

***Candida antarctica* Lipase B Synthetic Gene: A Bioinformatics Analysis**

Febriana Dwi Wahyuni ^{1*}, Seprianto ¹

Program Studi Bioteknologi, Universitas Esa Unggul, Jakarta¹

febriana@esaunggul.ac.id

Abstract. *Enzyme exploration is important to support the development of biotechnology. To facilitate microbial discovery, it can be done synthetically using bioinformatics methods. Candida antarctica lipase B (CALB) is one of enzyme derived from microorganisms and has been applied as a biocatalyst in several industry. The methods used in this study are the analysis of gene structure using NCBI, analysis of protein CALB sequence using Uniprot, analysis of 3D protein structure using Swiss model, analysis primary design using primer3 and analysis of restriction sites using snappgene. The construction of synthetic calB gene obtained based on the results of the CalB gene sequence in geneBank using NCBI with access number Z30645.1. CalB gene wildtype is modified by adding several restriction enzyme (XhoI, XbaI, ClaI, and BglI), 6 Histags, and 2x stops codon and produce 1083 base pairs. CALB protein has 342 amino acids. Based on 3D structure, CALB protein has three molecules in common with one homotrimer ring that will encircle the double helix DNA. Primers used for calB gene amplification are CALB forward 5'-TCCCCAGTATCAGGTCCAAG-3' and CALB reverse 5'-GACACCTGAGGCTGAACGAT-3'.*

Katakunci: Bioinformatika, gen sintetik, *Candida antarctica* Lipase B



This is an open access article distributed under the Creative Commons 4.0 Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. ©2018 by author.

1. PENDAHULUAN

Enzim merupakan makromolekul yang berperan dalam mempercepat reaksi kimia dalam sel hidup. Eksplorasi enzim merupakan hal yang penting untuk menunjang perkembangan teknologi bio yang ramah lingkungan. Pencarian mikroba secara tradisional dengan kultivasi sekarang ini menjadi lebih tidak efisien karena menghabiskan biaya dan tenaga (Narita, 2012). Solusi yang bisa dilakukan untuk eksplorasi enzim secara efektif yaitu dengan metoda kajian bioinformatika. Menurut kamus Oxford, bioinformatika adalah ilmu dalam mengumpulkan dan menganalisa data-data biologi seperti kode genetik (Saraswati, 2017). Bioinformatika saat ini menjadi alternatif pencarian sekuen gen atau enzim baru karena didalamnya terdapat berbagai informasi biologis yang tersaji dalam

suatu *online* database. *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).

Berdasarkan laporan dari *Global Industry Analysts Inc.*, pasar dunia untuk enzim diperkirakan mengalami peningkatan sampai 2,9 miliar USD pada tahun 2012 dan meningkat sampai 7 miliar USD pada tahun 2013 dengan rata-rata peningkatan sekitar 6,3% (Ulfah, 2013). Penggunaan enzim yang semakin meningkat, salah satunya yaitu di dalam bidang industri, mendorong para peneliti untuk mengembangkan enzim secara rekombinan (Wahyuni, 2016). Para peneliti telah mencoba untuk mengembangkan enzim, misalnya lipase yang berasal dari mikroorganisme. Lipase yang dikembangkan secara rekombinan diharapkan memiliki sifat yang ramah terhadap lingkungan, dapat mengurangi biaya produksi, serta dapat meningkatkan kualitas produk (Hasan, dkk. 2006).

Candida antarctica lipase B (CALB) merupakan salah satu jenis lipase yang berasal dari mikroorganisme dan telah banyak diaplikasikan sebagai biokatalis dalam beberapa bidang industri (Wahyuni, 2016). Pemanfaatan enzim CALB dalam bidang industri ini karena sifat enzim CALB yang dapat diaplikasikan pada suhu rendah dan dapat memiliki termostabilitas yang baik terhadap peningkatan suhu serta memiliki stabilitas yang tinggi dalam pelarut organik (Joseph *et al.* 2008; Pfeffer *et al.* 2006).

Lipase yang digunakan pada bidang industri di Indonesia masih diperoleh melalui impor dari negara asing. Harga lipase tersebut relatif lebih mahal, misalnya lipase *Rhizomucor miehei*, yaitu *Lipozyme IM* yang telah diproduksi oleh perusahaan Novo, Nordisk *Bioindustry Inc*, Denmark. Hal tersebut menyebabkan lipase impor tidak ekonomis untuk diaplikasikan dalam skala industri. Oleh karena itu, dilakukan usaha untuk membuat lipase secara rekombinan. Lipase yang akan dibuat adalah enzim CALB yang berasal dari gen sintetik *Candida antarctica* Lipase B (*CalBsyn*).

2 BAHAN DAN METODE

Beberapa metode yang dilakukan dalam penelitian ini adalah: (1) Analisis struktur, fungsi, ekspresi gen *CalBsyn* dengan menggunakan database bio-informatika dari situs <http://www.ncbi.nlm.gov> dan PubMed, <http://www.uniprot.org> Sekuen gen *cbm* yang diperoleh kemudian dioptimasi dengan kodon *Pichia pastoris* menggunakan program DNAWorks 3.2; (2) Analisis sekuen protein CALB (domain protein, karakteristik fisiko-kimia, profil skala asam amino, prediksi signal peptide, prediksi target peptide) dengan menggunakan situs <http://www.expasy.org>, Prosite, Protparam, Pro-Scale, Psipred, SignalP, TargetP dan PeptideCutter; (3) Analisis struktur 3D protein CALB dengan menggunakan situs Protein Data Bank, <http://bioinf.cs.ucl.ac.uk/> menu PSIPRED dan 3D

protein menggunakan situs <http://www.pdb.org> menu Swiss Model dan software PYMOL; dan (4) Analisis desain primer dengan menggunakan software Primer3 dari sekuens gen yang memiliki similaritas tinggi (pada daerah basa domain cds) dengan keunikan basa untuk dibuat desain primer. Untuk menghindari hairpin pilih sekuens hasil desain primer software PerlPrimer (5) Analisis situs restriksi dengan menggunakan software Snapgene.

3 HASIL DAN PEMBAHASAN

Analisis sekuens gen *calB*

Berdasarkan hasil pencarian sekuens gen *calB* yang ada di geneBank melalui NCBI dengan nomor akses Z30645.1, diperoleh sekuens gen *calB* dari *Pseudozyma antarctica*, yang terdiri dari 1029 pasang basa (Gambar 1.). Gen *CalB wildtype* kemudian dimodifikasi dengan menambahkan beberapa situs restriksi (*XhoI*, *XbaI*, *Clal*, dan *BglI*), 6Histag, dan 2x stop sehingga menghasilkan ukuran 1083 pasang basa. Sekuens DNA tersebut kemudian dioptimasi dengan kodon *Pichia pastoris* menggunakan program DNAWorks 3.2. Setelah diperoleh sekuens DNA gen *calB*, selanjutnya dicari sekuens protein CALB melalui program Uniprot. Protein CALB mempunyai 342 asam amino.

```

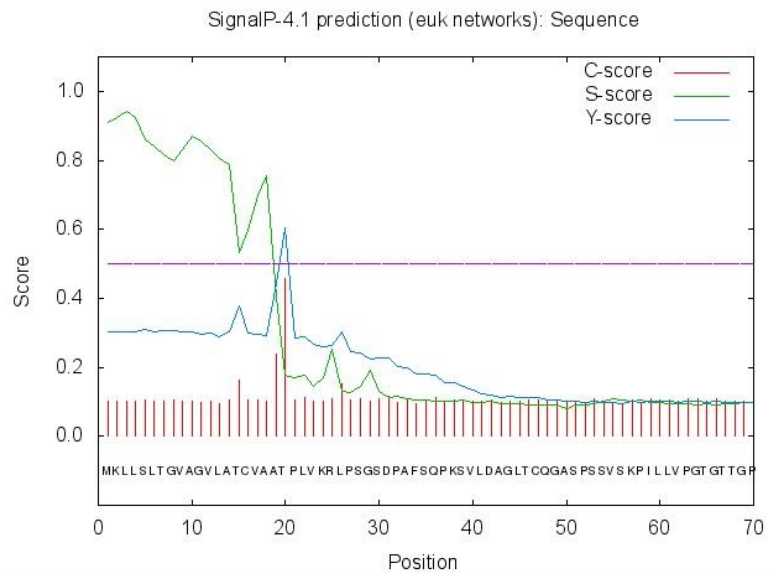
1 atgaagctac tctctctgac cggtgtggct ggtgtgcttg cgacttgcgt tgcagccact
61 ccttttggatg agcgtctacc ttccgggttcg gaccctgcct tttcgcagcc caagtcggtg
121 ctcgatgcgg gtctgacctg ccagggtgct tcgccatcct cggtctccaa acccatcctt
181 ctcgtccccc gaaccggcac cacaggtcca cagtcgttcg actcgaactg gatccccctc
241 tcaacgcagt tgggttacac accctgctgg atctcaccac cgccgttcat gctcaacgac
301 acccaggtca acacggagta catgggtcaac gccatcaccg cgctctacgc tggttcgggc
361 aacaacaagc ttcccgtgct tacctgggtcc cagggtggtc tggttgcaca gtgggtctg
421 accttcttcc ccagtatcag gtccaaggtc gatcgactta tggccttgc gcccgactac
481 aagggcaccg tcctcgccgg ccctctcgat gcaactcggg ttagtgacc ctccgtatgg
541 cagcaaacca ccggttcggc actcaccacc gcaactcggaa acgcaggtgg tctgaccag
601 atcgtgcccc ccaccaacct ctactcggcg accgacgaga tcgttcagcc tcaggtgtcc
661 aactcgccac tcgactcatc ctacctctc aacggaaaga acgtccaggc acaggccgtg
721 tgtgggcccgc tgttcgtcat cgacctgca ggctcgtca cctcgagtt ctctacgtc
781 gtcggtcgat ccgccctgcg ctccaccacg ggccaggctc gtagtgaga ctatggcatt
841 acggactgca accctcttcc cgccaatgat ctgactccc agcaaaaggc cgccggcgt
901 gcgctcctgg cgccggcagc tgcagccatc gtggcgggtc caaagcagaa ctgcgagccc
961 gacctcatgc cctacgcccg cccctttgca gtaggcaaaa ggacctgctc cggcatcgtc
1021 accccctga

```

Gambar 1. sekuens gen *CalB wildtype*

Prediksi signal peptide (SignalP) protein CALB

Untuk mengetahui adanya signal peptide pada protein CALB, dilakukan dengan analisis menggunakan program SignalP (Gambar 2). Dari hasil analisis, dapat diketahui bahwa protein CALB mempunyai signal peptide dengan situs pemotongan ada di posisi 19-20.



Gambar 2. Hasil analisis signal peptide pada protein CALB dengan program signalIP

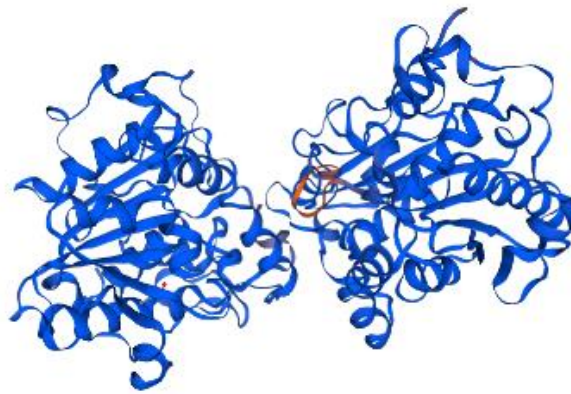
Grafik signalIP menunjukkan tiga skor yang berbeda antara C, S dan Y, untuk setiap posisi dalam sekuens protein. Nilai *C-score* yang merupakan situs pemotongan memiliki nilai maksimum 0,456 terletak pada posisi protein ke 20. Nilai *S-score* (sinyal peptida) terletak pada posisi protein ke 3 dengan nilai maksimum 0,942. Sedangkan *Y-score* (gabungan antara *C-score* dan *S-score*) terletak pada posisi protein ke 20 dengan nilai maksimum 0,606. Informasi situs pemotongan signalIP sangat penting dalam proses ekspresi pada spesies yang berbeda (*heterologous expression*). Sistem ekspresi gen *CaB* dapat diinduksi ke dalam *escherichia coli* BL21 dengan mengetahui urutan signal peptida pada kondisi suhu rendah (20°C) dengan menggunakan sistem pET (Blank et al, 2006). Produksi protein CALB meningkat sesuai dengan aktivitasnya sebesar 480 U/ml diperoleh dengan peptida sinyal yang asli (*wildtype*), sedangkan signalIP alfa memiliki aktivitas maksimum sebesar 160 U/ ml. Selain itu, CalB disekresikan sebagai protein dengan N-terminus ketika berikatan dengan signalP Native, sementara CALB yang tidak proses dengan ekstensi N-terminal dideteksi dengan peptida sinyal alfa (Vadhana et al, 2013).

Analisis struktur sekunder 3D protein (Psipred)

Struktur 3 dimensi protein CALB dapat diketahui menggunakan software Swiss model (Gambar 3). Gambar 3 menunjukkan struktur 3D protein CALB terdapat tiga molekul identik berikatan satu sama lain dengan membentuk sebuah cincin homotrimer yang akan melingkari DNA *double helix*. Molekul Loops (koil) membentuk lipatan yang mana struktur ini mudah mengalami mutasi. Permukaan dalam dari lingkaran tersebut berbentuk α heliks dan bermuatan positif sehingga dapat berinteraksi dengan DNA. Permukaan luar dari

lingkaran ini berbentuk β sheet dan bermuatan negatif. Struktur tersier 3D protein mengacu pada hubungan spasial antar struktur sekunder. Struktur ini distabilkan oleh empat macam ikatan, yakni ikatan hidrogen, ikatan ionik, ikatan kovalen, dan ikatan hidrofobik.

Candida antarctica lipase B (CALB) banyak dimanfaatkan dalam dunia industri karena memiliki selektivitas yang tinggi, jangkauan substrat yang luas, stabil dalam kondisi panas serta dalam berbagai pelarut organik. Struktur CALB stabil terhadap semua pelarut organik karena memiliki α -helix pendek (residu 139–150), α -helix panjang terdapat pada residu 266–289 yang berikatan dengan sisi aktif dengan membentuk tiga lipatan (*loop*) di permukaan sisi aktif (residu 26-30, dan 92–97, 215–222), hal ini menunjukkan adanya perbedaan fleksibilitas yang lebih tinggi antara air dan pelarut non-polar. Substrat dapat berikatan dengan alkohol sekunder (residu 39, 42, 47, 104, 225) dan dapat mengikat gugus asil (residu 134, 138, 157, 189, 190) dengan fleksibilitas rendah, sedangkan lipatan (*loop*) yang mengikat alkohol sekunder (residu 141.144.154.285.289.290) menunjukkan fleksibilitas yang lebih tinggi.



Gambar 3. Struktur 3D protein CALB

Analisis profil skala asam amino (ProtScale)

Untuk mengetahui presentasi asam amino, berat molekul, isoelectric point (pI) dan sifat fisiko kimia lain dari protein CALB maka dilakukan analisis pada sekuens asam amino CALB dengan menggunakan Protparam Expsy:

Tabel 1. Sifat fisika dan kimia protein CALB

Parameter	Protein CALB		
Berat Molekul	35517,57		
pH isoelektrik	8,12		
Komposisi Asam Amino	ala (A)	40	11.7%
	Cys (C)	7	2.0%
	Gly (G)	28	8,2%
	Thr (T)	30	8,8%
Komposisi Atom	Carbon	C	1584
	Hydrogen	H	2507
	Nitrogen	N	419
	Oxygen	O	482
	Sulfur	S	12
	C ₁₅₈₄ H ₂₅₀₇ N ₄₁₉ O ₄₈₂ S ₁₂		
Total jumlah atoms: 5004			
Index aliphatic	88,19		
<i>Grand average of hydropathycity</i> (GRAVY)	0,121		

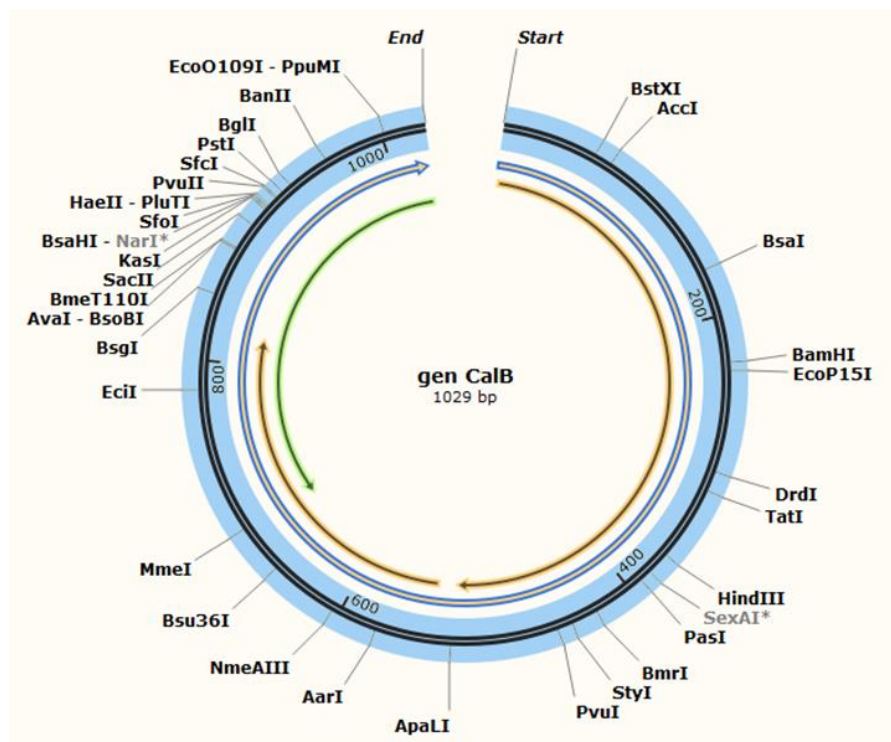
Desain Primer

Desain primer merupakan bagian dari bioinformatika yang menjadi faktor yang paling penting dalam menentukan sekuen DNA yang belum diketahui. Desain primer menggunakan software "Primer3" kemudian dikonfirmasi dengan BLAST NCBI untuk melihat spesifisitas primer. Untuk melihat struktur sekunder (hairpin loop, dimer) yang dihasilkan oleh primer dilakukan analisis dengan menggunakan PerlPrimer. Dilakukan desain primer dari sekuen yang dipilih dan dimasukkan ke dalam program primer3 output, dan diperoleh beberapa alternatif desain primer sebagai berikut:

Dari hasil analisis primer, yang memenuhi syarat adalah oligo yang keempat karena memiliki sekuen yang similaritasnya tinggi dengan ukuran produk 232 pasang basa. Pemilihan kandidat oligo 4 selain menghasilkan produk yang lebih panjang dari kandidat lainnya, primer ini sudah memenuhi kriteria dalam mendesain sebuah primer yang baik. Primer yang unik untuk urutan target yang akan diamplifikasi harus memenuhi kriteria tertentu seperti panjang primer, GC%, suhu *annealing* dan nilai *temperature melting* (tm), stabilitas akhir 5', spesifisitas ujung 3' (Dieffenbach et al., 1995 dan Syamsurizal, 2018). Panjang basa primer sebesar 20 pasang basa dan memiliki nilai tm antara forward dan reverse yang tidak jauh berbeda. Nilai tm (*temperature Melting*) yang optimal untuk primer yang baik dalam kisaran 52-60 °C, umumnya menghasilkan hasil yang lebih baik daripada primer dengan suhu leleh yang lebih rendah. Primer dengan tm di atas 65°C juga harus dihindari karena potensi terjadinya penempelan primer dua kali pada cetakan DNA (*secondary annealing*) (Elsalam, 2003).

Analisis Situs Restriksi

Untuk mengetahui situs restriksi yang terdapat pada gen *CaB* maka dilakukan analisis dengan menggunakan software *SnapGene*. Tujuan diketahuinya situs restriksi tersebut agar gen dapat dipotong dengan salah satu enzim restriksi endonuklease yang diinginkan. Hasil analisis disajikan pada Gambar 5.



Gambar 5. Pemetaan Enzim Restriksi pada gen *caB* dengan software *SnapGene*

4. KESIMPULAN

Gen sintetik CalB yang telah dibuat mempunyai ukuran 1083 pasang basa dan 342 asam amino. Protein CALB mempunyai signal peptide dengan situs pemotongan ada di posisi 19-20. Primer yang digunakan yaitu CalB forward 5'-TCCCCAGTATCAGGTCCAAG-3' dan CalB reverse 5'-GACACCTGAGGCTGAACGAT-3'.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada Ristekdikti yang telah memberikan dana untuk penelitian ini melalui Hibah Penelitian Dosen Pemula untuk anggaran tahun 2018.

DAFTAR PUSTAKA

- Ausubel FM, Brent R, Kingston RE, Moore DD, Seidman JG, Smith JA, Sruhl K. 2002. *Short Protocol in Molecular Biology*. 5th Edition. New York (US): John Wiley & Sons Inc.
- Blank K, Morfill J, Gump H, Gaub HE. 2006. Functional expression of *Candida antarctica* lipase B in *Eschericia coli*. *J of Biotechnol*. 125:474-483.
- Claverie JM. dan Notredame C. 2006. *Bioinformatics for Dummies*. New Jersey: John Willey and Sons.
- Dieffenbach CW, Lowe TMJ, Dveksler GS. 1995. General Concepts for PCR Primer Design. In: PCR Primer, A Laboratory Manual, Dieffenbach CW, Dveksler GS Ed., Cold Spring Harbor Laboratory Press, New York, 133-155.
- Elsalam KA. 2003. Bioinformatic tools and guideline for PCR primer design. *African Journal of Biotechnology*, 2 (5): 91-95.
- Fernandez L, Beerthuyzen, MM, Brown J, Siezen RJ, Coolbear T, Holland R, Kuipers OP 2000. Cloning, characterization, Controlled overexpression and inactivation of the major tributyrin esterase gene of *Lactococcus lactis*. *Appl. Environ. Microbial.*, 66 : 1360-1368.
- Gunasekaran V, Das D. 2005. Lipase Fermentation: Progress and Prospect. *Indian J of Biotechnol*. 4: 437-445.
- Gupta. 2004. Bacterial Lipases: an overview of production, purification and biochemical properties. *Appl Microbial biotechnol*, 64:763-782.
- Hasan F, Shah AA, Hameed A. 2006. Industrial Aplications of Microbial Lipase. *Enzyme and Microbial Technol*, 39 (2): 235-251.
- Joseph B, Ramteke PW dan Thomas G. 2008. Cold active microbial lipases: some hot issues and recent developments. *Biotechnology Advances*, 26 (5): 457-470.
- Mount DW. 2001. *Bioinformatics: Sequence and Genome Analysis*, Cold Spring Harbor: Cold Spring Harbor Laboratory Press.
- Nardini M, Dijkstra BW. 1999. α/β hydrolase fold enzymes : the family keeps growing. *Current opinion in Structural Biology*, 9 : 732-737.
- Narita V, Arum AL, Isnaeni S, Fawzya NY. 2012. Analisis Bioinformatika Berbasis WEB untuk eksplorasi Enzim Kitonase Berdasarkan Kemiripan Sekuens. *Jurnal Al-Azhar Indonesia Seri Sains dan Teknologi*, 1 (4): 197-203.

- Pfeffer J, Richter S, Nieveler J, Hansen CE, Rhild RB, Schmid RD, Rusnak M. 2006. High yield expression of lipase A from *Candida antarctica* in the methylotrophic yeast *Pichia pastoris* and its purification and characterisation. *Appl Microbiol and Biotechnol*, 72: 931-938.
- Pleiss, J, Fischer M. Schmid RD. 1998. Anatomy of lipase binding sites : the scissile fatty acid binding site. *Chemistry and Physics of Lipids*, 93: 67-80.
- Saraswati H. 2017. Analisa Bioinformatika Gen E1 dan E2 dari Virus Hepatitis C (HCV) Genotipe 1, 2, 3, dan 6 sebagai Kandidat Vaksin *Viral-Like Particles* (VLP). *Indonesian Journal of Biotechnology and Biodiversity*, 2 (2): 48-57.
- Sharma, Rohit. *et al.* 2001. Production, Purification, Characterization, and Application of Lipases. *Biotechnology Advances*. 19 : 627-662
- Syamsurizal, & Kadri, H. (2018). *Genotyping SNP Rs12255372 TCF7L2 Gene Using Three-Primer ARMS-PCR for Detection T2DM n Indonesian Batak Ethnic*. Paper presented at the International Conference on Mathematics and Natural Sciences (IConMNS), Bali, Indonesia.
- Ulfah, M. 2013. Subkloning gen sintetik *Candida antarctica* Lipase B (*CalBsyn*) dan Fusi Gen *CalBsyn-egfp* ke dalam Vektor Ekspresi pGAPZ α pada *Eschericia coli* TOP10F'. Skripsi. Jakarta: Universitas Indonesia.
- Vadhana AK1, Samuel P, Berin RM, Krishna J, Kamatchi K, Meenakshisundaram S. 2013. Improved secretion of *Candida antarctica* lipase B with its native signal peptide in *Pichia pastoris*. *Enzyme Microb Technol*. 5 (3): 177-83.
- Wahyuni, FD. 2016. Konstruksi Vektor dan Ekspresi Protein Rekombinan Lipase *Candida antarctica* (calB) dengan Cellulose Binding Module (Cbm) pada *Pichia pastoris*. Tesis. Bogor: Institut Pertanian Bogor.
- Wahyuni, FD., Fuad, AM. 2016. Constitutive Expression of *Candida antarctica* Lipase B (CALB) in *Pichia pastoris* Using pGAPZ α Vector. *Annales Bogorienses*. 20 (1): 29-36.