

PERAN BIOINFORMATIKA DALAM KAJIAN INTERAKSI PROTEIN- PROTEIN

Arli Aditya Parikesit

Departemen Ilmu Komputer, Universitas Leipzig, Jerman

Härtelstr. 16-18

D-04107 Leipzig

email: arli@daad-alumni.de

Abstract

Molecular Biology is advancing through the protein-protein interaction studies. Its wet experiment has generated a massive amount of data. Bioinformatics was set up for converting those data into useful information. Protein-protein interaction wet experiment was a long and expensive labours. Bioinformatics are here to remedy those labours. Online experimentations at protein-protein interaction database are possible. The available databases are InterPare, ClusPro, and PROTORP. InterPare is an open and public database server for protein interaction interface information. ClusPro is an algorithm for filtering docked protein conformations, and rank them. While PROTORP is a database to calculates a series of physical and chemical parameters of the protein interaction sites that contribute to the binding energy of the association. The online databases are useful tools for aiding the wet laboratory protein-protein interaction experiment.

Keywords: *Bioinformatics, database, InterPare, ClusPro, PROTORP*

1. Pendahuluan

Eksperimen basah biologi molekuler telah menghasilkan data biologi dalam jumlah banyak. Data tersebut diserahkan kepada *online* database. Bioinformatika adalah ilmu gabungan antara ilmu biologi dan ilmu komputer. Tugas bioinformatika adalah mengolah data tersebut menjadi informasi yang berguna. *Online* database yang tersedia adalah GenBank dari Amerika Serikat (<http://www.ncbi.nlm.nih.gov>), DDBJ dari Jepang (<http://www.ddbj.nig.ac.jp/>), dan EBI dari Uni Eropa (<http://www.ebi.ac.uk/>) (Claverie *et al*, 2003).

Protein adalah kelas molekul yang sangat penting dalam sel. Sebagian besar fungsi protein adalah berinteraksi dengan molekul lain, terutama dengan sesama protein. Interaksi protein mengalami regulasi dan terkonservasi pada evolusi. Hal ini terjadi karena interaksi tidak sempurna yang dipicu oleh mutasi random, dapat menyebabkan disfungsi molekul. Oleh sebab itu, area interaksi interface molekul berada dibawah tekanan oleh seleksi alam dan lebih terkonservasi dibandingkan dengan bagian lain pada protein. Sejalan dengan semakin lengkapnya sekuens genom, 'interaktomik struktural' protein untuk memetakan semua interaksi domain protein menjadi semakin penting. Sekarang ilmuwan dapat memetakan seluruh interaktom manusia secara bioinformatika, menggunakan data eksperimen yang berlimpah dari berbagai metode, seperti analisis *yeast two hybrid* (Gong *et al*, 2005).

Beberapa metode penting prediksi interaksi protein membutuhkan elusidasi struktur dan kompleks protein yang telah dipecahkan secara eksperimen, dan hal ini dapat diperoleh dengan kristalisasi. Struktur protein dapat dipecahkan dengan menggunakan hasil sintesis kristal, menggunakan difraksi sinar X atau analisis difraksi neutron. Di sisi lain, *nuclear magnetic resonance* (NMR) adalah teknik yang digunakan pada protein larutan, biasanya untuk protein yang tak dapat dikristalisasi. Penggunaan teknik NMR untuk memecahkan berbagai struktur protein ternyata mengalami keterbatasan, karena teknik ini hanya mampu memecahkan struktur atau kompleks berukuran kecil (Sekitar 30 KDa). Spektrum NMR diperoleh dengan meletakkan sampel pada medan magnetik, dan mengaplikasikan pulsa frekuensi radio (Taylor *et al*, 2005).

Walaupun selama 30 tahun terakhir sudah ada upaya keras untuk membongkar misteri interaksi protein-protein, namun tidak ada hasil yang signifikan. Data interface dari berbagai pendekatan yang ada, tidak diatur dengan baik dan tidak dibagi dengan ilmuwan yang lain. Namun, dengan bantuan kristlografi sinar X dan NMR yang lebih baik pada area biologi struktural, terjadi peningkatan signifikan pada jumlah struktur tiga dimensi protein yang diketahui. Informasi struktur 3D ini adalah sumber yang terbaik untuk data kajian interface protein (Gong *et al*, 2005).

Struktur protein yang telah dipecahkan, bersama model teoritis, didepositkan pada bank data struktur protein yang disebut *Protein Data Bank* (PDB). Bank data ini dapat diakses pada <http://www.rcsb.org>. Pada 25 Agustus 2009, ia memiliki 59790 struktur yang terdepositkan, namun hanya sebagian kecil dari semua ini yang merepresentasikan kompleks protein-protein. Hal ini terjadi, karena kristalisasi merupakan prosedur yang sangat rumit dan memakan waktu, dimana melibatkan eksperimen secara *trial and error* untuk menemukan kondisi optimum. Oleh karena itu, dalam mempelajari interaksi protein-protein, diperlukan bantuan bioinformatika secara lebih ekstensif (Taylor *et al*, 2005).

Tantangan untuk memprediksi kedudukan protein (protein docking) dimulai dengan kordinat dari komponen molekul tak terikat dan memperoleh model untuk model ikatan. Dengan pengembangan pendekatan korelasi Fourier, maka menjadi mungkin secara komputasi, untuk mengembangkan

dan mengevaluasi sekian milyar konformasi dudukan yang dimungkinkan, dengan fungsi angka sederhana. Dimulai dari protein tak terikat, metode ini dapat sekaligus menghasilkan struktur yang mendekati aslinya dan banyak positif palsu yang memiliki komplemen permukaan. Hal ini terutama jika faktor elektrostatis dimasukkan dalam potensial, namun tetap berbeda bentuknya dari kompleks asli. Alasan utama dari kejadian ini, adalah karena ketidakpastian intrinsik dari dudukan struktur protein, misalnya pada posisi gugus fungsi yang terpapar oleh pelarut (Comeau *et al*, 2003).

Dalam beberapa tahun terakhir ini, hasil yang penting telah dicapai dalam rangka mengembangkan metode untuk mengurutkan ulang konformasi dudukan dan mencoba untuk memilih dudukan yang mendekati konformasi asli, biasanya menggunakan potensial yang menghitung afinitas kimiawi antara molekul, dan kemungkinan mengatur juga permukaan interaksi. Prosedur ini meningkatkan diskriminasi, dan 'hits' (konformasi dengan RMSD dibawah 10 Å) dapat ditemukan pada 10 struktur yang paling baik (Comeau *et al*, 2003).

Asosiasi protein-protein dapat diklasifikasikan sebagai permanen atau sementara, tergantung pada waktu hidup dan tipe asosiasi. Berbagai tipe kelas asosiasi menunjukkan properti fisis dan kimiawi yang berbeda pada situs interaksi. Kombinasi dari berbagai properti situs interaksi tersebut telah digunakan untuk prediksi situs interaksi (Reynolds *et al*, 2008).

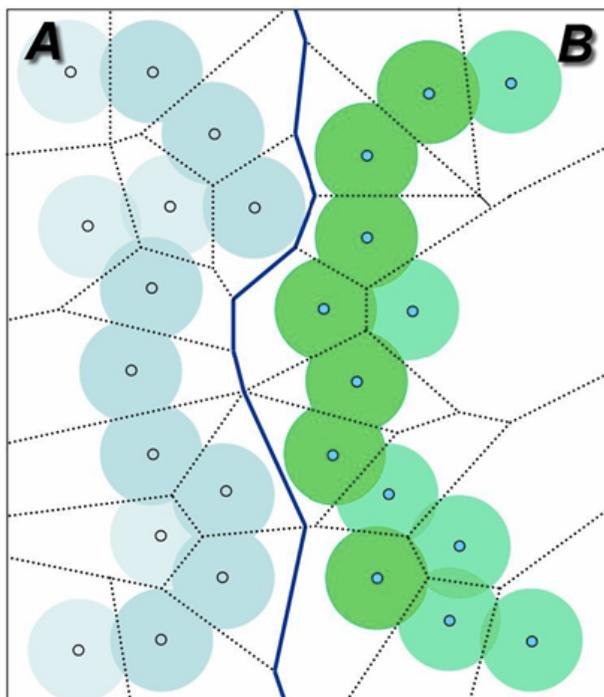
2. Database Interaksi Protein-Protein

Eksperimen basah interaksi protein-protein memerlukan waktu yang panjang dan regen biokimia yang mahal. Diperlukan metode yang mampu mempersingkat waktu dan menekan kebutuhan regen. Bioinformatika dapat membantu tercapainya kedua hal tersebut. Ada tiga database *online* yang akan dibahas dalam ulasan ini. Mereka adalah InterPare, ClusPro, dan PROTORG.

2.1 InterPare

Database interaksi interface protein skala besar telah diperkenalkan pada web <http://www.interpare.net> atau <http://psimap.org>. Nama database tersebut adalah InterPare. Database ini menyajikan interface antara protein domain, yang diidentifikasi oleh tiga metode. Pertama, interface dideteksi dengan menghitung jarak geometrik antara subunit protein multidomain atau kompleks protein pada PDB. Pada pendekatan kedua, area protein yang terkubur di dalam diidentifikasi dengan menghitung *accessible surface area* (ASA) ketika mereka membentuk kompleks atau agregat dengan subunit atau domain lain. Area tersebut dapat diakses air ketika mereka menjadi subunit bebas. Akhirnya, interface dapat didefinisikan oleh pendekatan geometrik dan topologis menggunakan diagram voronoi (gambar 1). Struktur interface dari protein antrian, dalam konteks dari seluruh konfigurasi protein, dapat dilihat dengan menggunakan tiga *molecular viewer* berbeda.

Mereka adalah Chime, Jmol, dan InterFacer. InterPare juga menyediakan file kordinat atom untuk permukaan protein, interior, dan interface untuk analisis lebih lanjut (Gong *et al*, 2005).



Gambar 1: Diagram Voronoi. Diagram power antara dua domain berbeda dengan representasi 2D. Lingkaran biru muda (atom) berada di domain A, dan atom hijau berada di domain B. Garis titik-titik menandakan ujung Voronoi antara dua atom yang bertetangga, sementara garis solid menandakan interface geometris Voronoi antara dua domain. Semua poligon yang berhadapan dengan setidaknya satu interface geometris Voronoi adalah interface sel. Jika sebuah sel adalah interface sel, maka kita sebut atom pada sel tersebut sebagai atom interface. Atom interface adalah sedikit lebih gelap daripada yang non interface. Database InterPare menyimpan semua informasi atom interface.

Protein interaktom dapat didefinisikan sebagai seluruh set dari interface interaksi protein yang ditemukan di sel. Ada banyak metode untuk mendefinisikan set data interface. Konsep yang digunakan adalah klasifikasi hirarkis domain protein dari SCOP. Klasifikasi SCOP diperluas pada interface molekuler. Keunggulan pendekatan ini adalah setiap interface dapat didefinisikan pada konteks evolusi domain. Superfamili SCOP adalah tingkat klasifikasi dimana struktur protein diketahui secara jelas untuk berhubungan didalam grup klasifikasi. Tingkat famili protein di SCOP adalah kelas yang relevan secara fungsional, dimana setiap anggota famili adalah berhubungan dan mirip secara fungsional. Dibawah famili, terdapat domain individual. Sistem klasifikasi protein apapun seperti FSSP dan CATH, dapat juga digunakan. Kontribusi utama pada bioinformatika struktural adalah interface dapat dicari dan dibandingkan dengan komputer. Hal tersebut dapat dilakukan dengan bantuan InterPare (Gambar 2). Kluster yang secara hirarkis serupa pada keseluruhan interface diharapkan supaya sangat terkonservasi, untuk menjaga partner interaksi mereka (Gong *et al*, 2005).

a

OPTIONS	
PDB ID	(Ex: 1A25)
Relative ASA(%)	<input type="radio"/> 5 <input type="radio"/> 10 <input type="radio"/> 15 <input type="radio"/> 20 <input type="radio"/> 25 <input type="radio"/> 30
Interface Threshold	Δ ASA (\AA^2)
	Distance (\AA)
	<input type="text" value="1"/> \AA^2 <input type="text" value="5.0"/> \AA
<input type="button" value="search"/> <input type="button" value="clear"/>	

b

1 Hit from PDB

PDB ID	RSA file	ASA file	Shell(Exterior)	Core(Interior)	Face(Interface)	Full View
1a25	1a25.rsa	1a25.asa	1a25.surface <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	1a25.interior <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	1a25.interface <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	<input type="button" value="CHIME"/> <input type="button" value="JMOL"/>

2 Hit from SCOP domain

SCOP ID	RSA file	ASA file	Shell(Exterior)	Core(Interior)	Face(Interface)	Full view
d1a25a	d1a25a.rsa	d1a25a.asa	d1a25a.surface <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	d1a25a.interior <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	d1a25a.interface <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	<input type="button" value="CHIME"/> <input type="button" value="JMOL"/>
d1a25b	d1a25b.rsa	d1a25b.asa	d1a25b.surface <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	d1a25b.interior <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	d1a25b.interface <input type="button" value="CHIME"/> <input type="button" value="JMOL"/>	<input type="button" value="CHIME"/> <input type="button" value="JMOL"/>

c

MDL

PDB	: 1a25
RED	: protein surface(Shell)
BLUE	: protein interior(Core)
SPACE FILL	: protein interface(Face)

d

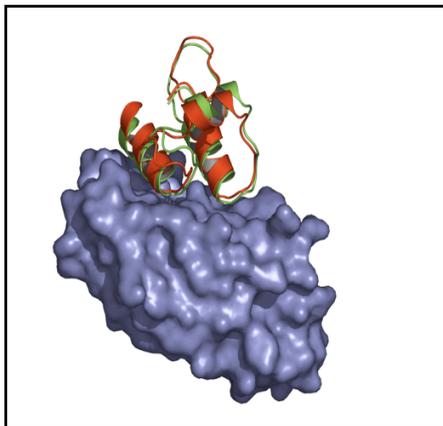
PDB	: 1a25
RED	: protein surface(Shell)
BLUE	: protein interior(Core)
SPACE FILL	: protein interface(Face)

Gambar 2: Layar dari halaman hasil situs web InterPare. (a) Boks pertanyaan dengan beberapa opsi. “Relative ASA” adalah kriteria untuk memahami area interior protein dari area permukaan. Ada dua kriteria independen untuk definisi interface: “ Δ ASA” untuk metoda ASA, dan ‘Distance’ untuk algoritma PSIMAP. (b) Hasil menanyakan Database dengan ‘1a25’ sebagai ID PDB. 1a25 adalah domain C2 dari protein kinase C beta. Ia terdiri dari dua domain identik (d1a25a_ dan d1a25b_). Hasil dari metode ASA ditunjukkan disini. Koordinat atom pada area permukaan protein (Shell), interior (Core), dan Interface (Face) dapat diunduh dari halaman ini. (c), (d) Struktur 3D protein ‘1a25’ oleh Chime dan Jmol. Area interior berwarna biru, dan permukaan berwarna merah. Area Interface ditunjukkan dalam model *space fill* dan area non interface ditunjukkan pada *wireframe* dengan mode *backbone tracing*.

2.2 ClusPro

Algoritma diskriminasi dan otomatisasi dudukan *rigid body* dapat menyaring konformasi dudukan, dan memberikan peringkat berdasarkan properti kluster. Metode ini telah diimplementasikan pada server web yang bernama *ClusPro*. Penyaringan melibatkan penggunaan metode evaluasi empiris energi bebas, yang memilih konformasi dengan desolvasi terendah dan energi elektrostatik. Metode Kluster, yang pertama kali diimplementasikan oleh Camacho dan Gatchell untuk eksperimen 2001 *Critical Assessment of PRedictions of Interactions* (CAPRI), dimotivasi oleh pen-

gamatan bahwa energi bebas dari kompleks reseptor-ligan yang tersolvasi secara parsial menunjukkan atraktor energi bebas situs ikatan. Server ClusPro dapat diakses di <http://nrc.bu.edu/cluster/> (Comeau *et al*, 2003).



Gambar 3 : Prediksi CAPRI Target 12 dengan server Cluspro, kompleks Cohesin/Dockerin. Cohesin dari prediksi kami dan struktur kristal telah di *superposed* (pada permukaan biru). Posisi hasil prediksi untuk dockerin adalah hijau, sementara dockerin dari struktur kristal berwarna merah.

Server ClusPro telah digunakan untuk membantu eksperimen basah pada reseptor CLEC-2. Eksperimen ini telah dilakukan oleh Watson *et al* pada tahun 2006. Kajian tersebut berguna untuk memberikan kerangka untuk memahami efek dari pengikatan racun rhodositin pada CLEC-2 dan untuk memahami ligan endogenus yang terkait. Diharapkan kajian ini akan menjadi basis bagi desain rasional obat untuk memblokir pengikatan ligan, dalam rangka mendeaktivasi racun.

2.3 PROTORP

PROTORP adalah server yang menghitung properti interaksi dari struktur 3D protein individu dan untuk seluruh set data secara *real time*. Hal ini memberikan keunggulan, yaitu pengguna tidak perlu dibatasi oleh struktur yang disimpan di database dan dapat memanipulasi file input untuk menganalisis interface alternatif yang mereka pilih (seperti asosiasi domain-domain). Properti yang dihitung dapat dibandingkan dengan parameter distribusi dari data non-homolog kelas asosiasi protein. Informasi ini dapat berguna untuk membedakan asosiasi biologis dari kontak kristal. Dengan pengetahuan yang bertambah terhadap properti asosiasi protein-protein, maka adalah mungkin untuk memprediksi interaksi protein secara biologis. PROTORP dapat diakses di <http://www.bioinformatics.sussex.ac.uk/protorp/> (Reynolds *et al*, 2008).

Protein Interface Parameter		Value
List of interface residues		
Interface Residue Segments		2
Interface Accessible Surface Area (Å ²)		469.23
% Interface Accessible Surface Area		11.74
Atoms in Interface		43
% Polar Atoms Contribution to Interface		45.02
% Non-Polar Atoms Contribution to Interface		35.06
% Neutral Atoms Contribution to Interface		18.96
Residues in Interface		15

PROTEIN	FACE1	FACE2	SEGMENTS	ASA	%_ASA	ATOMS
1CDT	A	B	2	469.23	11.74	43
1VSG	A	B	9	4238.08	23.39	338
1FDL	L	H	7	1900.96	16.59	155
3HFM	H	Y	3	473.31	4.00	45

Gambar 3: Output Protorp. Server Protorp menghitung dan menunjukkan parameter interface untuk asosiasi protein-protein tunggal sebagai halaman HTML (Bagian atas gambar) dan sebagai multi asosiasi untuk output text yang dapat diunduh (bagian bawah gambar).

Server PROTORG adalah berbasis file struktur data dari Protein Data Bank (PDB). Pada PDB, partner protein pada asosiasi dibedakan berdasarkan identifikasi subunit (www wwpsdb.org/docs.htm). Sebagai contoh, protein Cardiotoksin (kode PDB 1cdt) adalah protein homodimerik dengan dua subunit identik, dalam bentuk identifikasi A dan B pada format file PDB. Server PROTORG menghitung properti interaksi untuk satu pemasangan subunit-subunit pada satu waktu. Dengan cara ini, kontak biologis dan fisis dapat dianalisis (Reynolds *et al*, 2008).

PROTORG berperan penting untuk mempelajari hubungan antara adhesi sel dan sitoskeleton. Hal ini penting untuk mengetahui peran hubungan tersebut pada bentuk sel, motilitas, dan signaling. Dalam percobaan yang dilakukan oleh Lad *et al* pada tahun 2008. dengan menggunakan bantuan PROTORG, telah diketahui bahwa terjadi kompetisi antara ligan-ligan filamen untuk mencapai situs pengikatan pada domain IgFLN. Hal ini memberikan pemahaman pada modulasi interaksi dan signaling filamin.

3. Kesimpulan

InterPare adalah server database publik untuk informasi interaksi interface protein. Database tersebut memiliki data interface skala besar untuk protein yang struktur 3D nya telah diketahui. Interface tersebut berbasis pada domain protein yang tercatat pada database SCOP.

ClusPro adalah server terintegrasi pertama yang memiliki langkah docking dan diskriminasi untuk memprediksi struktur kompleks protein-protein. Server tersebut dapat digunakan untuk mendiskriminasi satu set struktur kompleks potensial dari beberapa algoritma docking.

PROTORP adalah server yang dapat menghitung parameter fisis dan kimia dari situs interaksi protein. Perhitungan tersebut adalah untuk memprediksi energi ikatan asosiasi. Parameter yang dilibatkan adalah ukuran dan bentuk, ikatan intermolekuler, komposisi gugus fungsi dan atom, dan struktur sekunder.

Ketiga database online tersebut dapat dimanfaatkan untuk membantu eksperimen basah interaksi protein-protein

Kepustakaan

Gong, Sungsam., Park, Changbun., Hansol, Choi., Ko, Junsu., Jang, Insoo., Jungsul, Lee., Bolser, Dan M., Oh, Donghoon., Kim, Deok-Soo., and Bhak, Jong. 2005. "A Protein domain interaction interface database: InterPare". *BMC Bioinformatics* **Vol.6**: 207.

Claverie, Jean Michel., Notredame, Cedric. 2006. *Bioinformatics for Dummies*. New Jersey: John Wiley and Sons

Comeau, Stephen R., Gatchell, David W., Vajda, Sandor and Camacho, Carlos. 2003. "ClusPro: an Automated docking and Discrimination Method for the Prediction of Protein Complexes". *Oxford Journal of Bioinformatics* **Vol 20** no.1: 45-50.

Lad, Yatish., Jiang, Pengju., Ruskamo, Salla., Harburger, David S., Ylanne, Jari., Campbell, Lain, D., and Calderwood, David A. 2008. "Structural Basis of the Migfilin-Filamin Interaction and Competition with Integrin β Tails". *The Journal of Biological Chemistry* **Vol 283** No 50: 35154-35163

Reynolds, Christopher., Damerel, David and Jones, Susan. 2008. "ProtorP: a Protein-Protein Interaction Analysis Server". *Oxford Journal of Bioinformatics* **Vol 25** no 50: 413-414.

Taylor, A-Walker., Jones, D.T., 2005. "Computational Methods for Predicting Protein-Protein Interaction". *Protein Review* **Vol 3**: Proteomics and Protein-Protein Interaction: 89-114

Watson, Alexandra A., Brown, James., Harlos, Karl., Eble, Johannes A., Walter, Thomas S., O'Callaghan, Christopher A. 2006. "The Crystal Structure and Mutational Binding Analysis of the Extracellular Domain of the Platelet-activating Receptor CLEC-2". *The Journal of Biological Chemistry* **Vol 282** No 5: 3165-3172

ARLI ADITYA PARIKESIT

Lahir di Jakarta, penulis menyelesaikan S1 dan S2 pada Departemen Kimia, Fakultas Matematika dan Ilmu Pengetahuan Alam, Universitas Indonesia. Setelah itu, sempat membantu menjadi asisten peneliti dan mengajar di alamatnya secara tidak tetap. Pada tahun 2008, penulis mendapat kesempatan untuk melanjutkan studi S3 di Jerman, dengan bantuan beasiswa DAAD (Dinas pertukaran akademis Jerman). Sekarang terdaftar sebagai kandidat doktor di Departemen Ilmu Komputer, Fakultas Matematika dan Informatika, Universitas Leipzig, Jerman. Minat penelitian terutama pada aplikasi bioinformatika pada penelitian biomedis, dan desain algoritma/struktur data bioinformatika. Penulis adalah co-autor pada artikel ilmiah 'In Silico Analysis of Envelope Dengue Virus-2 and Envelope Dengue Virus-3 Protein as the Backbone of Dengue Virus Tetravalent Vaccine by Using Homology Modeling Method', yang telah dipublikasi oleh *OnLine Journal of Biological Sciences*.