	159-167	100
N3	GETALALLL	GDAALALLL	N	215-224	80
N4	LALLLDRL	LALLLDRL	N	219-227	100
N5	LLDRLNQL	LLDRLNQL	N	222-230	100
N6	RLNQLESKV	RLNQLESKM	N	226-234	89
N7	TKQYNVTQAF	TKAYNVTQAF	N	265-274	90
N8	GMSRIGMEV	GMSRIGMEV	N	316-324	100
N9	MEVTPSGTWL	MEVTPSGTWL	N	322-331	100
N10	QFKDNVILL	NFKDQVILL	N	345-353	78



Gbr. 2 Metode analisis epitope sel T.

pada lokasi yang sama. Metode ini pernah dilakukan dan berhasil memetakan 48 sekuen epitope SARS-Cov ke SARS-Cov2 [8]. Pada makalah ini dipilih 29 epitope yang memiliki tingkat kemiripan lebih dari 70%. Hasil pemetaan sekuen epitope SARS-Cov pada SARS-Cov2 disajikan pada Tabel IV.



Gbr. 3 Pemetaan lokasi epitope pada lokasi 304-321.

Contoh pemetaan lokasi epitope sel T SARS-Cov pada SARS-Cov2 ditunjukkan pada Gbr. 3.

B. Perhitungan Karakteristik Physicochemical

Protein yang menyusun organisme (virus, bakteri, tumbuhan, hewan) terbuat dari dua puluh asam amino [15]. Setiap asam amino memiliki karakteristik physicochemical yang menjelaskan ciri fisik dan kimianya. Beberapa karakteristik tersebut yaitu berat molekuler, antigenisitas, hydrophilicity, hydrophobicity, indeks alifatik, indeks kestabilan, dan titik isoelectric. Berat molekuler berarti berat molekuler suatu protein. Antigenisitas atau reaktivitas antigenik mengacu pada kapasitas virus untuk mengikat molekul antibodi tertentu [16]. Hydrophilicity merupakan sifat molekul yang dapat berikatan dengan air, sedangkan hydrophobicity merupakan sifat molekul yang takut/tidak dapat berikatan dengan air. Indeks alifatik merupakan volume relatif yang ditempati oleh rantai samping alifatik seperti alanine,

TABEL V
HASIL PERHITUNGAN *PHYSICOCHEMICAL*

Kode	Berat Molekul (BM)		Titik <i>Isoelectric</i> (TI)		Indeks Alifatik (IA)		GRAVY (GRA)		Indeks Kestabilan (IK)		Antigenisitas (Ant)	
	SARS	SARS-2	SARS	SARS-2	SARS	SARS-2	SARS	SARS-2	SARS	SARS-2	SARS	SARS-2
S1	2.131,42	2.132,40	8,5	9,70	75,56	53,89	-0,406	-0,694	-2,36	-8,34	0,3829	0,3790
S2	1.887,18	1.823,09	9,19	9,90	45,88	45,88	0,394	0,129	5,10	23,69	0,6165	0,8820
S3	1.124,34	1.094,25	4,21	4,21	68,00	68,00	0,230	-0,030	-2,85	-11,34	-0,4911	-0,5308
S4	1.877,14	1.918,15	5,83	4,00	97,78	97,78	0,556	0,428	36,39	48,83	0,3717	0,4592
S5	984,18	998,20	5,24	5,24	130,00	141,11	1,400	1,433	38,84	48,28	0,6198	0,5234
S6	1.904,18	1.920,18	4,00	4,00	97,65	91,76	0,406	0,253	71,08	71,08	0,7436	0,6982
S7	2.226,62	2.099,46	9,83	9,53	35,79	59,44	-0,379	-0,367	33,49	21,30	-0,0029	0,1499
S8	1.762,98	1.690,87	4,37	4,21	108,82	114,71	0,141	0,276	41,26	52,59	0,7630	0,6237
S9	999,22	999,22	8,80	8,80	173,33	173,33	0,644	0,644	22,57	22,57	-0,5716	-0,5716
S10	1.042,24	929,08	5,81	5,81	205,56	182,50	0,867	0,500	51,69	56,90	-0,8524	-0,8576
S11	1.000,21	1.000,21	9,75	9,75	173,33	173,33	0,667	0,667	84,50	84,50	0,8238	0,8238
S12	1.897,21	1.897,21	10,84	10,84	120,00	120,00	0,130	0,130	20,14	20,14	0,3527	0,3527
S13	1.884,17	1.865,08	9,47	6,75	78,00	48,57	0,140	-0,850	44,65	39,27	0,6513	0,4830
S14	1.056,23	1.056,23	8,75	8,75	130,00	130,00	-0,589	-0,589	8,89	8,89	0,1820	0,1820
S15	1.030,14	1.030,14	3,67	3,67	173,33	173,33	0,122	0,122	30,29	30,29	0,6827	0,6827
S16	916,17	916,17	5,52	5,52	227,78	227,78	3,056	3,056	-0,54	-0,54	0,3162	0,3162
M1	1.115,36	1.129,38	5,52	5,52	118,89	130,00	1,456	1,489	22,60	22,60	0,2548	0,2537
M2	1.007,26	1.007,26	5,18	5,18	166,00	166,00	2,600	2,600	52,94	52,94	1,2318	1,2318
M3	1.021,21	1.053,23	9,76	9,77	97,78	141,11	-0,089	-0,067	44,00	30,29	0,4547	0,2446
N1	1.008,14	1.008,14	6,79	6,79	97,78	97,78	-0,700	-0,700	48,28	48,28	-0,8423	-0,8423
N2	970,13	970,13	5,52	5,52	130,00	130,00	0,111	0,110	78,48	78,48	0,0480	0,0480
N3	1.013,24	1.013,24	4,00	4,00	215,00	215,00	1,800	1,800	1,46	1,46	0,5482	0,5482
N4	1.039,33	1.039,33	5,84	5,84	271,11	271,11	1,844	1,844	0,51	0,51	0,5933	0,5933
N5	1.097,32	1.097,32	5,84	5,84	216,67	216,67	0,444	0,444	-7,87	-7,87	0,1566	0,1566
N6	1.086,26	1.118,32	8,75	8,75	118,89	86,67	-0,878	-1,133	12,48	58,13	0,3912	0,5632
N7	1.199,33	1.142,00	8,26	8,26	39,00	49,00	-0,830	-0,300	9,11	-7,03	0,3666	0,4932
N8	979,18	979,18	6,00	6,00	75,56	75,56	0,322	0,322	79,11	79,11	0,6287	0,6287
N9	1.120,29	1.120,29	4,00	4,00	68,00	68,00	0,130	0,130	8,59	8,59	0,6342	0,6342
N10	1.089,30	1.089,30	5,84	5,84	162,22	162,22	0,522	0,522	5,21	-11,49	0,5844	1,1677

valine, *isoluesin*, dan *leusin*. Indeks kestabilan adalah perkiraan kestabilan protein dalam suatu tabung uji [17]. Yang terakhir, titik *isoelectric* merupakan pH larutan ketika muatan bersih protein menjadi nol [18].

Perhitungan ini dilakukan menggunakan *tool* bioinformatika yang tersedia *online* di internet, yaitu ProtParam dan VaxiJen. Dalam mendesain sebuah obat dan vaksin, beberapa karakteristik *physicochemical* yang digunakan antara lain berat molekuler, jumlah residu positif dan negatif, indeks kestabilan, indeks alifatik, *Grand Average of Hydropathicity* (GRAVY), titik *isoelectric*, antigenisitas, dan *hydrophilicity* [19]. *Hidropathicity* merupakan angka relatif yang mencerminkan *hydrophilicity* dan *hydrophobicity* [20].

Perhitungan *physicochemical* terkait berat molekuler, titik *isoelectric*, indeks kestabilan, indeks alifatik, dan GRAVY dilakukan menggunakan *tool* ProtParam [17], sedangkan perhitungan antigenisitas dilakukan menggunakan *tool* VaxiJen [21]. Hasil perhitungan karakteristik *physicochemical* diperlihatkan pada Tabel V.

C. Analisis Perbedaan Karakter

Analisis karakter SARS-Cov dan SARS-Cov2 dilakukan menggunakan metode statistik dasar dan matrik korelasi. Setiap karakter dihitung rata-rata dan standar deviasinya, mengacu

pada (1) dan (2). Kemudian, seluruh karakteristik dihitung korelasinya dengan matriks korelasi berdasarkan teori Pearson mengacu pada (3). Matriks korelasi memiliki rentang antara -1 hingga 1. Nilai -1 untuk relasi yang sangat lemah, sedangkan nilai 1 untuk relasi yang sangat kuat [22].

$$\bar{x} = \frac{x_1 + x_2 + \dots + x_n}{n} \quad (1)$$

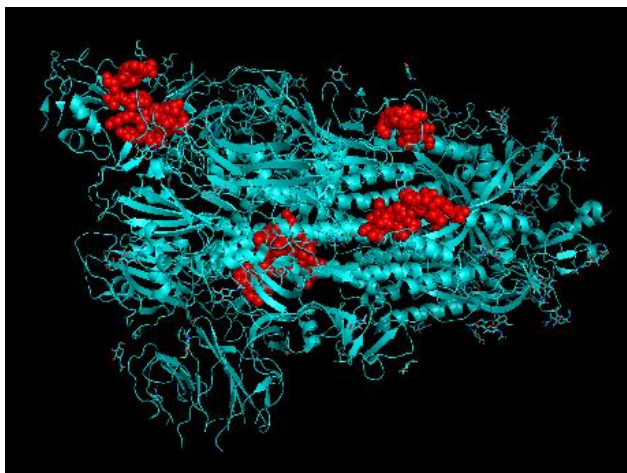
$$s = \sqrt{\frac{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}{n(n-1)}} \quad (2)$$

$$r = \frac{n(\sum xy) - (\sum x)(\sum y)}{\sqrt{[\sum x^2 - (\sum x)^2][\sum y^2 - (\sum y)^2]}} \quad (3)$$

D. Molecular Modeling

Visualisasi daerah *epitope* pada protein SARS-Cov2 dilakukan menggunakan *software* Pymol. Data protein SARS yang digunakan berasal dari *Protein Data Bank* (PDB) dengan ID 6VSB [23], [24]. Pada proses visualisasi ini hanya ditampilkan *epitope* pada protein S, karena pemodelan molekul protein M dan N untuk SARS-Cov2 dengan sekuen yang lengkap belum tersedia.

Pada Gbr. 4 diperlihatkan pemodelan molekul lokasi *epitope* dengan PDB ID=6VSB. Daerah yang berwarna merah adalah lokasi *epitope* pada S2, S6, S8, dan S11, yaitu sekuen



Gbr. 4 Pemodelan molekuler lokasi *epitope*.

‘SASFSTFKCYGVSPTKL’, ‘TECSNLLLQYGSFCTQL’, ‘DSLSTASALGKLQDVV’, dan ‘LITGRLQSL’. Empat lokasi yang dipilih merupakan lokasi dengan nilai antigenisitas tertinggi.

IV. HASIL DAN PEMBAHASAN

Pada Tabel VI dapat dilihat hasil perhitungan rata-rata untuk seluruh karakteristik pada SARS-Cov dan SARS-Cov2. Hasil rata-rata untuk indeks alifatik dan antigenisitas lebih besar pada SARS-Cov2 daripada SARS-Cov, sedangkan pada karakter lainnya, rata-rata SARS-Cov2 lebih kecil daripada SARS-Cov.

Dapat dikatakan bahwa *epitope* SARS-Cov2 memiliki rantai samping alifatik dan antigenisitas lebih besar daripada *epitope* SARS-Cov, tetapi berat molekulnya lebih ringan dan kestabilannya lebih kecil. Dari hasil ini, dapat diketahui bahwa *epitope* sel T SARS-Cov2 lebih mudah berikatan dengan antibodi, tetapi kurang stabil.

Hasil perhitungan standar deviasi disajikan pada Tabel VII. Standar deviasi berat molekul dan GRAVY pada SARS-Cov2 lebih kecil, sedangkan karakter lainnya lebih besar. Dari rata-rata dan standar deviasi, dapat dikatakan bahwa karakteristik sekuen protein pada *epitope* sel T SARS-Cov maupun SARS-Cov2 memiliki kemiripan > 90%. Padahal pemetaan sekuen dipilih dengan nilai kemiripan >70%, sehingga perubahan asam amino terjadi dengan asam amino lain yang memiliki karakter berdekatan.

Kemudian pada Tabel VIII dapat dilihat matriks korelasi antar karakter SARS-Cov dan SARS-Cov2. Warna biru menunjukkan adanya relasi yang sangat kuat, warna oranye menunjukkan relasi yang kuat, warna hijau menunjukkan relasi yang cukup kuat, warna abu-abu menunjukkan lemah, sedangkan yang tidak berwarna menunjukkan relasi yang sangat lemah, bahkan tidak ada relasi. Terlihat bahwa pada karakter yang sama terdapat relasi yang sangat kuat antara SARS-Cov dan SARS-Cov2. Lalu, relasi yang kuat terjadi antara karakter kestabilan dan GRAVY, relasi lemah ditunjukkan antara karakter antigenisitas dengan kestabilan, sedangkan karakter yang tidak memiliki relasi yang kuat dengan karakter lain adalah karakter indeks alifatik.

TABEL VI
HASIL PERHITUNGAN RATA-RATA

	Rata-Rata					
	BM	TI	IK	GRA	IK	Ant
Cov	1.292,01	6,72	28,89	128,19	0,487	0,333
Cov2	1.280,60	6,62	29,73	127,95	0,419	0,356

TABEL VII
HASIL PERHITUNGAN STANDAR DEVIASI

	Standar Deviasi					
	BM	TI	IK	GRA	IK	Ant
Cov	421,67	2,18	27,16	60,74	0,95	0,489
Cov2	410,56	2,25	29,66	60,50	0,99	0,513

TABEL VIII
MATRIKS KORELASI

	BM	TI	IK	GRA	IK	Ant
BM	0,996	0,269	-0,027	-0,568	-0,389	0,153
TI	0,336	0,955	0,001	-0,293	-0,444	-0,071
IK	-0,031	-0,059	0,919	-0,122	-0,063	0,013
GRA	-0,532	-0,265	-0,097	0,972	0,642	-0,014
IK	-0,315	-0,410	-0,054	0,665	0,970	0,229
Ant	0,087	-0,140	0,159	0,062	0,276	0,961

Pertimbangan bahwa karakter *epitope* SARS-Cov2 lebih ringan, lebih mudah berikatan dengan antibodi, dan kurang stabil dapat membantu proses pembuatan vaksin berbasis *epitope*. Karakter-karakter tersebut nantinya akan berpengaruh pada desain struktur protein yang akan direspons oleh antibodi.

V. KESIMPULAN DAN SARAN

Pemetaan *epitope* SARS-Cov pada SARS-Cov2 sebagai tahap awal pembuatan desain vaksin dapat dipertimbangkan, karena adanya relasi yang kuat antara karakter *physicochemical* SARS-Cov dan SARS-Cov2. Adanya perubahan asam amino akibat mutasi tidak berdampak signifikan pada *epitope* penelitian kali ini. Dari karakter yang telah diteliti ternyata indeks alifatik tidak memberikan perbedaan yang signifikan, sehingga karakter tersebut dapat dihilangkan pada penelitian selanjutnya. Penelitian selanjutnya dapat membandingkan lokasi *epitope* SARS-Cov dengan SARS-Cov2 yang khusus berasal dari Indonesia dengan bantuan mesin pembelajaran karena keberagaman genom yang sangat tinggi.

UCAPAN TERIMA KASIH

Ucapan terima kasih disampaikan kepada Ilmu Komputer UGM yang telah memberikan dasar ilmu Bioinformatika, kemudian kepada Magister Teknologi Informasi UGM yang telah mengajarkan metode analisis yang dalam, dan terakhir disampaikan kepada UAD yang terus menyemangati untuk terus melanjutkan penelitian.

REFERENSI

[1] P. Zhou, X.-L. Yang, X.-G. Wang, B. Hu, L. Zhang, dkk., "A Pneumonia Outbreak Associated with a New Coronavirus of Probable Bat Origin," *Nature*, Vol. 579, No. 7798, hal. 270–273, 2020.
 [2] (2020) "BBC News" [Online], <https://www.bbc.com/news/world-51318246>, tanggal akses: 9-Jun-2020.
 [3] (2020) "WHO Coronaviruses Disease Dashboard" [Online], <https://covid19.who.int/>, tanggal akses: 9-Jun-2020.

- [4] M.A. Shereen, S. Khan, A. Kazmi, N. Bashir, dan R. Siddique, "COVID-19 Infection: Origin, Transmission, and Characteristics of Human Coronaviruses," *J. Adv. Res.*, Vol. 24, hal. 91–98, 2020.
- [5] R.A. Khailany, M. Safdar, dan M. Ozaslan, "Genomic Characterization of a Novel SARS-CoV-2," *Gene Reports*, Vol. 19, hal. 1-6, 2020.
- [6] S.F. Ahmed, A.A. Quadeer, dan M.R. McKay, "Preliminary Identification of Potential Vaccine Targets for the COVID-19 Coronavirus (SARS-CoV-2) Based on SARS-CoV Immunological Studies," *Viruses*, Vol. 12, No. 3, hal. 1–15, 2020.
- [7] A. Wu, Y. Peng, B. Huang, X. Ding, X. Wang, P. Niu, J. Meng, Z. Zhu, Z. Zhang, J. Wang, J. Sheng, L. Quan, Z. Xia, W. Tan, G. Cheng, dan T. Jiang, "Genome Composition and Divergence of the Novel Coronavirus (2019-nCoV) Originating in China," *Cell Host Microbe*, Vol. 27, No. 3, hal. 325–328, 2020.
- [8] A. Grifoni, J. Sidney, Y. Zhang, R.H. Scheuermann, B. Peters, dan A. Sette, "A Sequence Homology and Bioinformatic Approach can Predict Candidate Targets for Immune Responses to SARS-CoV-2," *Cell Host Microbe*, Vol. 27, No. 4, hal. 671-680.e2, 2020.
- [9] R. Kogay dan C. Schönbach, "Epitope Predictions," *Encycl. Bioinforma. Comput. Biol.: ABC Bioinforma.*, Vol. 2, No. 1, hal. 952–971, 2018.
- [10] D.A. Aprijani dan M.A. Elfaizi, "BIOINFORMATIKA: Perkembangan, Disiplin Ilmu dan Penerapannya di Indonesia," Universitas Gunadarma, Jakarta, Indonesia, Kuliah, hal. 1-25, 2004.
- [11] I.Y. Abdurakhmonov, "Bioinformatics: Basics, Development, and Future," dalam *Bioinformatics - Updated Features and Applications*, I.Y. Abdurakhmonov, Ed., London, UK: IntechOpen, 2016.
- [12] S.K. Dhanda, S. Mahajan, S. Paul, Z. Yan, H. Kim, M.C. Jespersen, V. Jurtz, M. Andreatta, J. A. Greenbaum, P. Marcatili, A. Sette, M. Nielsen, dan B. Peters, "IEDB-AR: Immune Epitope Database - Analysis Resource in 2019," *Nucleic Acids Res.*, Vol. 47, No. W1, hal. W502–W506, 2019.
- [13] Y. Kim, A. Sette, dan B. Peters, "Applications for T-cell Epitope Queries and Tools in the Immune Epitope Database and Analysis Resource," *J. Immunol. Methods*, Vol. 374, No. 1–2, hal. 62–69, 2011.
- [14] Z. Song, Y. Xu, L. Bao, L. Zhang, P. Yu, Y. Qu, H. Zhu, W. Zhao, Y. Han, dan C. Qin, "From SARS to MERS, Thrusting Coronaviruses into the Spotlight," *Viruses*, Vol. 11, No. 1, hal. 1-28, 2019.
- [15] A.J. Cozzzone, "Proteins: Fundamental Chemical Properties," *Encycl. Life Sci.*, hal. 1–10, 2010.
- [16] M.H.V Van Regenmortel, "Antigenicity and Immunogenicity of Viral Proteins," dalam *Encyclopedia of Virology*, 3rd ed., B.W.J. Mahy dan M.H.V. Van Regenmortel, Eds., Cambridge, USA: Academic Press, hal. 137–142, 2008, hal. 137–142.
- [17] E. Gasteiger, C. Hoogland, A. Gattiker, S. Duvaud, R. Appel, dan A. Bairoch, "Protein Identification and Analysis Tools on the ExpASY Server," dalam *The Proteomics Protocols Handbook*, 1st ed., W.M. John, Ed., Totowa, USA: Humana Press, hal. 571–607, 2005.
- [18] P. Novák dan V. Havlíček, "Protein Extraction and Precipitation," dalam *Proteomic Profiling and Analytical Chemistry*, 2nd ed., P. Ciborowski dan J. Silberring, Eds., New York, USA: Elsevier, 2016, hal. 51–62.
- [19] S. Kumar, "Drug and Vaccine Design against Novel Coronavirus (2019-nCoV) Spike Protein through Computational Approach," *Preprints*, hal. 1-16, Feb. 2020.
- [20] J. Kyte dan R.F. Doolittle, "A Simple Method for Displaying the Hydropathic Character of a Protein," *J. Mol. Biol.*, Vol. 157, No. 1, hal. 105–132, 1982.
- [21] I.A. Doytchinova dan D.R. Flower, "Identifying Candidate Subunit Vaccines Using an Alignment-Independent Method Based on Principal Amino Acid Properties," *Vaccine*, Vol. 25, No. 5, hal. 856–866, 2007.
- [22] A.P. King dan R.J. Eckersley, "Descriptive Statistics II: Bivariate and Multivariate Statistics," dalam *Statistics for Biomedical Engineers and Scientists*, Cambridge, USA: Academic Press, hal. 23-56, 2019..
- [23] H.M. Berman, J. Westbrook, Z. Feng, G. Gilliland, dkk., "The Protein Data Bank," *Acta Crystallogr. Sect. D Biol. Crystallogr.*, Vol. 58, No. 6I, hal. 899–907, 2002.
- [24] D. Wrapp, N. Wang, K. Corbett, J.A. Goldsmith, C. Hsieh, O. Abiona, B. S. Graham, dan J.S. McLellan, "Cryo-EM Structure of the 2019-nCoV Spike in the Prefusion Conformation," *Science*, Vol. 367, No. 6483, hal. 1260-1263, Mar. 2020.