

Re-Scoring Skor Grid Hasil Docking Molekular Ligan Asli dari Protein Target T47D dan WiDr dengan Metode Skor GBSA-Hawkins-Zou dan Amber menggunakan Program DOCK6

Broto Santoso

Fakultas Farmasi, Universitas Muhammadiyah Surakarta

Email: Broto.Santoso@ums.ac.id

Abstrak

Keywords:

DOCK6, Amber-Score, Solvent-Score, T47D, WiDr

Prediksi energi ikatan ligan protein dapat dilakukan dengan teknik komputasi. Model skoring kekuatan ikatannya ditentukan oleh software dan metode perhitungan yang menyertainya. Selain penggunaan dataset ligan pembeding dan pengumpan, ligan terpilih yang potensial dapat diuji dengan berbagai model skoring untuk mendekatkan hasil dengan kondisi yang sesungguhnya. Re-Scoring dalam penelitian ini ditujukan untuk memperoleh karakteristik hasil ikatan ligan protein terbaik di semua model skoring yang digunakan. Protein target yang memiliki mekanisme aksi molekular terhadap sel T47D dan WiDr telah diunduh dari www.rcsb.org. Sebanyak 36 protein terpilih telah dilakukan pemisahan untuk didapatkan protein dan ligan native yang telah ditambahkan hidrogen dan muatannya (Amber-ff14SB) menggunakan bantuan program UCSF Chimera dan layanan online dari Docking Blaster (<http://blaster.docking.org>). Semua ligan native telah dihitung nilai skor Gridnya menggunakan DOCK6 terhadap semua protein target. Nilai RMSD-h ligan menjadi acuan validasi konformasi ruang hasil docking dengan konformasi pengukuran kristalografi untuk masing-masing protein target. Masing-masing skor dan interaksi ligan protein yang dipilih terbaik divisualisasi dan dianalisis menggunakan PyMOL, PLIP, dan PoseView online. Skor Amber dan Hawkins telah diperoleh untuk semua ligan native, namun hanya 47,22% protein target yang mempunyai skor Zou terhadap ligan native-nya. Hasil pengukuran afinitas ikatan ligan native secara silang menunjukkan bahwa terdapat 61,11 dan 80,56 % protein yang nilai afinitas ikatan ligan native-nya lebih rendah dibandingkan dengan ligan native dari protein lainnya. Teknik re-scoring telah terbukti memberikan penilaian lebih baik dalam melakukan evaluasi afinitas ikatan ligan protein dan konformasi ruang ligan dalam binding site pocket setiap protein.

1. PENDAHULUAN

Nilai prediksi interaksi ligan protein hasil skrining molekular dalam komputasi rancang obat dapat diperoleh melalui beberapa teknik skoring. Hasil pengukuran ini tergantung dari software dan metode

algoritma yang digunakan dan menyertai. Selain menyandingkan pengukuran dataset ligan pembeding dan pengumpan (*decoy*), ligan terpilih yang potensial dan telah dilakukan optimasi konformasi ruang tiga dimensinya atau merupakan konformasi

hasil pengukuran secara laboratorium seperti ligan *native* dapat diuji dengan metode skoring yang berbeda untuk mendapatkan hasil dengan kondisi yang sesungguhnya.

Beberapa skoring diantaranya adalah menghitung interaksi ligan dan protein dengan kehadiran pelarut, yang umumnya adalah air dan perlu dinilai besaran kontribusinya terhadap kekuatan interaksi yang terjadi dan skoring Amber yang telah terkenal dalam teknik molekular dinamik. Teknik skoring pelarut yang dapat digunakan secara gratis adalah *Molecular Mechanics Generalized Born Surface Area* (MM-GBSA) Hawkins dan Zou [1-3]. Keduanya telah tersedia dalam program DOCK6 [4-5] bersama dengan kumpulan perintah pengukuran skoring Amber [6].

Skoring ligan hasil terbaik dapat dilanjutkan dengan *solvent-score* dan *Amber-score*. Hal ini akan memperpendek waktu pengukuran dan meminimalisir persentase ketidakpastian dan perbedaan hasil prediksi dengan uji laboratorium nantinya dikarenakan adanya pelibatan pelarut dalam skoring dan inisiasi keberlanjutan pengukuran dengan teknik molekular dinamik, Amber. *Re-scoring* telah berhasil dilakukan untuk menilai hasil molekular *docking* ligan *native* yang nantinya dapat diaplikasi terhadap semua program *docking* seperti vina, PLANTS, FlexAID (NRG Suite). Ligan *native* setiap protein target hasil pengukuran dengan kristalografi x-ray harus memiliki nilai skoring tertinggi untuk ketiga metode. *Re-Scoring* dalam penelitian ini ditujukan untuk memperoleh karakteristik hasil ikatan ligan protein terbaik di semua model skoring yang digunakan.

2. METODE

Komputasi *re-scoring* telah berhasil dilakukan menggunakan workstation berbasis Ubuntu Trusty Tahr dengan spesifikasi Intel® core i5-4460, 16GB RAM, dan NVIDIA GeForce GTX 750 Ti beserta program DOCK6 versi 6.9, Chimera versi 1.13, PyMOL versi 2.1.0, PLIP versi 1.4.1. Terdapat 36 protein yang dipilih dan memiliki mekanisme aksi molekular terhadap sel T47D dan/atau WiDr (Tabel 1). Semua berkas protein telah diunduh dari www.rcsb.org dan dipersiapkan untuk mendapatkan berkas baru protein dan ligan *native* yang telah ditambahkan atom hidrogen dan muatan secara Amber-ff14SB menggunakan program UCSF Chimera dan sebagian lainnya merupakan hasil *online* dari laman preparasi protein pada *Docking Blaster* (<http://blaster.docking.org>).

Semua ligan *native* telah dihitung nilai skor Gridnya menggunakan DOCK6 terhadap semua protein target baik terhadap protein target asalnya dan secara bersilang terhadap 35 protein lainnya. Nilai RMSD-h ligan menjadi acuan validasi konformasi ruang hasil *docking* dibandingkan dengan konformasi ruang ligan hasil pengukuran kristalografi untuk setiap protein target. Masing-masing skor yang diperoleh dan interaksi ligan protein dipilih yang terbaik untuk dilakukan analisis lanjutan berdasarkan visualisasi PyMOL dan hasil prediksi interaksi menggunakan PLIP, dan PoseView *online*. Ligan *native* hasil *docking* terbaik digabungkan dengan protein target dan disimpan sebagai berkas baru untuk diukur interaksi yang terjadi antara keduanya.

Tabel 1. Protein target, nilai deskriptor ligan *native* berupa bobot molekul, muatan formal, donor dan akseptor ikatan hidrogen (HB), dan *polar surface area* (# = data diperoleh dari DOCK6, * = data diperoleh dari laman Pubchem)

Protein ID	Ligan ID	DOCK Rotatable Bonds#	Bobot Molekul#	Muatan Formal#	Pusat Massa#	Dimensi GridBox	Donor HB*	Akseptor HB*	Polar Surface Area*
S001 1x9q	FLU	3	331,304	-0,999	9,148 43,862 47,542	30,712 32,023 29,709	2	5	83,8
S002 3hb4	E2B	6	405,537	-0,002	11,678 6,832 - 11,045	36,226 43,215 36,830	3	3	83,6
S003 3hb5	E2B	6	405,537	0,000	11,101 6,470 - 11,678	33,742 40,063 36,343	3	3	83,6
S004 4xi3	29S	8	468,595	0,002	-4,773 - 7,038 33,572	29,870 30,954 31,656	2	4	57,9
S005 4zw1	4V1	2	342,442	1,000	16,468 12,619 - 19,882	23,137 26,368 24,672	2	5	93,2
S006 5aau	XBR	4	354,836	0,002	18,358 - 13,790 13,016	33,099 33,781 34,189	2	3	56,3
S007 5aav	GW5	8	353,44	-0,998	32,575 13,699 10,424	34,010 32,783 31,974	1	2	37,3
S008 5acc	KE9	6	442,481	0,001	14,149 23,452 65,533	33,830 35,153 34,653	2	6	56,3
S009 5ha9	TP0	4	276,401	0,996	-11,513 7,120 - 5,999	34,756 32,179 30,451	0	1	3,2
S010 5nqr	958	4	395,421	0,000	82,097 17,117 110,064	29,580 28,078 28,259	0	7	101
S013 5nwh	9CH	5	494,36	1,000	11,264 - 13,407 - 17,774	34,532 31,931 32,779	1	8	113

S014 5nwh	9CH	5	494,36	1,005	-2,050 - 10,542 - 24,959	29,896 28,031 26,134	1	8	113
S015 5t1z	Q97	9	360,453	-0,003	26,061 - 0,430 8,927	36,446 35,514 37,285	2	3	49,7
S016 5t92	77W	6	402,445	-0,998	-39,557 7,037 - 26,512	33,190 37,990 32,760	2	5	60,8
S017 5t97	782	7	426,535	-1,001	-40,082 6,767 - 25,145	32,355 37,395 32,859	2	4	60,8
S018 5twl	H91	4	499,439	0,001	-4,604 8,239 - 6,052	31,277 32,636 35,988	2	4	68,7
S019 5u2b	6WV	3	345,484	0,003	17,325 6,088 17,555	34,040 35,106 33,844	2	2	32,3
S020 5u2d	OBH	7	436,485	0,000	-4,876 20,624 4.873	38,681 33,928 34.326	2	6	101
S021 5w9c	OHT	10	387,522	-0,001	13,643 - 10,510 - 27,190	33,037 36,162 36,263	1	3	32,7
S022 5w9d	9XY	10	374,503	1,000	17,698 26,891 - 15,915	20,137 26,525 27,845	2	3	41,5
S023 6d0f	FYS	8	334,415	-0,002	-14,054 4,952 49,598	38,414 33,056 34,642	3	3	60,7
S024 4hki	FLN	1	222,243	0,000	7,605 - 13,339 13,211	29,509 32,801 36,034	0	2	26,3
S025 4hkk	AGI	4	270,241	0,001	-8,671 - 14,696 - 14,344	26,805 27,631 33,091	3	5	87
S026 4hkn	LU2	5	286,241	0,000	4,996 - 13,915 - 12,764	27,863 25,266 29,306	4	6	107

S027 4hl5	15W	3	268,269	-0,001	-4,599 - 12,657 - 18,051	35,717 29,623 36,419	1	4	55,8
S028 4hlf	15Z	4	270,241	0,001	-6,258 - 37,191 12,256	30,457 32,092 36,263	3	5	87
S029 4hlg	20B	2	238,243	0,002	-7,022 - 37,606 12,228	31,202 31,888 36,865	1	3	46,5
S030 4hlh	20D	1	240,234	-0,002	-38,507 - 12,961 10,661	27,777 31,060 33,724	0	3	26,3
S031 4hlk	431	1	236,27	-0,001	-40,410 - 11,493 12,191	25,204 28,575 37,176	0	2	26,3
S032 4hlm	16S	3	254,242	-0,003	-9,098 37,111 - 13,625	26,989 31,766 31,956	2	4	66,8
S033 4hmh	F94	3	254,242	0,000	6,476 - 13,592 - 13,910	30,802 24,567 31,862	2	4	66,8
S034 5ax9	4KT	5	416,48	0,000	-11,890 - 54,903 - 29,281	42,131 28,449 35,157	1	7	79,6
S035 5d7a	58C	5	375,43	0,002	264,491 - 8,758 13,080	24,068 32,842 34,258	3	6	96
S036 5ety	K56	5	432,502	-0,997	-34,231 21,394 - 20,900	31,421 32,673 23,911	1	6	79,8
S037 5fto	YMX	7	555,607	0,996	4,777 20,358 6,749	36,532 36,313 34,011	3	8	85,5
S038 5ftq	U4W	9	575,681	1,002	-20,770 9,251 - 6,690	35,831 36,071 33,914	4	8	96,5

3. HASIL DAN PEMBAHASAN

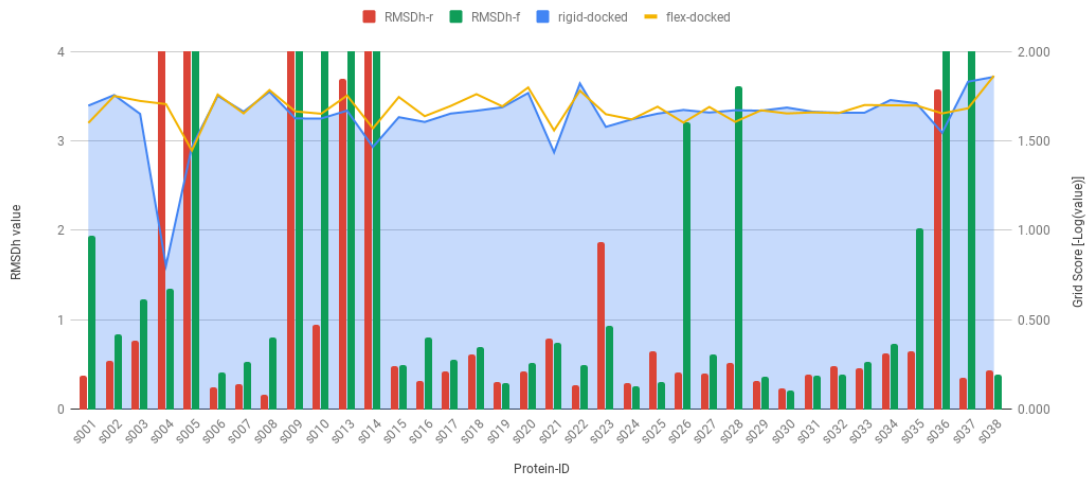
Terdapat hanya dua dari 36 ligan asli protein target yang memiliki bobot molekul (BM) lebih dari 500 kDa. Deskriptor ini digunakan untuk mengetahui secara awal seberapa besar volume *binding pocket site* ligan asli dari

suatu protein. Semua ligan asli yang diujikan telah memenuhi minimal dua nilai deskriptor dalam aturan Lipinski [7] yang digunakan untuk memprediksi kemampuan absorpsi atau permeasi ligan (Tabel 1). Nilai tiga dimensi untuk pusat massa pengukuran setiap protein telah diperoleh

secara otomatis menggunakan DOCK6 berdasarkan posisi dan konformasi ligan asli di masing-masing proteinnya.

Validasi sistem *docking* molekular memperlihatkan bahwa 3 ligan asli tidak mempunyai nilai RMSD-h kurang dari 4 (ideal) baik dengan teknik rigid atau fleksibel. Hal ini menunjukkan bahwa konformasi ligan hasil *docking* tidak mirip dengan ligan hasil kristalografi. Protein

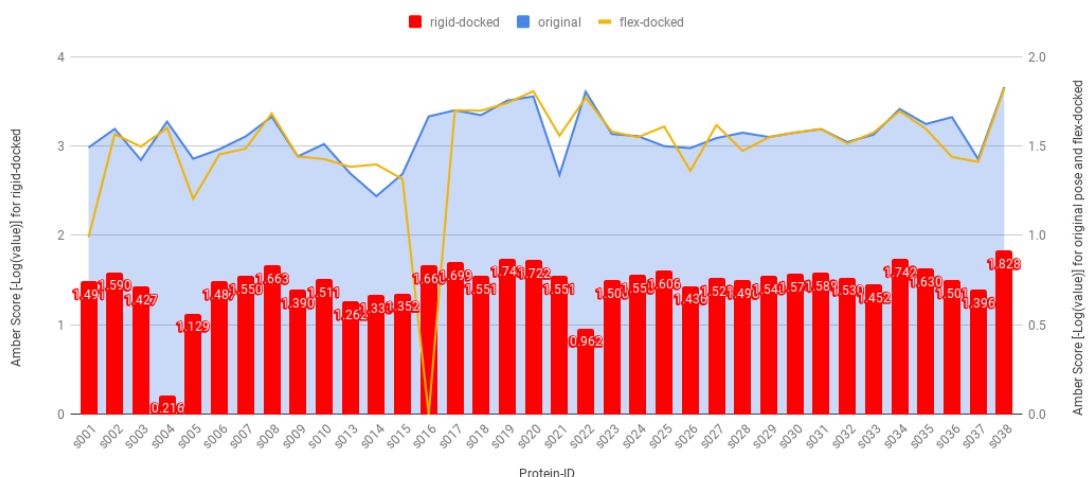
s004 memiliki RMSD-h rigid melewati nilai yang diperbolehkan dan terdapat 11% protein dengan nilai RMSD-h fleksibelnya yang memiliki karakteristik yang sama (Gambar 1). Secara keseluruhan, sistem *docking* molekular untuk protein target (lebih dari 75%) telah memenuhi nilai RMSD dengan ciri telah memiliki kemiripan konformasi dengan ligan kristalografi.



Gambar 1. Nilai Skor Grid dan RMSD-h hasil *docking* dengan ligan yang dipertahankan rigid dan fleksibel selama proses *docking* molekular.

Skor grid hasil *docking* untuk rigid dan fleksibel secara berurutan berada di rentang nilai antara -6,4-(-72,1) dan -28,0-(-73,1) kkal/mol (Gambar 1). Nilai terbaik diperoleh untuk ligan dengan BM lebih dari 500, yaitu protein s038. Hal ini

disebabkan ligan memiliki *binding site pocket* dengan volume spesifik dalam protein yang dikhususkan untuk ligan asli dan sesuai dengan volume permukaan molekularnya [8], yang berlaku hal yang sama untuk protein s037.



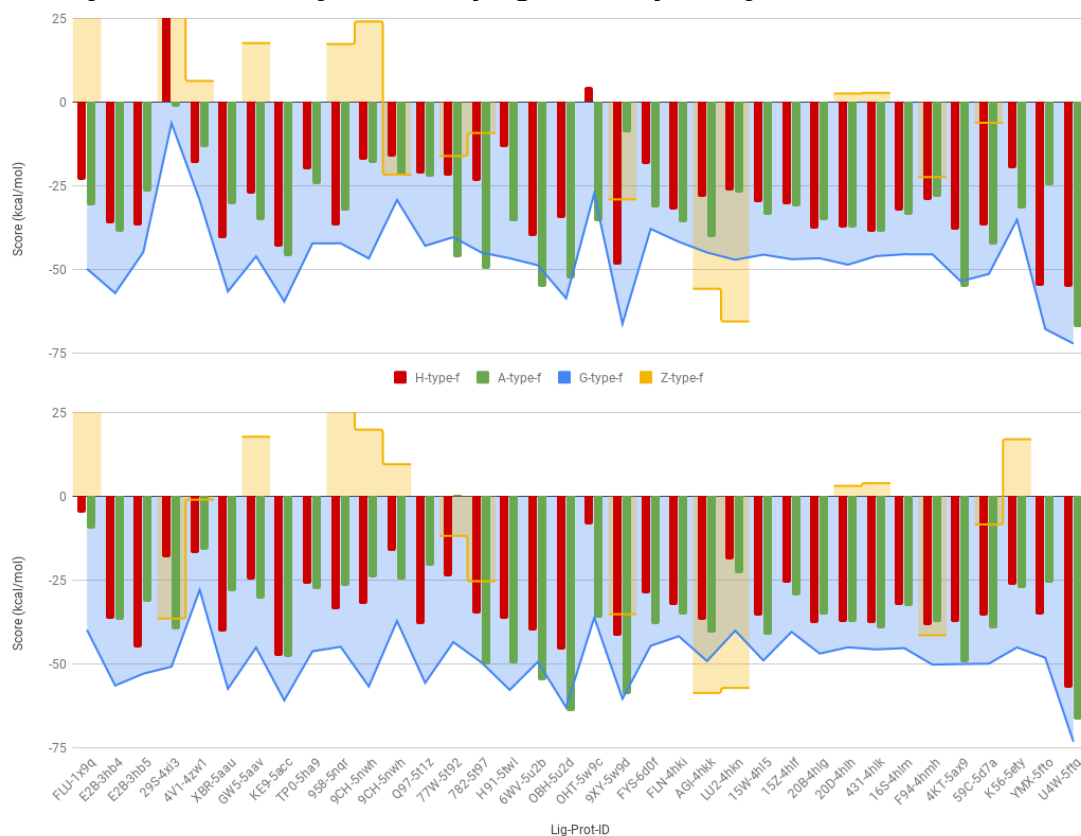
Gambar 2. Nilai Skor Amber ligan kristalografi (*original*) dan ligan asli hasil *docking* yang dipertahankan rigid dan fleksibel.

Skor Amber yang diperoleh pada ligan hasil *docking* memperlihatkan bahwa

adanya perubahan konformasi ruang ligan yang tidak saling mempengaruhi dan

berhubungan. Namun terdapat 16 protein menunjukkan hasil skor amber ligan aslinya yang konsisten (Gambar 2). Terdapat faktor luas permukaan yang

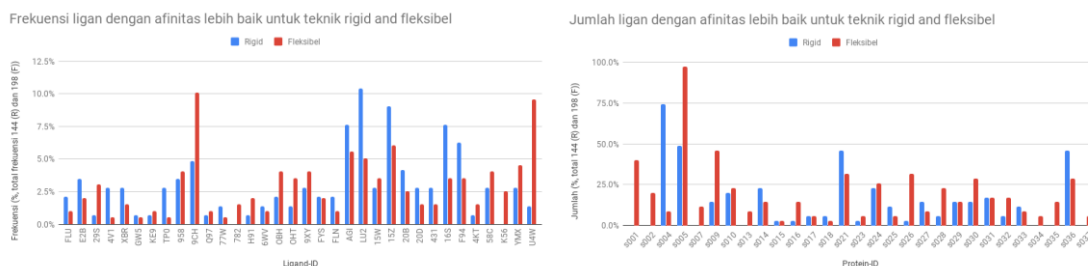
bersifat polar pada ligan dan nilai RMSD pada validasi sistem yang berkontribusi terhadap hasil skor perhitungan. Hal ini menjadi perhatian dalam *docking*.



Gambar 3. Perbandingan nilai skor Grid, Hawkins, Zou, and Amber pada konformasi hasil docking teknik rigid (atas) dan fleksibel (bawah).

Skor grid menunjukkan nilai paling rendah dibandingkan dengan ketiga skor lainnya (Gambar 3). Bahkan terdapat ligan asli yang tidak diperoleh nilai skor pelarut Zou. Berbeda dengan skor grid, Hawkins

dan Amber, skor Zou memiliki profil yang tidak konsisten dikarenakan tergantung dengan spesi ligan yang dapat memberikan kontribusi energi elektrostatis [1]. Hal ini akan diabaikan pada interaksi hidrofobik.



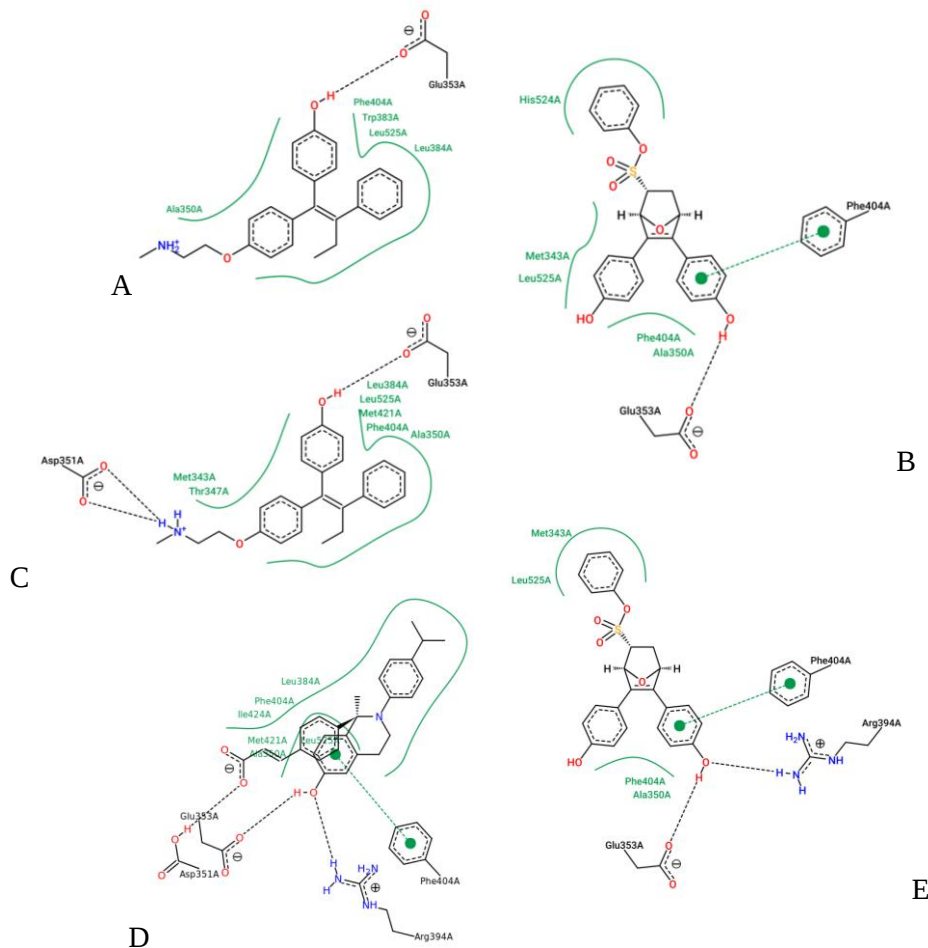
Gambar 4. Frekuensi dan jumlah ligan dengan afinitas lebih baik dibandingkan dengan ligan asli setiap protein pada teknik rigid (biru) dan fleksibel (merah).

Frekuensi kemunculan dan jumlah ligan yang memiliki afinitas ikatan lebih baik dibandingkan dengan ligan asli pada protein target ditampilkan pada Gambar 4. Ligan U4W dan 9CH mempunyai frekuensi tertinggi pada teknik fleksibel

diikuti dengan ligan LU2 dan 15Z untuk teknik rigid pada proses pengukuran *docking* molekular. Protein target s005 dan s004 secara berurutan mempunyai ligan dengan afinitas ikatan terbaik terbanyak untuk pemodelan silang menggunakan

teknik fleksibel (97%) dan rigid (74%)

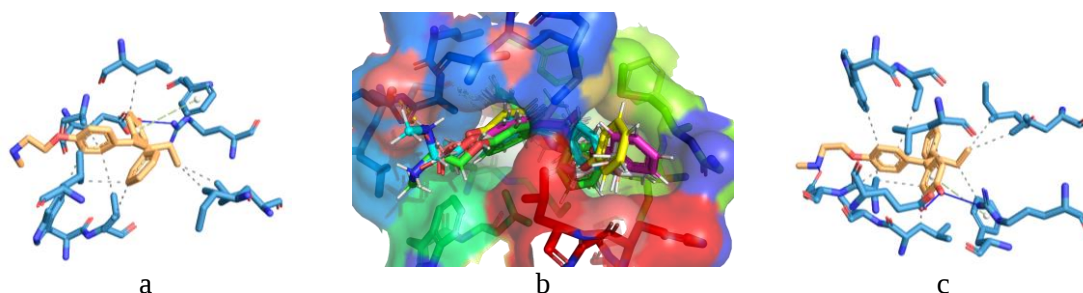
dibandingkan dengan 35 protein lainnya.



Gambar 5. Prediksi interaksi ligan protein untuk ligan 9XY (A: rigid; C: fleksibel), OBH (B: rigid; E: fleksibel) dan 782 (native, D) pada protein s017 hasil pengukuran dengan PoseView online [9].

Skor Amber dan Hawkins telah diperoleh untuk semua ligan native, namun hanya 47,22% protein target yang mempunyai skor Zou terhadap ligan *native*-nya (Gambar 3). Hasil pengukuran afinitas ikatan ligan native secara silang menunjukkan bahwa terdapat 61,11 dan 80,56 % protein yang nilai afinitas ikatan ligan *native*-nya lebih rendah dibandingkan dengan ligan native dari protein lainnya. Gambar 5 memperlihatkan terdapat tiga jenis interaksi yang terjadi antara ligan dan protein pada s017, yaitu interaksi hidrofobik, ikatan hidrogen, dan interaksi

phi-stack. Tipe interaksi phi-stack tidak teridentifikasi pada hasil PoseView namun PLIP menunjukkan hal yang berbeda. Interaksi tersebut diperoleh pada hasil akhir prediksi secara *online* seperti yang telah ditampilkan pada Gambar 6. Residu fenilalanin PHE-404 yang bertanggung jawab terhadap interaksi tersebut. Ikatan hidrogen pada kedua ligan terbaik diperoleh dari hasil interaksi gugus fungsional pada ligan dengan residu GLU-353. Residu leusin pada posisi LEU-525 menjadi penentu kekuatan interaksi hidrofobik yang perlu diteliti.



Gambar 6. Konformasi ligan native dan 2 ligan terbaik lainnya pada protein s017 (b) dengan hasil prediksi interaksi ligan protein untuk ligan 9XY (a: rigid; c: fleksibel) menggunakan PLIP (*Protein Ligand Interaction Profiler*) online [10]. Ligan OBH dan *native* tidak ditampilkan.

Ketiga ligan dalam hal ini 782, 9XY, dan OBH yang dijadikan contoh profil interaksi menginformasikan secara visual bahwa ketiganya pada konformasi yang berhimpit pada *binding site pocket* protein s017 (Gambar 6b). Nilai skor grid, Hawkins, Zou, dan Amber untuk ketiga ligan (782, 9XY, OBH) pada teknik rigid (r) dan fleksibel (f) secara berurutan adalah -44,97; -23,59; -9,25; -49,95 (782r); -45,86; -29,29; -24,38; -38,37 (9XYr); -45,00; -25,69; -22,36; -41,26 (OBHr); -49,58; -34,98; -25,35; -50,07 (782f); -55,94; -42,21; -38,09; -39,85 (9XYf); dan -50,63; -33,80; -34,17; -38,73 (OBHf). Skor grid, Hawkins, dan Zou menunjukkan hubungan antara urutan ligan terbaik dengan ligan aslinya. Namun terjadi hal sebaliknya dengan skor Amber dimana ligan asli memiliki skor tertinggi dibandingkan dengan ligan terbaik pada skor grid. Adanya perbedaan ini harus menjadi perhatian dan pertimbangan dalam memutuskan apakah suatu ligan perlu dilanjutkan ke tahap uji coba laboratorium.

4. KESIMPULAN

Teknik *re-scoring* telah terbukti memberikan penilaian lebih baik dalam melakukan evaluasi afinitas ikatan ligan protein dan konformasi ruang ligan dalam binding site pocket setiap protein.

REFERENSI

- [1] Hawkins GD, Cramer CJ, Truhlar DG. Pairwise solute descreening of solute charges from a dielectric medium. *Chemical Physics Letters*; 1995; 246(1-2): 122-129. doi:10.1016/0009-2614(95)01082-K
- [2] Hawkins GD, Cramer CJ, Truhlar DG. Parametrized Models of Aqueous Free Energies of Solvation Based on Pairwise Descreening of Solute Atomic Charges from a Dielectric Medium. *The Journal of Physical Chemistry*; 1996; 100 (51): 19824-19839. doi: 10.1021/jp961710n
- [3] Tsui V, Case, DA. Theory and applications of the generalized born solvation model in macromolecular simulations. *Biopolymers*; 2000; 56: 275-291. doi:10.1002/1097-0282(2000)56:4<275::AID-BIP10024>3.0.CO;2-E
- [4] DOCK 6.9. University of California at San Francisco; San Francisco, CA. November 2018.
- [5] Allen WJ, Balias TE, Mukherjee S, Brozell SR, Moustakas DT, Lang PT, Case DA, Kuntz ID, Rizzo RC. DOCK 6: Impact of New Features and Current Docking Performance. *J. Comput. Chem.* 2015;36(15):1132-56. doi:10.1002/jcc.23905
- [6] Brozell SR, Mukherjee S, Balias TE, Roe DR, Case DA, Rizzo RC. Evaluation of DOCK 6 as a pose generation and database enrichment tool. *J. Comput. Aided. Mol.* 2012; 26(6):749-773. doi:10.1007/s10822-012-9565-y
- [7] Benet LZ, Hosey CM, Ursu O, Oprea TI. BDDCS, the Rule of 5 and drugability. *Adv Drug Deliv Rev.* 2016;101:89-98. doi:10.1016/j.addr.2016.05.007
- [8] Santoso B. Validasi Ruang Komputasi Senyawa Native Tidak Memenuhi Aturan Lipinski menggunakan PLANTS. In: *Proceeding of The 8th University Research Colloquium 2018: Bidang*

- MIPA dan Kesehatan*. Purwokerto; 2018. p. 98–102.
- [9] Stierand K, Maaß P, Rarey M. Molecular Complexes at a Glance: Automated Generation of two-dimensional Complex Diagrams. *Bioinformatics*. 2006; 22:1710-1716.
- [10] Salentin S, Schreiber S, Haupt VJ, Adasme MF, Schroeder M. PLIP: fully automated protein-ligand interaction profiler. *Nucl Acids Res*. 2015; 43 (W1): W443-W447. doi: 10.1093/nar/gkv315